Shape-correlated Statistical Modeling and Analysis for Respiratory Motion Estimation

Xiaoxiao Liu

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science.

Chapel Hill
2011

Approved by:

Stephen M. Pizer, Advisor

Marc Niethammer, Co-principal Reader

J.S. Marron, Reader

Dinggang Shen, Reader

David S. Lalush, Reader

Bradley C. Davis, Reader

Abstract

Xiaoxiao Liu: Shape-correlated Statistical Modeling and Analysis for Respiratory Motion Estimation.
(Under the direction of Stephen M. Pizer.)

Respiratory motion challenges image-guided radiation therapy (IGRT) with location uncertainties of important anatomical structures in the thorax. Effective and accurate respiration estimation is crucial to account for the motion effects on the radiation dose to tumors and organs at risk. Moreover, serious image artifacts present in treatment-guidance images such as 4D cone-beam CT cause difficulties in identifying spatial variations. Commonly used nonlinear dense image matching methods easily fail in regions where artifacts interfere.

Learning-based linear motion modeling techniques have the advantage of incorporating prior knowledge for robust motion estimation. In this research shape-correlation deformation statistics (SCDS) capture strong correlations between the shape of the lung and the dense deformation field under breathing. Dimension reduction and linear regression techniques are used to extract the correlation statistics. Based on the assumption that the deformation correlations are consistent between planning and treatment time, patient-specific SCDS trained from a 4D planning image sequence is used to predict the respiratory motion in the patient's artifact-laden 4D treatment image sequence.

Furthermore, a prediction-driven atlas formation method is developed to weaken the consistency assumption, by integrating intensity information from the target images and the SCDS predictions into a common optimization framework. The strategy of balancing between the prediction constraints and the intensity-matching forces makes the method less sensitive to variation in the correlation and utilizes intensity information besides the lung boundaries. This strategy thus provides improved motion estimation accuracy and robustness.

The SCDS-based methods are shown to be effective in modeling and estimating respiratory motion in lung, with evaluations and comparisons carried out on both simulated images and patient images.

In honor of

my parents & Li

Acknowledgments

Five years of graduate school at UNC-CH has been an incredible journey for me. So many people have helped and inspired me both in research and in life.

I would mostly like to thank my adviser Steve Pizer, who taught me scientific principles in the medical image analysis field, helped me to achieve all the doctoral milestones and inspired me with his outstanding passion and dedication in research. It has been a privilege to be one of his doctoral advisees. Steve heavily influenced my dissertation direction and initiated the research collaboration with Memorial Sloan-Kettering Cancer Center. A special thanks goes to Gig Mageras, whose expertise and dedication made this collaboration smooth and fruitful.

My other committee members, Marc Niethammer, Brad Davis, Steve Marron, David Lalush and Dinggang Shen all have been very constructive and helpful in numerous discussions regarding my dissertation research. Their research expertise and critical thinking were extremely valuable and helped me in clarifying specific research aims at different stages of the research.

The entire MIDAG group has been supportive and helpful over those years. Josh Levy, Ja-Yeon Jeong, Joshua Stough, Beatriz Paniagua, Sungkyu Jung, Qiong Han, Dereck Merck, Gregg Tracton, Graham Gash, Rohit Saboo, Mark Foskey, Sha Chang and Brandon Frederick have helped me in different projects. In particular, I want to thank Ipek Oguz for coaching me in adopting her correspondence methodology in my dissertation work. Last but not least, my office mate for the past two years, Ilknur Kabul, has been a very good companion at work. Positive energy in the office definitely helped me to handle dissertation writing stress.

In the end I would like to express my gratefulness to my family. My parents have always been unconditionally supportive and boundlessly patient for what I choose to accomplish. My loving husband Li has been the greatest support and my inspiration along this journey.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CBCT	cone-beam CT	6
COG	center of gravity	48
DVF	displacement vector field	41
EE	end expiration	5
EI	end inspiration	7
$HDLSS$	high dimensional low sample size	6
$IGRT$	image-guided radiation therapy	2
LDD	large diffeomorphic deformation	18
$LDDMM$	large diffeomorphic deformation mapping metric	21
$M\text{-}rep$	medial representation	32
NST	nanotube stationary tomosynthesis	7
$NCAT$	nurbs-based cardiac-torso	61
PCA	principal component analysis	8
PDF	probability density function	23
PDM	point distribution model	5
$RCCT$	respiration-correlated CT	6
$RC\text{-}CBCT$	respiration-correlated cone-beam CT	6
$SCDS$	shape-correlated deformation statistics	4

List of Symbols

I	planning or training images	37
J	treatment or test images	59
v_t	time-varying flow of velocity fields	19
L	differential operator	19
K	Green's function operator	19
$\mathbf{u}$	displacement vector fields	41
ϕ	diffeomorphic transformations among training images	19
φ	diffeomorphic transformations among test images	63
p	a surface point set extracted from a training image	41
q	a surface point set extracted from a test image	59
$\mathbb{R}$	the real numbers	19
Ω	a subset of $\mathbb{R}^3$ wherein images and deformations are defined	19
e_i	the ith eigenvector	25
x	3D coordinates of a voxel	19
O	tumor (or region of interest) contours	63
σ	a weighting factor between terms in image matching functional	21
∇^2	Laplacian operator	19
∇	gradient operator	19
$\nabla\cdot$	divergence operator	19

Chapter 1

Introduction

1.1 Motivation

Lung cancer, the leading lethal cancer type, causes 1.3 million deaths each year worldwide (WHO, 2006). More than half of those diagnosed with non-small cell lung cancer will receive radiation therapy during their treatments. The goal for any cancer radiation therapy is to attain the highest probability of cure with the least morbidity. To increase this therapeutic ratio with radiation is to expose all cancer cells with sufficient doses of radiation while sparing surrounding healthy tissues to reduce the probability of complications.

Over the past decade enormous progress has been made in improving radiation therapy techniques. On the one hand, the application of radiography (x-ray imaging), computed tomography, magnetic resonance imaging, ultrasound, positron emission tomography, and electronic portal imaging have greatly improved the possibility of identifying tumors and organs at risk. On the other hand, high-precision dose delivery systems, such as intensity-modulated radiation therapy (IMRT), which utilizes computer-controlled linear accelerators to deliver precise radiation, enable more and more accurate dose distributions.

Accurately identifying and localizing the target objects are crucial. In image-guided

radiation therapy (IGRT) applications, the intended area for treatment is usually outlined on a set of diagnostic images by radiation oncologists at the time of *planning*. Another set of images is then acquired during the time of *treatment*, immediately prior to the radiation delivery, to adjust the patient position with respect to the planning data set. Achieving spatial accuracy of the radiation dose relies on achieving a close match between the anatomic evidence from the planning and treatment images, respectively.

For target areas such as the head and neck or the prostate, which are static during each imaging, the misalignment between the *planning* and the *treatment* images can be removed by image alignment or registration. For the thorax and abdomen regions, respiratory motion challenges the standard treatment planning procedure. Sizable errors in organ volume, position, and shape can occur during breathing, which can significantly affect the calculation of dose-volume histograms (Balter et al., 1996). In fact, positional uncertainties caused by the respiratory motion have been shown to have a large impact on radiation dose (Minohara et al., 2000; Keall et al., 2006). Efficient respiratory motion modeling and analysis is necessary for removing the motion-induced uncertainties.

The advances in 4D imaging, with *respiration phase* as the fourth dimension, enables capturing and storing the motion in the form of a phase-stamped 3D image sequence. In this thesis a 4D image sequence refers to a respiration-correlated image sequence. Each 3D image volume is respiratory-phase-stamped. The phase is determined by a gating device that monitors the breathing with measurements used for sorting the projection data for 3D reconstruction (see Section 2.2.2 for details).

How to accurately and efficiently identify objects of interest and track their spatial changes for IGRT with respiratory motion is the driving problem for this thesis. From the anatomy and the physics points of views, strictly repeatable breathing patterns are unrealistic assumptions even for healthy patients, due to the hysteresis effects, i.e., different trajectories during inhalation and exhalation. Regular breathing varies from cycle to cycle with different amounts of contraction and expansion. Thus, although the

breathing motion is periodic, the conformation of objects in the thorax does not remain associated with a fixed time point in the breathing cycle. Respiratory motions could be approximated by periodic functions of surrogate signals measured by monitoring devices during the breathing cycle. Still, for patients with irregular breathing a periodic function is inadequate to describe the internal organ motion (Vedam et al., 2003; Neicu et al., 2003; Nehmeh et al., 2004).

Recognizing the potential benefits of respiratory motion models that do not assume reproducible respiration patterns, various motion models have been developed for characterizations of breathing motions from 4D image data (Low et al., 2005; Yang et al., 2008; Zeng et al., 2007; Dirk et al., 2003; Zhang et al., 2007). A greater challenge for the respiratory motion estimation arises when various imaging artifacts are present. For example, severe streak imaging artifacts appear in cone-beam CT images due to the limited angle reconstruction at each respiratory phase. The reconstruction artifacts can lead to uncertainties in computing the motion trajectory from those images.

In situations where the image intensities are not very informative, a learning-based approach is a natural choice for introducing proper constraints and useful knowledge. Especially for medical image analysis where sophisticated geometry and mobility reside in various anatomical structures, statistical models extracted from training samples are extremely valuable and often necessary. However, considering both anatomical and pathological variations in terms of the breathing patterns, training statistics collected from other patients are significantly less useful than patient-specific data. For each patient the CT image sequence acquired prior to the treatment offers good contrast and resolution for the purposes of diagnosis and planning. It therefore can be used as the training data for predicting the motion during treatment time for the same patient. The underlying assumption for the learning-based approach is that the breathing pattern extracted for a specific patient is *consistent* between the *planning* and the *treatment* time. The consistency emphasizes the stability of the mobility for the anatomical structures.

In modeling with surrogate signal, the correlation between the surrogate signal and the image deformation is assumed to be consistent between planning time and treatment time.

This dissertation proposes a statistical shape-correlated deformation system for patient-specific respiratory motion modeling and estimation in the thorax for 4D image-guided radiotherapy. There are three primary contributions in this thesis:

1. A shape-correlated deformation statistics (SCDS) framework is developed to model the respiratory motion with three components: a Fréchet mean atlas image, the diffeomorphic trajectories from all the phases to the atlas image, and the linear correlations between the deformations of the shape surrogate and the deformations of the whole image space.

2. The SCDS is applied to predict the respiratory motion from 4D cone-beam CT images with image reconstruction artifacts.

3. A prediction-driven atlas formation method is developed for a robust respiratory motion estimation, by integrating SCDS prediction constraints into the intensity-based Fréchet mean atlas formation framework.

1.1.1 Shape-correlated deformation statistics (SCDS)

Recognizing the hysteresis of respiration and irregular breathing patterns, various auxiliary tracking devices, such as the pencil-beam navigator (Dirk et al., 2003) and the spirometer (Low et al., 2005), have been used to obtain external motion parameters for motion estimation. The diaphragm position of the lung has also been used as an internal motion parameter for modeling (Zhang et al., 2007). These various motion parameters are *surrogates* of the underlying motion: the amount of positional change of the surrogate is correlated to the amount of the geometric change of the moving structures in the image.

In thorax images air-filled lung regions provide high-contrast boundaries that can be accurately detected by image processing techniques. The shape of the lung can therefore be used as an internal surrogate for the respiratory motion. Also, due to lung's large area within the imaging field, the lung boundary can be used to model localized details of the respiratory motion. The idea of the SCDS model is to model the image deformations as a function of the shape surrogate. Statistical analyses of the shape variation space and the image deformation space as well as of the correlations between the two are carried out in SCDS.

Shape modeling of anatomical objects has been studied for decades to fit geometric models to image intensity profiles. In this spatio-temporal problem it is equally critical that the shape models across the breathing cycle have nice geometric correspondence to assure the tightness of the statistics. A state-of-the-art point-based correspondence system is adopted in this thesis to efficiently build the point distribution models (PDMs) from the 4D image sequence with satisfactory group-wise correspondence. This surface shape representation allows the object to have arbitrary topology, making this method widely applicable to various anatomical structures.

To quantify the deformation over the whole image space, non-linear image registrations are often used to compute the trajectories of the motion. A reference image is usually used for contouring the regions of interest and repositioning the patient prior to the treatment. The breathing motion can be characterized by the non-linear transformations that bring those time-stamped images onto the reference image. This reference image should be consistent in different acquisitions. The end expiration (EE) phase is usually used as the reference phase for registration due to its relatively stable repeatability. The total amount of transformations to bring all other images to the EE phase image is much larger than to any of the phases in the middle of inspiration of expiration. Computationally, small deformations are preferred in terms of both accuracy and efficiency. A Fréchet mean image that has the minimized the total amount of transfor-

mations to all phases represents an averaged spatial configuration over the breathing cycle. Note that the Euclidean mean of all the image intensities is inappropriate and leads to blurred edge features. The Fréchet mean (non-linear mean) image is calculated by minimizing the sum of geodesic distance-squared among images on the manifold constructed by diffeomorphic deformations. In the process of generating the Fréchet mean image, the deformations that transform all the phases to the mean image are optimized.

From the statistical modeling point of view, although the correlation in this sophisticated breathing motion is certainly non-linear, a linear approximation is sufficient when the phase-to-phase deformation is not large. The challenge of building SCDS is the high dimensional low sample size (HDLSS) nature of the data representations: both shape and image deformations are characterized by large-dimension (on the order of thousands and millions, respectively) vectors. However, the sample size (number of phases during one breathing cycle) is usually no more than 10, determined by the 4D imaging protocol. Dimension reduction techniques are therefore necessary to avoid prohibitive and sampling-noise-sensitive computations.

1.1.2 Respiratory motion prediction

In 4D IGRT two different imaging modalities are often used for planning and treatment, respectively. Standard multi-slice respiration-correlated CT (RCCT) scans are usually taken for diagnosis and planning, for its good contrast and high resolution and large field of view. Immediately prior to each treatment, another set of images is acquired for radiation guidance and dose calculation. Factors such as the speed of the scan, the imaging dose and the configuration of the treatment devices determine the feasibility of a certain image modality being adopted at treatment time.

Cone-beam CT (CBCT) exploits flat panel technology to integrate the imaging system directly into the treatment accelerator, allowing acquisition of a volumetric image in the treatment position (Jaffray et al., 2002). However, degradation of image qual-

ity (Figure 1.1) during the image reconstruction resulting from sparse projections at each respiratory phase imposes limitations on its IGRT application (Sonke et al., 2005). Nanotube stationary tomosynthesis (NST) technology can provide unprecedented imaging speed and good spatial resolution in the plane perpendicular to the radiation field (Maltz et al., 2009). However, the resulting tomographically reconstructed images have interfering patterns in comparison to the standard CT scans, making the soft tissue appear with low contrast and with blur (Figure 1.2).

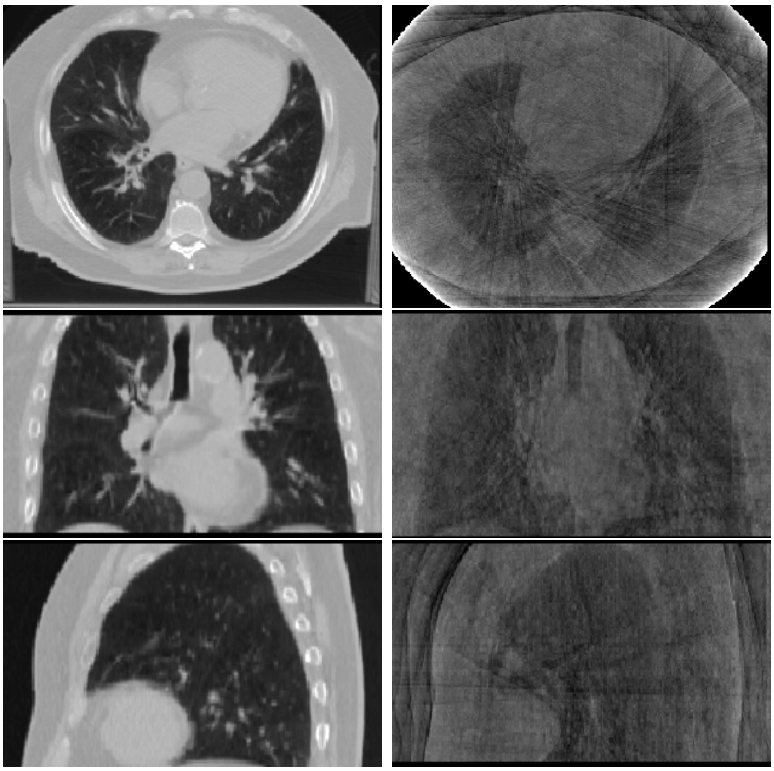


Figure 1.1: **4D Cone-beam CT image artifacts.** Left: The axial, coronal and sagittal views of a diagnostic CT image at the end of inspiration (EI) phase taken at the treatment planning time. Right: The axial, coronal and sagittal views of a one-minute CBCT scan at the EI phase, having significant streak artifacts due to the limited angle reconstruction.

In order to apply the trained correlation to the artifacts-laden images, accurate surrogate extraction is the key. Since there are artifacts interfering with the lung boundaries, simple thresholding is not sufficient for shape extraction. This thesis introduces a learning-based approach to regularize the segmentation against the artifacts. With

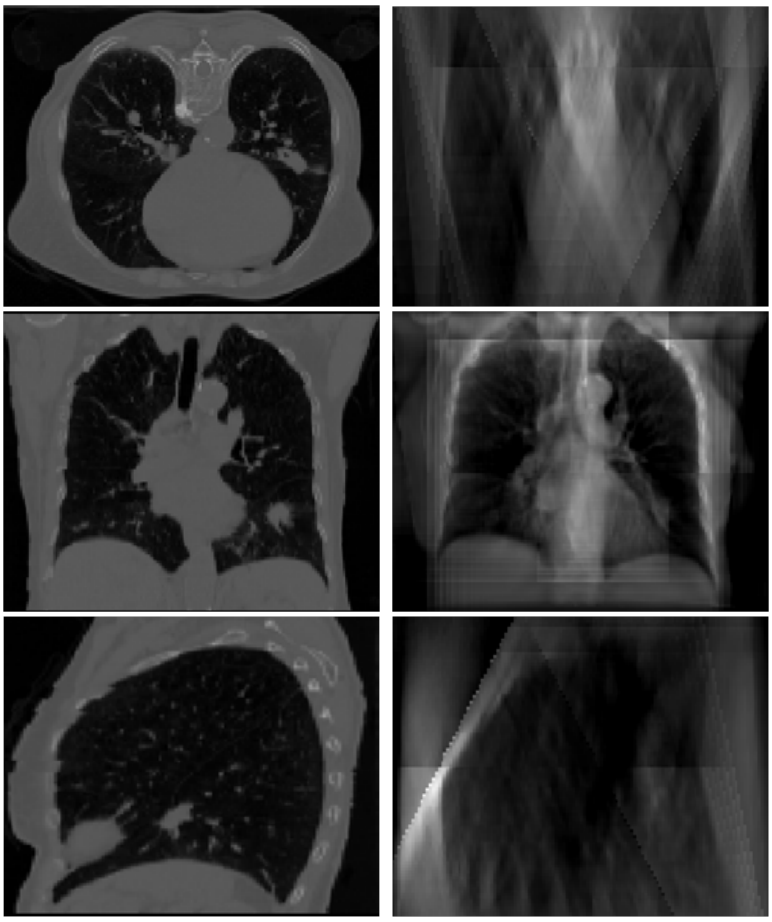


Figure 1.2: **Nanotube Stationary Tomosynthesis (NST) image artifacts**. Left: The axial, coronal and sagittal views of a diagnostic CT scan taken at the end of inspiration (EI) phase taken at the treatment planning time. Right: The axial, coronal and sagittal views of a NST image at EI phase, having interfering blurry patterns due to the limited angle reconstruction.

the planning data available, a shape prior is trained from the segmentations of the RCCTs via principal component analysis (PCA) . And a posterior probability optimization scheme is used to generate the shape models that fit into enhanced edge features in the treatment images while staying within the trained shape space.

1.1.3 Prediction-driven deformation atlas formation

When using the proposed SCDS prediction, the correlation between surrogate deformations and underlying image deformations among the training data is assumed to be the

same as among the target data. This assumption simplifies the complicated breathing system and enables the estimation method of applying the training statistics to target data. However, the assumption is too simplistic for patients who have difficulty in regularizing their breathing. Besides, noise kept in the SCDS trained from HDLSS data might result in predictions that are not consistent with the actual anatomical conformations in some local regions. Facing both problems, improvement of the motion estimation can be made by utilizing intensity features in the target data beyond the shape surrogate for better intensity matching.

A prediction-driven deformation atlas formation method driven by a combination of prediction constraints and the image matching force is developed in this thesis. The deformations predicted from the SCDS model are used as a soft constraint during the optimization. The balancing force between the prediction and the noisy intensity profile can be adjusted via a weighting factor, selected upon the credibility of the training statistics and the quality of the treatment images.

1.2 Thesis and Contributions

Thesis: *To account for respiratory motion effects on the radiation dose to tumors and organs at risk in image-guided radiotherapy for lung cancer, shape-correlated deformation statistics (SCDS) trained from planning CT images can be used to effectively estimate the respiratory motion in artifact-laden treatment images, when breathing patterns are consistent between planning time and treatment time. A more robust motion estimation can be achieved by combining the SCDS-prediction constraints and intensity-matching forces into a unified atlas formation framework.*

The contributions of this dissertation are as follows:

1. *I use the shape of the lung as the surrogate for respiratory motion estimation in the thorax. The high contrast of the air-filled lung regions provides comprehensive*

motion evidence.

2. *I apply an entropy-based particle system to 4D lung CT images to obtain surface lung models with group-wise correspondence. The surface correspondence across the 4D sequence is crucial for statistical shape modeling.*

3. *I apply a Fréchet mean image formation method to generate the respiratory motion atlas from 4D respiratory CT images. The motion atlas contains an atlas image and the transformations that match each individual phase image to the atlas.*

4. *I present shape-correlated deformation statistics (SCDS) to capture correlations between the shape deformation and the dense image deformation fields from a 4D CT sequence.*

5. *I apply patient-specific SCDS trained from planning 4D CT images in predicting the respiratory motion from cone-beam CT images that contain artifacts due to image reconstructions with sparse projections.*

6. *I present a prediction-driven atlas formation method that combines SCDS-prediction constraints and intensity-matching forces into a unified framework for robust respiratory motion estimation.*

1.3 Overview of Chapters

The remainder of this dissertation is organized as follows:

Chapter 2 introduces the background of the thesis work, including the clinical application of 4D image-guided radiotherapy, the respiratory motion mechanism and the large-deformation diffeomorphic image registration method used for characterizing the motion, and the statistics of multivariate Gaussian distributions.

Chapter 3 introduces the core method of the thesis, shape-correlated deformation statistics (SCDS), for approximating a linear correlation between the shape deforma-

tions of the surrogate object and dense image deformations. Techniques used for shape modeling with group-wise correspondence and the motion characterization using the Fréchet mean image are introduced in detail. Various statistical correlation methods are investigated. The effectiveness of the SCDS is demonstrated using planning CT images.

Chapter 4 applies the SCDS model in predicting the breathing motion from 4D cone-beam CT (CBCT) at treatment time. The key for the prediction is an accurate segmentation of the shape from CBCT images that contain severe imaging artifacts. A posterior deformable model segmentation method is presented to account for the imaging artifacts. Experiments are carried out and results are evaluated on both phantom and real patient data.

Chapter 5 presents an prediction-driven atlas formation method for respiration-motion, which integrates the SCDS prediction results into the intensity-based Fréchet mean formation framework. It aims to combine the advantages of the SCDS-prediction method and a pure intensity-based motion estimation method towards more robust estimation results.

Chapter 6 contains a discussion of the contributions of this thesis and an outline of future work.

Chapter 2

Background

2.1 Radiation Therapy: From 3D to 4D

Traditionally, 3D medical images are reconstructed from 2D tomography using various imaging systems. Each of the slices is treated independently and is acquired as a self-contained data set. In efforts to improve the temporal resolution of 3D imaging techniques, the basic imaging element has been changed from 2D slice imaging to 3D slab imaging. There are two different imaging techniques in 3D CT: multi-slice CT (MSCT) and cone beam CT (CBCT). MSCT can simultaneously acquire and reconstruct 16 to 64 slices (Takuya et al., 2006). Anatomical regions can be imaged within a second per gantry rotation. CBCT uses an x-ray source that diverges both longitudinally and laterally, coupled with a flat-panel detector array in order to simultaneously image the entire volume (Endo et al., 2003). Depending on the 3D reconstruction algorithms, some artifacts, such as distortion or aliasing may occur in CBCT images (Köhler et al., 2002).

4D radiation therapy (4DRT) aims to track and compensate for target motion, such as respiratory motion, during radiation treatment. Motion compensation can therefore minimize normal tissue damage, especially critical structures adjacent to the target region, while maximizing radiation dose to the target. The fourth dimension is the time during the imaging process. Involuntary and voluntary patient motion has become a

major obstacle for achieving high-fidelity medical imaging and high precision radiation therapy (HPRT). The common goal of all the advances in intensity-modulated radiation therapy (IMRT), adaptive radiation therapy (ART) and image-guided radiation therapy (IGRT) is to precisely localize the target in space and in time to achieve HPRT (Keall, 2004; Sanghani & Mignano, 2006; Zerda et al., 2007; Xing et al., 2006). With the progress in image-guided radiation delivery systems, such as IMRT, on-board CBCT, helical tomotherapy and robotic linear accelerators, 4D radiation therapy, although still in the research phase, is promising to improve the therapeutic ratio.

2.2 Imaging of Respiratory Motion

2.2.1 Breathing mechanism

Human respiration is a process of exchanging air between the lung and the environment. It involves the motion of the diaphragm, heart, liver, stomach, spleen, thoracic cage, and lungs (Segars, 2001; Siebenthal, 2008). The pressure within the lungs and their volume is changed by the motion of the diaphragm and the ribsfor illustration of the anatomy in thorax). During exhalation, the diaphragm contracts to force the abdominal contents downward and forward and thereby increases the volume of the thorax. During inhalation the diaphragm relaxes and the abdominal contents move upward and inward and thus decrease the volume of the lung. The level of the diaphragm can move up and down from 1 to 10 cm during different levels of breathing (Fig.2.1) (Segars, 2001). The diaphragm moves about 1-2 cm during normal tidal breathing (West 1995). During respiration the lungs inflate and deflate with changes in the volume of the thoracic cavity. Besides the diaphragm, the ribs rotate about an axis through their costal necks to affect the anteroposterior and transverse diameters of the thoracic cavity (Fig.2.2).

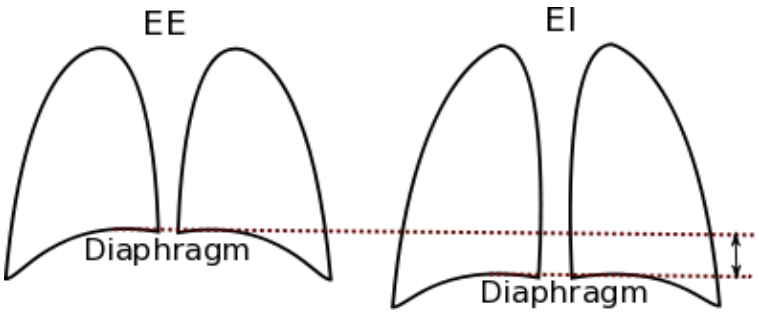


Figure 2.1: **Motion of the diaphragm during respiration.** During inspiration the diaphragm contracts, increasing the volume of the thoracic cavity. During expiration the diaphragm relaxes, decreasing the volume of the thoracic cavity. At the end of expiration (EE) phase (left), the thoracic volume decreases to the minimum, and at the end of inspiration (EI) phase (right), the lung volume increases to the maximum.

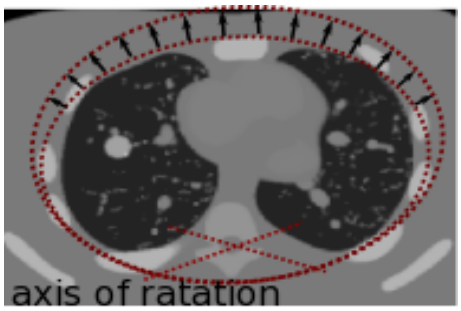


Figure 2.2: **Motion of the ribs during respiration.** During inspiration the external intercostal muscle contracts, pulling the ribs upward and forward increasing the transverse and anteroposterior diameters of the thorax.

Respiratory motion often has Irregular amplitude, frequency, and shape of the breathing trajectory (Siebenthal, 2008). Frequent causes of irregular motion are changes in the tension of the muscles involved in respiration, but also emotional changes, sighing, or coughing. For such irregular motion, which is often observed during free breathing, a sinusoidal function (Lujan et al., 2003) is a very coarse approximation.

2.2.2 Retrospective imaging

Two different interventional strategies have evolved to reduce the effect of respiratory motion in radiation treatments: controlled patient breathing and respiration gating of the accelerator while the patient breathes normally. In the former approach, breathing is

altered either voluntarily by instructing the patient or assisted by means of an occlusion valve . In the latter approach, a device monitors patient breathing and allows delivery of radiation only during certain time intervals, synchronous with the patient's respiratory cycle. However, roughly one-third to one-half of eligible patients could not perform deep inspiration breath-hold (DIBH) technique satisfactorily, and average session times for simulation and treatment of the initial patients were nearly double that for free-breathing treatment .

There are two categories of methods to carry out 4D imaging in free-breathing respiratory motion. One is prospective imaging, which obtains image projections using respiratory-gated image acquisition in reference to an external motion indicator, producing a single "motion-free" 3D image at the selected respiratory phase. The other, which is more true to the name "4D", is retrospective imaging, which acquires image projections from all respiratory phases and sorts them into appropriate phase bins based on the external motion indicator, producing a series of motion-free 3D images at different phases of the breathing cycle (Li et al., 2008). The fourth dimension refers to the respiration phase. The 4D data used in this work are all respiration-correlated data.

Each 3D volume of the 4D data is reconstructed at different temporal phases of a breathing cycle sorted according to a certain gating device, such as spirometry measurement (Lu et al., 2005), elastic belt amplitude (Guckenberger et al., 2007), video respiration monitor (Mageras et al., 2004), etc. Figure 2.3 illustrates the sorting scheme for RCCT reconstruction.

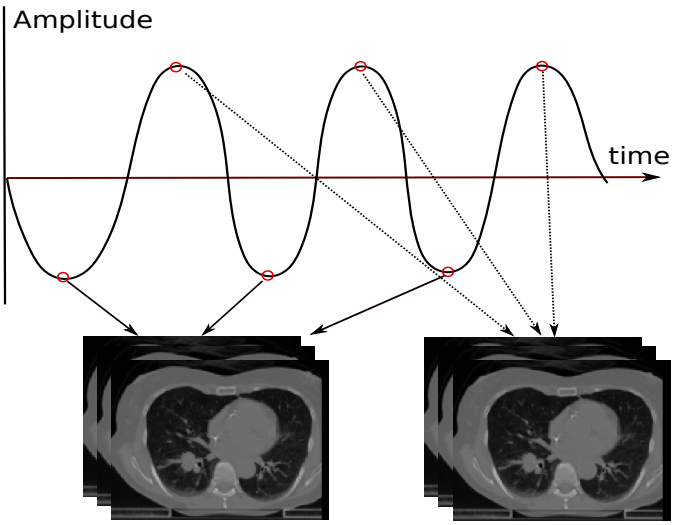


Figure 2.3: **Respiration-correlated CT reconstruction diagram.** The respiration-correlated CT reconstructed by retrospective sorting the amplitude of a sensor signal: the image projections acquired at a corresponding phase of the breathing cycle are used for a 3D reconstruction of that phase/time-point.

2.2.3 Respiration-correlated CBCT

The advent of large area flat panel x-ray detectors has led to the development of kilo-voltage image guidance systems in the treatment room (Balter et al., 2002; Hammoud et al., 2005). Several commercial in-room CT systems have become available, providing volumetric patient images for soft tissue identification (Barker et al., 2004; Uematsu et al., 2000). Recent developments in cone-beam CT mounted on the linear accelerator makes it possible to acquire patient images in the 1treatment room just prior to treatment (Jaffray et al., 2002; Oldham et al., 2005; Chang et al., 2006). CBCT potentially provides sufficient image quality for evaluating and improving the accuracy of patient treatment.

As with conventional CT, 3D CBCT is heavily influenced by respiratory motion. Respiration-correlated CBCT reduces respiratory motion artifacts. However, due to the significantly reduced number of CB projections used to reconstruct each 3D CBCT image, the resulting image quality is often degraded by view-aliasing artifacts easily seen in

the axial view. Acquisitions using multiple gantry rotations or slow gantry rotation can increase the number of projections and substantially improve the 4D images. However, the extra cost of the scan time may set fundamental limits to their applications in clinics. Motion-compensated CB reconstruction techniques have been proposed, enabling the use of all the acquired projections (Li et al., 2007; Rit et al., 2008). These methods require prior knowledge of the motion present during CB acquisition.

The gated radiotherapy system used for acquiring RC-CBCTs used in this thesis is the Real-Time Position Management (RPM) Respiratory Gating System, Varian Medical Systems (see Figure 2.4). It permits breathing-synchronized fluoroscopy on a treatment simulator, as well as gated treatment on a linear accelerator. The system uses a breathing monitor, consisting of a video camera and passive infrared reflective markers placed on the patients thorax, to synchronize radiation from a linear accelerator with the patients breathing cycle (Mageras et al., 2001).

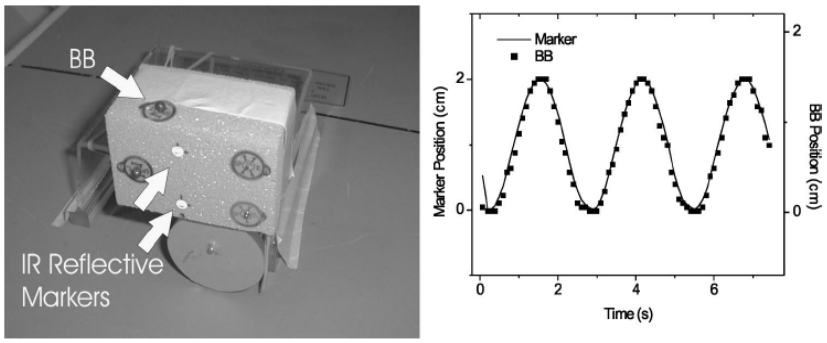


Figure 2.4: **RPM Respiratory gating system.** a) Mechanical motion phantom and reflective marker block. b) Comparison of the marker position waveform vs. time from the gating system camera, with the BB position observed from fluoroscopic images. (Image courtesy of Varian Medical Systems).

2.3 Image Registration for Motion Description

4D image registration, a series of 3D deformable image registrations, has been applied to capture the trajectories of image volumes in space from phase to phase within the

breathing cycle. Various 3D deformable image registration algorithms have been developed for matching anatomical structures, such as b-spline-based image registration (Rueckert et al., 2006), free-form deformation (Lu et al., 2004), optical flow based registration (Naqa et al., 2003), non-parametric Demons-based image registration (Vercauteren et al., 2007), the viscous fluid model (Foskey et al., 2005), etc. To handle relatively large deformations while maintaining spatial smoothness, large-deformation diffeomorphic (LDD) registration methods are desirable. In this thesis fluid-based LDD image registration is used to describe the 4D motion of the lung (chapter 3). The computation of the atlas image within a 4D sequence is built on top of the LDD registration scheme (chapter 5).

2.3.1 Large-deformation diffeomorphic image registration

For medical images diffeomorphic deformation of the image coordinates is used to describe the geometric deformation of the anatomical structures represented by the image intensities. The analysis of image deformations detects and quantifies shape changes or geometric differences in the objects of interest. The algebraic and differential geometric structure of diffeomorphisms provides a metric that has a well-defined notion of "the amount of geometric change" (Davis, 2008).

Large-deformation diffeomorphic image registration has originated in the pattern theory field (Grenander, 1996) and has been actively developed over the past decade by Christensen et al. (1996), Dupuis et al. (1998), Grenander & Miller (1998), Miller & Younes (2001), Beg et al. (2005), Miller et al. (2006), Hart et al. (2009) and more. The analysis of image deformations provides insight into shape changes or geometric differences in the underlying geometric structures. The image registration framework and its mathematical notions that are used later in Chapter 3 & 5 are introduced here in detail following the conventions in Davis's PhD thesis (Davis, 2008).

Manifold structures of LDD

Images are modeled as real-valued L^2 functions on the domain $\Omega \subset \mathbb{R}^3$. Spatial transformations are used to deform images by deforming the underlying coordinate space of Ω. These transformations ϕ are elements of a subgroup of diffeomorphisms $\mathrm{Diff}_V(\Omega)$, $\phi : \Omega \to \Omega$. They are generated by flows of smooth, and time-varying velocity fields, which enable LDD transformations while maintaing the diffeomorphic property.

These velocity flows v_t with a simulated time parameter $t \in [0, 1]$, are generated from velocity fields that are elements of a Hilbert space V with associated inner product $< , >_V$. For $u, v \in V$, this inner product is defined using a linear differential operator L (with associated adjoint $L^\dagger$):

$$< u, v >_V \equiv < Lu, Lv >_{L^2} = < L^\dagger Lu, v >_{L^2} = \int_\Omega < L^\dagger Lu(x), v(x) >_{E^3} dx \quad (2.1)$$

where $< \cdot, \cdot >_{E^3}$ is the Euclidean inner product. This inner product on velocity fields induces the norm $\|v\|_V \equiv \sqrt{< v, v >_V}$. The form of the differential operator L is taken from fluid mechanics (Christensen et al., 1996; Dupuis et al., 1998) to be

$$L = \alpha\nabla^2 + \beta(\nabla\cdot)\nabla + \gamma, \quad (2.2)$$

where the parameters α and β control the viscous properties of the deforming medium and γ ensures that L is invertible.

The operator L is associated with the compact self-adjoint operator K by $< u, v >_{L^2} = < Ku, v >_V$, which implies that $u = KL^\dagger Lu$. The flow v_t is related to the diffeomorphism ϕ via the Lagrangian ODE

$$\frac{d}{dt}\phi_t(x) \equiv \dot{\phi}_t = v_t(\phi_t(x)). \quad (2.3)$$

In particular, ϕ is generated from v_t according to

$$\phi_t(x) = x + \int_0^t v_t \circ \phi_t(x) dt, \tag{2.4}$$

subject to

$$\phi_0(x) = x, \qquad \phi(x) = \phi_1(x), \qquad \dot{\phi}_t(x) = v_t \circ \phi_t(x), \tag{2.5}$$

for all $x \in \Omega$.

A differentiable manifold structure is defined for $\mathrm{Diff}_V(\Omega)$, where V is the tangent space at the identity $Id_{\mathrm{Diff}_V(\Omega)}$. The combination of group structure and differentiable structure allows $\mathrm{Diff}_V(\Omega)$ to behave very much like a Lie group. In particular, a right-invariant Riemannian distance is defined on $\mathrm{Diff}_V(\Omega)$ based on $< , >_V$ at the identity:

$$d_{\mathrm{Diff}_V(\Omega)}(Id_{\mathrm{Diff}_V(\Omega)}, \phi) = \inf_{v:\dot{\phi}_t = v_t(\phi_t)} \int_0^1 \sqrt{< v_t, v_t >_V} dt = \inf_{v:\dot{\phi}_t = v_t(\phi_t)} \int_0^1 \|v_t\|_V dt, \tag{2.6}$$

subject to

$$\phi(x) = x + \int_0^1 v_t \circ \phi_t(x) dt \quad for\ all\ x \in \Omega. \tag{2.7}$$

The distance between any two elements of $\mathrm{Diff}_V(\Omega)$is defined by

$$d_{\mathrm{Diff}_V(\Omega)}(\phi_1, \phi_2) = d_{\mathrm{Diff}_V(\Omega)}(Id_{\mathrm{Diff}_V(\Omega)}, \phi_2 \circ \phi_1^{-1}). \tag{2.8}$$

With this structure the length of curves can be measured along the manifold $\mathrm{Diff}_V(\Omega)$. This distance provides a metric space structure for $\mathrm{Diff}_V(\Omega)$.

LDD image registration

The LDD metric described above can be applied to computing transformations that warp one image to another. Intuitively, the distance between two images is given by the amount of deformation required for one image to match the other. Consider a diffeomorphism $\phi \in \mathrm{Diff}_V(\Omega)$, the action of ϕ on an image $I \in \mathcal{I}$ is defined by

$$I_\phi \equiv I \circ \phi^{-1}. \tag{2.9}$$

Given fixed and moving (to-be-deformed) images, I^F and I^M in $\mathcal{I}$, the image registration goal is to generate a deformation ϕ that best aligns I^M_ϕ with I^F. It is common to allow some mismatching between the image pair, considering exact matching too ideal for most real applications. An image dissimilarity metric can be defined based on image modalities and image noise models. For the the same image modality, the L^2 norm is often used for its simplicity:

$$d^2(I^F, I^M) = \inf_{v:\dot{\phi}_t = v_t(\phi_t)} \int_0^1 \|v_t\|_V^2 dt + \frac{1}{\sigma^2}\|I^M_\phi - I^F\|^2_{L^2} \tag{2.10}$$

where σ controls the relative weight of the two terms. Small values of σ increase the importance of the image dissimilarity metric, forcing the images to match as well as possible; large values of σ produce deformations that require less "energy" according the metric on $\mathrm{Diff}_V(\Omega)$.

Greedy solution for the image registration

Beg et al. (2005) derived the LDD Metric Mapping (LDDMM) solution for Equation (2.10) using the Euler-Lagrange equations for the energy functional:

$$2v_t - K(\frac{2}{\sigma^2}|D\phi_{t,1}|\nabla I^M_{\phi_{0,t}} - I^F_{\phi_{1,t}}) = 0, \tag{2.11}$$

where $D\phi_{t,1}$ is the 3×3 Jacobian matrix of the transformation $\phi_{t,1} \equiv \phi_1 \circ \phi_t^{-1}$. For any particular time point $t \in [0, 1]$ the gradient of the energy functional is

$$\nabla_{v_t} E_t = 2v_t - K(\frac{2}{\sigma^2}|D\phi_{t,1}|\nabla I^M_{\phi_{0,t}} - I^F_{\phi_{1,t}}) \tag{2.12}$$

Christensen et al. (1996) proposed a greedy solution to equation (2.10). This solution separates the time dimension of the problem from the space dimensions. At each iteration, a new velocity is computed that optimizes the functional in (2.10) given that the current deformation is fixed (i.e., the past velocity fields are fixed). Unlike the LDDMM approach, this optimization does not update velocity fields once they are first estimated; nor does it take future velocity fields into account. Using a step-size ϵ, these velocity fields are integrated to produce the final deformation. In this case the gradient is

$$v_t = K(\frac{2}{\sigma^2}\nabla I^M_{\phi_{0,t}} - I^F_{\phi_{1,t}}) \tag{2.13}$$

From the view points of implementation, computation storage and speed, the greedy solution is often preferred to LDDMM solution.

2.4 Statistics of Multivariate Gaussian Distribution

In probability theory and statistics, the multivariate normal distribution or multivariate Gaussian distribution is a generalization of the one-dimensional (univariate) normal distribution to higher dimensions. A random vector is said to be multivariate normally distributed if every linear combination of its components has a univariate normal distribution. A vector of n dimensions $\mathbf{x}$ drawn from the multivariate Gaussian distribution

is expressed by

$$\mathbf{x} \sim N(\mu, \Sigma), \tag{2.14}$$

whereμ is the mean and Σ is the non-singular covariance matrix. The probability density function (PDF) for this distribution is

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\Sigma|^{1/2}} \exp(-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)). \tag{2.15}$$

The Mahalanobis distance, which is denoted by d_{mahal}, can be calculated by using the logarithm of the PDF above as follows:

$$-\log(p(\mathbf{x})) = \log((2\pi)^{k/2}|\Sigma|^{1/2}) + \frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu) \tag{2.16}$$

$$d_{mahal}(\mathbf{x}) = \sqrt{(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)} \tag{2.17}$$

$$\begin{aligned} d^2_{mahal}(\mathbf{x}) &= (\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu) \\ &= 2(-\log(p(\mathbf{x}))) + Constant \end{aligned} \tag{2.18}$$

The Mahalanobis distance is a statistical measurement of the distance from a sample to a population mean. It accounts for the anisotropy and orientation of the population's distribution. It also has a straightforward interpretation in applications of principal component analysis (PCA) reviewed in the following.

2.4.1 Principal component analysis

Principal component analysis (PCA) is a mathematical procedure that transforms possibly correlated variables into a smaller number of uncorrelated variables called principal components. The first principal component (PC) accounts for as much of the variability in the data as possible, and each succeeding PC accounts for as much of the remaining

variability as possible.

PCA is a tool for learning a probability distribution from a set of training samples that are are assumed to be drawn from a multivariate Gaussian distribution. PCA projects the data onto a set of orthogonal bases such that the projection coefficients are standard (zero mean, unit variance), independent, univariate Gaussian random variables. This formulation is useful in a generative setting, as independent samples from a standard univariate Gaussian distribution can be combined to create new data instances. It is also useful in a discriminative setting where the Mahalanobis distance is a natural measure on the data.

Suppose $X = [\mathbf{x}_1 ... \mathbf{x}_N]$ is a set of N samples and each sample $\mathbf{x}_i$ is written as a M element column vector. The full set of data is written as an $M \times N$ matrix. To compute PCA, the data needs to be centered by subtracting the mean μ from each of them as $X' = [\mathbf{x}_1 - \mu, ..., \mathbf{x}_N - \mu]$. The $M \times M$ sample covariance matrix of the centered data is a symmetric and positive semi-definite matrix, computed by

$$\Sigma = \frac{1}{M-1} X' X'^T. \tag{2.19}$$

The eigenvectors and eigenvalues of the sample covariance matrix can be computed by singular value decomposition (SVD):

$$\Sigma = V \Lambda V^T \text{ and } \Lambda_{ij} = \begin{cases} \lambda_i; \ i=j \\ 0; \ i \neq j \end{cases}, \tag{2.20}$$

where V is a matrix composed of the eigenvectors $\mathbf{e}_i$ and Λ is the diagonal matrix whose diagonal entries are the eigenvalues λ_i of Σ. The eigenvectors and the eigenvalues are paired and sorted in descending order of the eigenvalues.

Related to Equation (2.19), the square Mahalanobis distance can be written as

$$d^2_{mahal}(\mathbf{x}) = \sum_{i=1}^{M} \frac{(< \mathbf{x} - \mu, e_i >)^2}{\lambda_i}. \tag{2.21}$$

It follows that given a data instance $\mathbf{a}$ that is represented by $\mathbf{a} = \mu + \sum_{i=1}^{M} \alpha\sqrt{\lambda_i}\mathbf{e}_i$, its square Mahalanobis distance can be computed as

$$d^2_{mahal}(\mathbf{a}) = \sum_{i=1}^{M} \alpha^2. \tag{2.22}$$

Chapter 3

Shape-correlated Deformation Statistics (SCDS)

3.1 Introduction

In 4D image-guided radiation therapy (IGRT) two different imaging modalities are often used at *planning* time and *treatment* time, respectively. Standard multi-slice respiration-correlated CT (RC-CT) scans are usually acquired at the planning time for diagnosis and dose planning purposes. RCCT images have good intensity contrast, high resolution and large field of view in general. Immediately prior to radiation treatment delivery, another imaging procedure is needed for in-room image guidance and dose validation. Factors such as the speed of the scan, the imaging dose and the spatial configuration of the treatment devices determine the feasibility of adopting a certain image modality at the treatment time.

Advanced in-room imaging technologies have became active research areas, and different imaging modalities have been applied in various radiation treatment applications. Due to the low image quality provided by megavoltage (MV) portal images, implanted gold markers are commonly used for treatment setup verification of tumors treated with high-dose, single-fraction radiotherapy (Kriminski et al., 2008). As mentioned in the introduction chapter (Section 1.1.2), cone-beam CT (CBCT) and nanotube stationary

tomosynthesis (NST), have shown great potential for clinical adoption with good localization accuracy during the treatment while eliminating the needs for implanted markers. However, the imaging qualities of both CBCT and NST degrade from tomographic reconstructions with limited spatial coverage of the projection data.

For respiratory image sequences, motion reduction techniques can be used to enhance those low-quality images for treatment image guidance. One technique is to compute an atlas image by averaging all the images after warping them onto a common reference image (Zhang et al., 2010). If the spatial transformations successfully co-register the geometries of all the images, the contrast-to-noise ratio of the anatomical structures will be increased in the resulting atlas image. The key is to accurately estimate the deformations from those artifact-laden treatment images.

There are two different categories of methods commonly used to compute the respiratory motion trajectory. One category is intensity-based image matching or tracking. Various non-linear dense image registrations can be used to calculate the spatial changes of each voxel between images by matching their intensity profiles (Reinhardt et al., 2007; Ehrhardt et al., 2008). However, when image artifacts are present, common image registration optimizations often get trapped in local minima resulting from over-fitting to the image artifacts.

The other category is motion prediction via surrogate-correlated linear modeling. Recognizing the hysteresis of respiration, various external and internal surrogate signals have been used for motion modeling and prediction in lung. Multiple pencil-beam navigators have been used for 3D affine respiratory motion model calibration in coronary MR angiography (Dirk et al., 2003). A function of five degrees of freedom, including tidal volume and its temporal derivative airflow, has been used to model the breathing motion in CT (Low et al., 2005). The diaphragm position of the lung has been used as the internal surrogate for motion prediction in CBCT-guided radiation therapy (Zhang et al., 2007). External skin markers have been used as the surrogate for motion prediction

for MR respiratory motion images (Gao et al., 2008).

In this thesis the shape of the lung is used as the surrogate for the respiratory motion modeling. The spatial changes of the lung boundaries during the respiration are the most visible evidence of motion due to the high-contrast between the intensities of the air-filled region and its surrounding tissues. Compared to the surrogates represented by several landmarks' positions, there are several main advantages in using the shape of the lung as the surrogate: 1) The large volume coverage of the lungs in thorax makes the shape of the lungs a global motion signal. 2) The dense point-sampling on the surface of the lungs enables detailed correlation descriptions of the respiratory motion. 3) The shape of the lung can be segmented from images automatically via its intensity contrast.

In this chapter the framework of shape-correlated deformation statistics (SCDS) is introduced to extract the linear correlation between the shape surrogate and the underlying dense image deformation fields, using planning 4D CTs of a specific patient. The learned-correlations will then be used to predict the motion from image sequences acquired at treatment time for the same patient.

The following section presents the system for training the SCDS. Specifically, section 3.2.1 presents methods to obtain surface models with group-wise correspondence from the training CT images. Section 3.2.2 describes an iterative image atlas formation method to characterize the dense image deformations from the training CT images. Section 3.2.3 introduces the linear correlation regression analysis to calculate the SCDS. Section 3.3 shows experimental results as to the effectiveness of the SCDS. The motion prediction techniques for CBCT treatment images are introduced separately in Chapter 4.

3.2 Methodology

Shape-correlated deformation statistics (SCDS) is designed to reveal the respiratory motion correlations between the surrogate object (the lung) and the overall dense deformation fields from a breathing sequence of time-stamped CTs acquired at planning time.

In the SCDS model, the boundary of the lung is automatically extracted and used as the surrogate object. The intensity differences among the sequence of images are mainly caused by underlying non-linear spatial deformations of the anatomical structures in motion. The residual differences are imaging noises and artifacts. The spatial changes of all anatomical structures can be tracked voxel by voxel via dense image registration.

To prepare relating the shape deformations to the the image deformations, PCA (Section 2.4.1) is used to first reduce the dimensionality of both representations. Then a linear regression is carried out to reveal the strong correlation between the lung shape deformation and the dense image deformation. The result of the regression is a matrix that maps a lung shape to its corresponding dense deformation field.

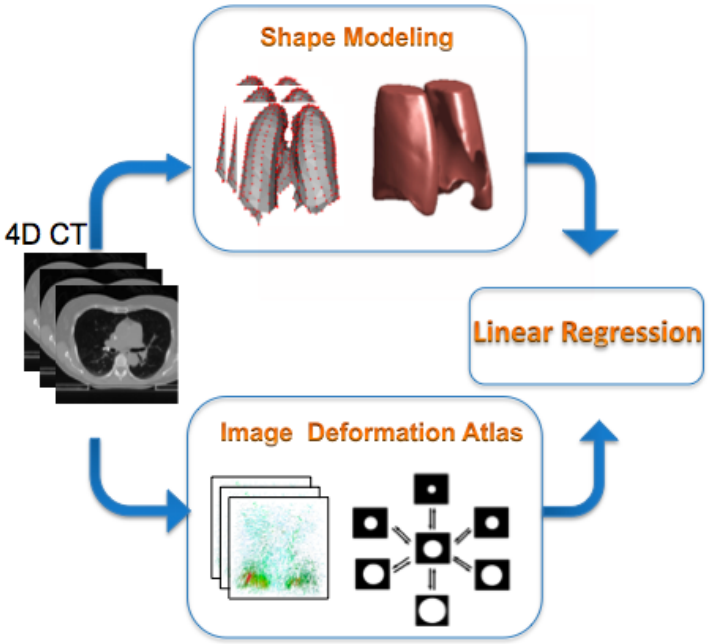
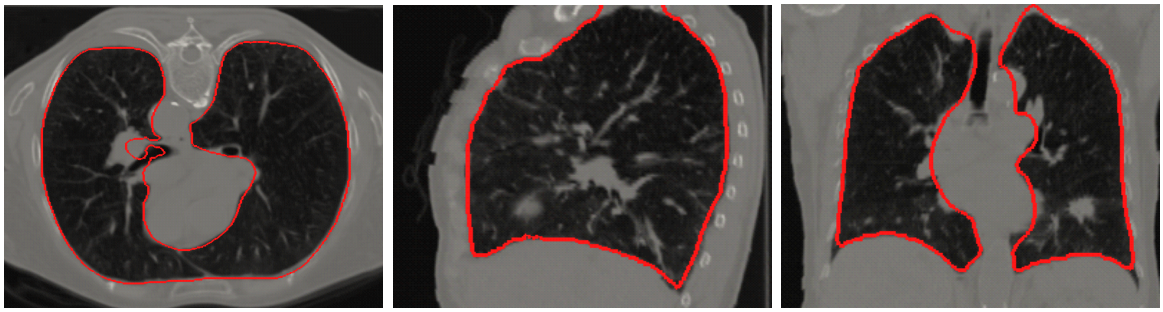


Figure 3.1: **SCDS pipeline diagram.** The pipeline diagram for computing the SCDS is shown with three major modules: the shape deformation extraction, the dense image deformation fields extraction and the linear correlation analysis.

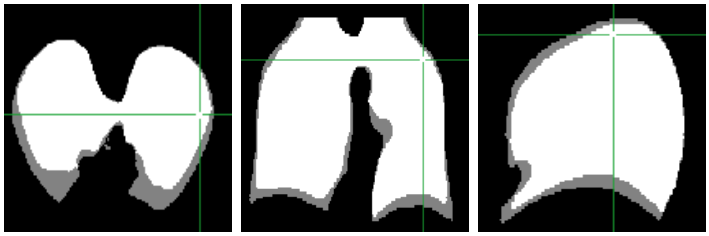
The pipeline for training the SCDS from planning CT images is shown in the diagram in Figure 3.1. Under the assumption that the breathing patterns of the same patient are consistent between the planning time and the treatment time, the correlation statistics trained from the planning time can be then directly applied to treatment time images for prediction.

3.2.1 Shape modeling with group-wise correspondence

In my method lungs are automatically segmented from the gray-scale CT images, similar to the approach in Heuberger et al. (2005). First, an intensity threshold is chosen carefully to exclude airway regions. A 3D mathmorphological "ball rolling" operator fills the holes after thresholding and removes unnecessary details in the connecting regions of bronchial structures, airways and lungs. An example of resulting lung segmentations is illustrated in Figure 3.2.



(a)



(b)

Figure 3.2: **Lung segmentation from planning CTs.** a) The axial, sagittal and coronal views of a lung segmentation from the EI phase of the 4D respiratory-correlated CT images. b) The axial, coronal and sagittal views of the binary segmentation of the EE (white) on top of the EI (gray) phase.

The shape of lung at the ith phase of the breathing sequence (usually 10 CT phases are collected in one breathing cycle) is represented by a surface point distribution model (PDM), denoted by

$$\mathbf{P}^i = [x_1^i, y_1^i, z_1^i, x_2^i, y_2^i, z_2^i ..., x_M^i, y_M^i, z_M^i], \tag{3.1}$$

where $[x_j^i, y_j^i, z_j^i]$ is the jth point's coordinates on the surface of phase i, and M is the total number of surface points for each shape object.

In order to produce tight shape statistics from a breathing sequence of shapes, the geometrical correspondence over the sampling population is crucial. A group-wise correspondence can be implicitly or explicitly enforced using different techinques. An implicit

approach generates all the shape models via warping a common initial model. As long as the deformations are not very large, and the topologies of the shapes are the same, the group-wise correspondence is implicitly enforced by the geometric configuration of the initial shape model. An explicit approach directly measures and maximizes the group-wise correspondence condition while generating the shape models for the entire group at once. Both approaches have been tried for the 4D lung models in the following sections, using a m-rep (medial representation) based surface diffusion method and an entropy-based particle system, respectively.

An implicit approach: m-rep surface diffusion

A single-figure m-rep model is a sheet of medial atoms represented by a quadrilateral mesh of atoms (see Figure 3.3), which carries geometric properties such as the object widening, bending and tapering, with the locality scale given according to the grid spacing (Pizer et al., 2003). Each atom controls two opposing local boundary regions implied by its spokes' end points.

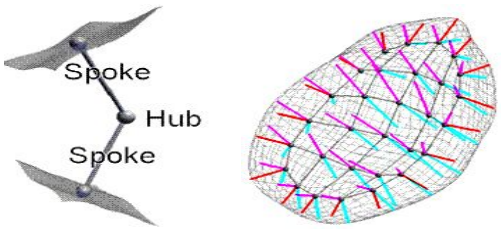
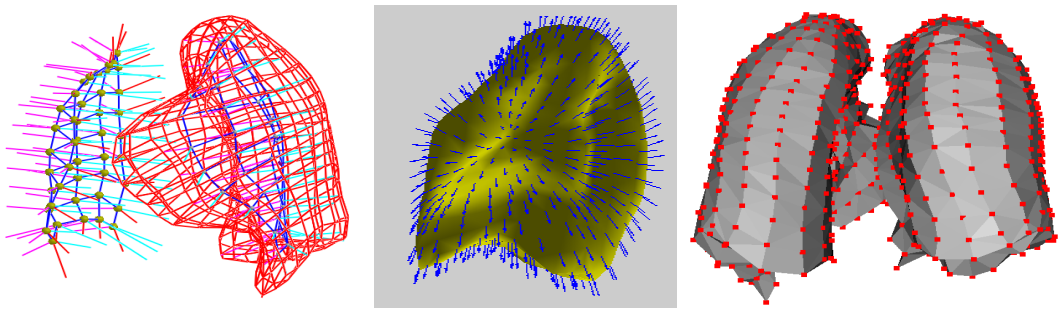


Figure 3.3: **M-rep model.** An m-rep model: An interior medial atom (left) with two spokes and an object (right). The object is composed of multiple medial atoms. Interior atoms have two spokes each. Exterior atoms (on the crest region) have three spokes each.

M-rep modeling has been successfully applied to various statistical shape analysis (Styner & Gerig, 2001) and challenging medical image segmentation tasks (Pizer et al., 2006; Siddiqi & Pizer, 2008). Given a group of binary lung images, the surface group-wise correspondence of the m-rep models is implicitly enforced by deforming an initial

model to fit each segmented lung.

A semi-manual generated m-rep model of the lung is generated and used as the initialization model, as shown in Figure 3.4a. A diffusion process deforms the surface along the surface normal directions to obtain a good fit to the lung boundary. The m-rep provides a fixed number of sparse surface points (the spoke ends) with surface normals (unit spoke vectors) for the diffusion process. The set of surface vertices of each diffused mesh is a PDM, as shown in Figure 3.4c). In the experiments each lung pair (the left lung and the right lung) has 900 surface points.



(a) m-rep (b) normal diffusion (c) PDM

Figure 3.4: **M-rep lung diffusion.** a) A posterior view of the initial m-rep model: the medial sheet of atoms is shown for the left lung, and the surface mesh is shown for the right lung; b) The diffusion surface normals illustrated on the left lung; c) An anterior view of a resulting surface PDM after the diffusion process.

An explicit approach: entropy-based particle system

A state-of-the-art entropy-based surface modeling method has been developed to construct compact shape models with group-wise surface correspondence, which is explicitly quantified and optimized in its objective function (Cates et al., 2007; Oguz et al., 2008). No surface parameterizations are required for applying this method. In this algorithm ordered point sets defining the shape-to-shape correspondences are modeled as sets of dynamic particles whose positions on the objects' surfaces are subject to entropy optimization. A fixed number of particles per lung are automatically generated for lung segmentations of all phases at once with a quick optimization convergence.

A surface $\mathbb{S} \in \mathbb{R}^3$ is sampled using a discrete set of surface points, called *particles*. They are considered to be random variables drawn from a probability density function (PDF). The tightness of the probability distribution is measured by so called *entropy*, which is computed as

$$H[X] = -\int_{\mathbb{S}} p(x) \log\ p(x) dx = -E\{\log\ p(x)\}), \tag{3.2}$$

where $E\{\cdot\}$ denotes expected value, and $p(x)$ is the probability of the particle x being at the position indicated by the current configuration. This entropy measures the spread of the particles across the surface of an object, and it is referred as *surface entropy* in the algorithm. $p(x)$ can be estimated via a nonparametric Parzen windowing method. Details of the optimization realization can be found in Cates et al. (2007).

To measure and enforce the group-wise correspondence, an *ensemble* Z is defined as a collection of N surfaces each with their own set of particles z^k. The ordering of the particles on each shape implies a correspondence among shapes. z^k is regarded as an instance of a random variable Z. The overall entropy-based particle system is designed to minimize the combined ensemble and surface entropy cost function:

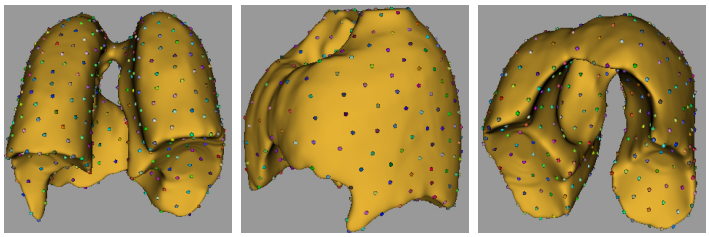
$$Q = H(Z) - \sum_{k=1}^{N} H(z^k), \tag{3.3}$$

which favors a compact ensemble representation balanced against a uniform distribution of particles on each surface.

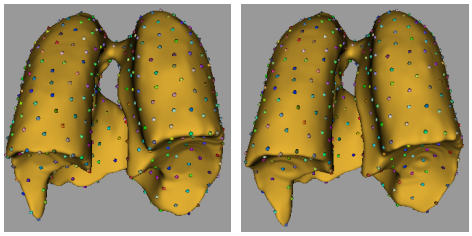
In summary, the optimization of the particle system moves the particles on the surfaces along the direction of the gradient of an energy functional that balances between an even point-sampling on each surface (characterized by a surface entropy) and a high spatial similarity of the corresponding samples across the population (characterized by an ensemble entropy). The particle system is initialized with two seed particles on the surface of the lungs (one on the left lung and the other on the right lung), which then

are recursively split into two to produce the desired number of particles. Please refer to Oguz (2009) for details of the method.

An example of the entropy-based particle optimization results is shown in Figure 3.5. Users can control the sampling density on the surface of the shape. A less dense sampling on the surface leads to a shorter optimization time but a coarser representation. Our experiments suggest that 1024 surface points for a lung pair are dense enough to represent the variation details in respiratory motion.



(a)



(b)

Figure 3.5: **PDMs resulting from the entropy-based particle system.** a) The axial, coronal and sagittal views of the particle sets at EE phase; b) Particles are shown on the surface of the lungs at the EE and EI phase with correspondence indicated by colors.

Compared with the deformable m-rep modeling, the entropy-based correspondence method has several advantages in modeling the lung shape: 1) it explicitly optimizes a group-wise correspondence measurement, which produces a more reliable correspondence condition that is crucial for statistical analysis; 2) it handles arbitrary topology, which makes the task of modeling two lungs at the same time much easier; 3) no initialization model is needed, and the surface sampling density can be adjusted as needed. The computation of the point-based optimization is also efficient, and it generates all the

training models at the same time. In the following studies of this dissertation , the entropy-based particle system is adopted for SCDS computation.

3.2.2 Motion atlas formation

To quantify the breathing motion from images, non-linear dense image registrations are often used to compute the spatial changes for each voxel in the image (Holden, 2008). The breathing motion can be quantified by the non-linear deformations that match each time-stamped image in a breathing cycle to an atlas image or a reference image. The atlas image, together with the deformations, form the respiratory *motion atlas* for this patient.

There are several aspects to be considered in choosing a proper atlas image. First of all, due to the large anatomical variations between patients, it is more practical for the atlas image to be patient-specific. Secondly, to be used for motion prediction, the conformation of the atlas image should be stable between the planning time and the treatment time. Thirdly, the total amount of transformations should be minimum for low total computation time and high computation accuracy.

The Fréchet mean image has the property that it minimizes the sum of squared distances on the Riemannian manifold of diffeomorphic transformations to a group of images. It represents an averaged spatial configuration of that group (Davis et al., 2007). Therefore, the patient-specific Fréchet mean image well satisfies the aforementioned criteria, and it is used in this thesis as the atlas image.

The EE phase has been commonly used as the reference image for registration due to its relatively stable repeatability (Zhang et al., 2010; Ehrhardt et al., 2008). Compared with the Fréchet mean image, it takes much larger total transformations to co-register all the input images. As a result, the PCA statistics of the deformation fields are also less tight. An experiment of improved statistics by using the Fréchet mean is given in Figure 3.12 of Section 3.2.3. Details of Fréchet mean image formation are introduced as

follows.

Fréchet mean image formation

The notion of Fréchet mean has been used to define the mean shape in the nonlinear shape space that has a metric space structure. It is a generalization of the statistical mean from a Euclidean space to a manifold. The mean in statistics captures the notion of central location, and the arithmetic mean is easily computed from data that are elements of a vector space. However, anatomical shape changes are not well described by vector spaces, and the naive approach of averaging an image voxel-wise clearly neither produces a realistic anatomical image nor captures the notion of mean geometric configuration.

Davis et al. (2007) proposed a Fréchet mean image framework for constructing a representative anatomical configuration for a population, using a geodesic distance metric measuring the large-deformation diffeomorphic (LDD) transformation (introduced in Section 2.3.1). They have successfully applied the regression of random design image data to study non-linear geometric changes and variabilities of anatomical structures.

Given a group of time-stamped images, geometric changes over time are represented as the actions of a group of diffeomorphisms on images. Let $\mathrm{Diff}_V(\Omega)$ be the group of diffeomorphisms that are isotopic to the identity. Each element $\phi : \Omega \to \Omega$ in $\mathrm{Diff}_V(\Omega)$ transforms an image I to another image $I \circ \phi$.

The Fréchet mean $\hat{I}$ is the image that requires the least amount of deformation to map onto the group of input images:

$$\hat{I} = \underset{I \in \mathcal{I}}{\operatorname{argmin}} \sum_{i=1}^{N} d(I, I^i)^2. \tag{3.4}$$

A diagram of the Fréchet mean on the manifold is illustrated in Figure 3.6.

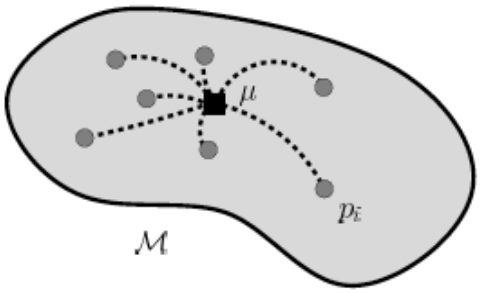


Figure 3.6: **A diagram of the Fréchet mean.** The filled circles represent individual point $\mathbf{p}^i$ on the Riemannian manifold $\mathcal{M}$. The Fréchet mean (filled square) is the point μ on the manifold that minimizes the sum of squared distances to the observations. Distances are measured along the manifold. Excerpt from Davis (2008).

The geodesic distance between a pair of images on the manifold is defined by the LDD image matching measure:

$$d^2(I^F, I^M) = \underset{\phi}{\operatorname{argmin}} \int_0^1 ||v_t||_V^2 dt + \frac{1}{\sigma^2}||I^M \circ \phi - I^F||_{L2}^2, \tag{3.5}$$

subject to $\phi(x) = x + \int_0^1 v_t dt$, as described in in Section 2.3.1. The first term defines a metric on the space of diffeomorphisms that are generated by integrating velocity fields. The second term penalizes residual image dissimilarity. Integrating the geodesic distance definition into (3.4), the optimization problem can be rewritten as

$$\begin{aligned} \hat{I}, \hat{\phi}^i = & \underset{I, \hat{\phi}^i \in \mathbb{I} \times \mathrm{Diff}_V(\Omega)^N}{\operatorname{argmin}} \sum_1^N \left[\int_0^1 ||v_t^i||_V^2 dt + \frac{1}{\sigma^2}||I - I^i \circ \phi^i||_{L2}^2 \right], \\ & \text{subject to } \phi_0^i = Id,\ \phi^i(x) = x + \int_0^1 v_t^i(\phi_t^i(x)) dt. \end{aligned} \tag{3.6}$$

The Fréchet mean image formation framework is implemented via an iterative procedure of continuous joint alignment of the input images: these images iteratively deform towards the evolving mean, as demonstrated in Figure 3.7. Starting from the identity transformation, the deformations are fixed first, and the mean image is updated; then

the mean image is fixed, and the N deformations are updated by taking incremental steps according to the velocities. These two steps are iteratively carried out until convergence. The details of the algorithm are listed in Algorithm 3.1.

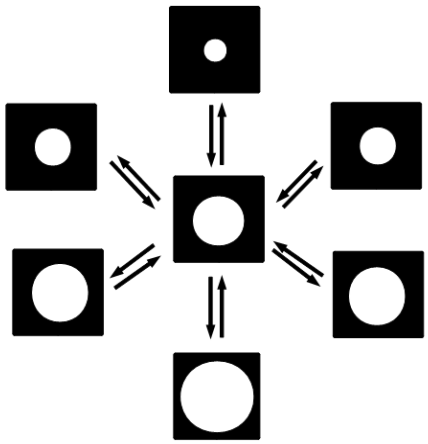


Figure 3.7: **Fréchet mean formation of breathing-spheres.** Iterative Fréchet mean image construction framework illustrated on images of spheres with varying radii. The mean image in the middle minimizes the sum of squared deformation distances required to match all the input images.

Algorithm 3.1 Fréchet mean image formation algorithm: A greedy solution (Davis, 2008)

Input: A collection of N imaegs : $\{I_i\}_{i=1}^{N}$

Output: A Fréchet mean image $\hat{I}$ and a collection of diffeomorphic deformations $\phi_{i=1}^{N}$ that map I_i to $\hat{I}$.

Initialize deformation with identity

for i=1:number of images **do**

  $\phi^i \leftarrow Id_{\mathrm{Diff}_V(\Omega)}$

end for

repeat

  $\hat{I} = \frac{1}{M}\sum_{i=1}^{M} I^i_{\phi^i}$

  for i=1:number of images **do**

    $v_{inc} \leftarrow K(\frac{2}{\sigma^2}\nabla I^i_{\phi^i_{0,1}}(I^i_{\phi^i_{0,1}} - \hat{I}))$

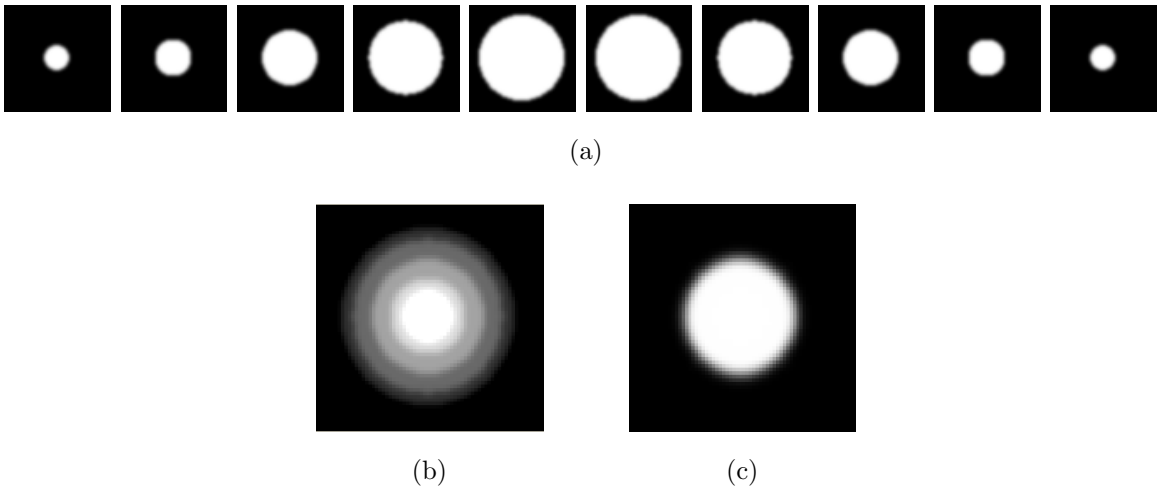
    $\phi^i \leftarrow Exp_{\phi_i}(\epsilon v_{inc})$

  end for

until convergence

A simple test and a comparison of the Euclidean mean and the Fréchet mean are

shown in Figure 3.8. A sequence of 3D images with sphere objects are used as the input data. The radius of the sphere objects changes along a sinusoidal wave, simulating the breathing scenario in 4D lung CTs. The Euclidean mean represents the intensity average of the images, while the Fréchet mean represents the geometry average that matters in analyzing the respiratory motion.



(a)

(b) (c)

Figure 3.8: **Comparison between the Fréchet mean and the Euclidean mean of a breathing-sphere images**: a) A sequence of 3D images of sphere objects with radii varying along a sinusoidal wave; b) The Euclidean mean of the sequence; c) The Fréchet mean of the sequence.

3.2.3 SCDS computation

After extracting the shape deformation and dense image deformation separately from the planning 4D CT images, linear statistical analysis is carried out to model the correlation between them. Considering that both spaces are of high dimensionality, any covariance analysis between the two spaces is computationally prohibitive. Therefore, PCA (Section 2.4) is used for dimension reduction before the correlation regression. Discussions of more sophisticated statistical methods designed specifically for high-dimensional-low-sample-size (HDLSS) data can be found in Lee (2007).

PCA Dimension Reduction

Both shape surrogates and dense deformation fields require dimension reduction. The dimension of the shape representation for one lung pair is typically 1024×3 (the number of points $\times$ 3). The dense deformation field for a 3D CT image is typically $512 \times 512 \times 100 \times 3$ (the image grid $\times$ 3).

The PCA reparameterizations of the shape and dense image deformation spaces are formalized as

$$\mathbf{p}^i = \mu_p + \sum_{j=1}^{k_p} x_j \sigma_{p,j} \mathbf{e}_{p,j} = \mu_p + X^i \mathbf{E}_p, \tag{3.7}$$

$$\mathbf{u}^i = \mu_u + \sum_{j=1}^{k_u} y_j \sigma_{u,j} \mathbf{e}_{u,j} = \mu_u + Y^i \mathbf{E}_u, \tag{3.8}$$

where $\mathbf{p}$ denotes the shape, $\mathbf{u}$ denotes the dense displacement vector field (DVF), i is the phase number, μ is the sample mean, $\mathbf{e}_{\cdot,j}$ is the jth normalized eigenvector, x_j and y_j are the corresponding coefficients or the projection scores of the data in the jth eigendirections, $\sigma_{\cdot,j}$ is the standard deviation of the jth principal component, $\mathbf{E}_\cdot$ is the matrix form of the eigenvectors, and k is the number of eigenmodes. With N training samples (at N respiratory phases), N-1 eigenmodes for both the shape deformation and image deformation spaces are resulting from the PCA calculation.

Given shape models with good group-wise correspondence, PCA treats each model as a sample point in a high dimensional shape space and produces uncorrelated directions, called principal component (PC) directions. The first principal component accounts for as much of the variability in the sampling data as possible, and each succeeding component accounts for as much of the remaining variability as possible. For both the shape and image deformations, the resulting first 1-3 PCs usually capture the majority of the total variation. In this thesis the decision on the number of PCs kept for reparameterization is made upon the minimum number of PCs that contain 90% of total variation.

The residual variation is regarded as sampling noise.

The PCA results from one patient data set, containing 10-phase planning CTs in one breathing cycle, are shown in Figure 3.9, Figure 3.10 and Figure 3.11. After the PCA calculation, the training samples are projected back to the first three principal modes to visualize the distributions of the 10 training phases in the training space. The three-dimensional spaces are spanned by their first three principal component directions, whose score units are normalized by the standard deviations of each principal mode. For example, the coordinate (1.0, 0.0, 0.0) in the PCA space represent an object that is one standard deviation away from the mean along the first PC direction. In the shape variation space the first, second and third modes take up 76.2%, 15.5% and 3.6% of the total variation, respectively; while in the image variation space the first three modes take 67.1%, 19.2% and 4.7% of the total variation. Strong linear correlation between the two spaces can be revealed from the first two modes that cover most of the variation. The correlation coefficient between PC scores of the the shape and the image deformation is 0.98 for the first PC and 0.96 for the second PC.

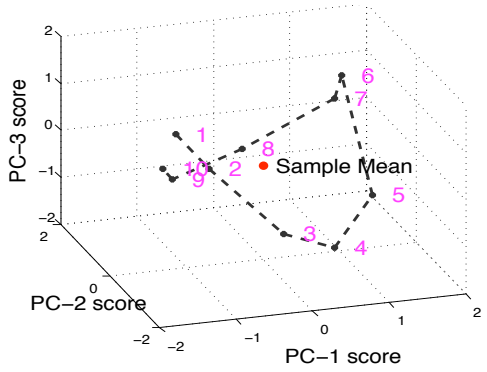


Figure 3.9: **PCA space of lung shapes trained from one sequence of planning CTs.** The first, the second and the third principal components' coefficients/scores of each shape are shown in the 3D coordinate. The numbers tagged along each point are the phase numbers within the breathing cycle.

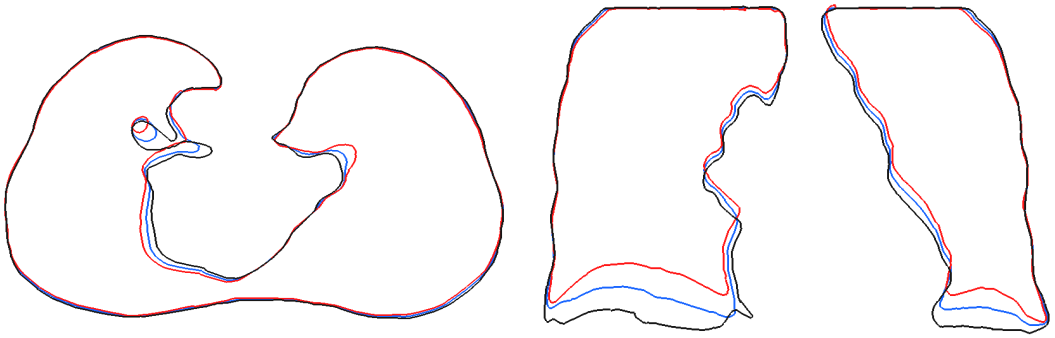


Figure 3.10: **Lung shapes generated from the trained PCA space.** Overlapping 2D slices of the lung shapes at mean+1.5 standard deviation of PC-1(in red), the mean (in blue) and mean-1.5 standard deviation of PC-1 (in black) in the training shape space. Left: An axial view. Right: A coronal view.

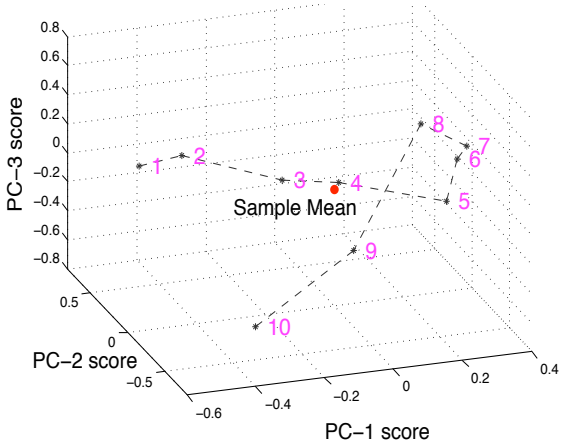


Figure 3.11: **PCA space of dense deformation fields trained from planning CTs.** The first, the second and the third principal components' coefficients are shown in the 3D coordinate. The numbers tagged along the points are the phase numbers within the breathing cycle.

Figure 3.12 shows a comparison of the PCA results of using different reference or atlas images for computing the motion atlas. The improvement in terms of the tightness of the statistics supports the strategy of adopting the Fréchet mean as the motion atlas image.

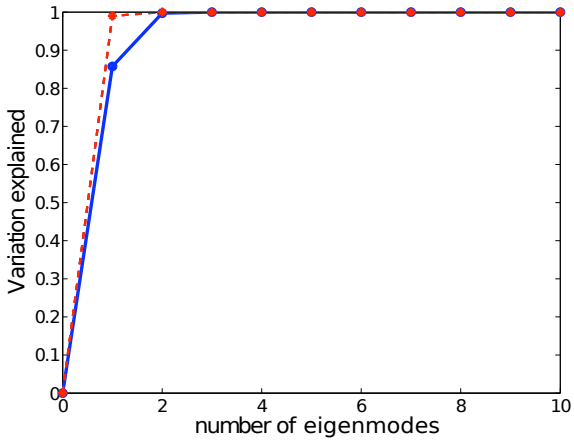


Figure 3.12: **Statistics advantage of using Fréchet mean image as the atlas image.** Tighter statistics are produced using the Fréchet mean as the atlas image (in red), compared with the results using the largest sphere image (the end inspiration phase in breathing) as the atlas image (in blue).

Linear Correlation Analysis

Linear correlation between the surrogate and the respiratory motion has been hypothesized for modeling the respiratory dynamics. Multiple linear regression (MLR) has been used to calculate the correlation parameters between the surrogate and the dense image deformation (Dirk et al., 2003; Zhang et al., 2007). Alternatives to MLR for analyzing the correlation statistics include canonical correlation analysis (CCA) and partial least square method (PLS) (Borga et al., 1997). A brief introduction to the three methods are given below.

MLR MLR, also referring to ordinary least squares (OLS), attempts to model the relationship between regressor (input) data X and response (output) data Y by fitting a linear transformation to observed data. Every value of the independent variable is associated with a value of the dependent variable via $Y = \beta X + \epsilon$, where ϵ is the sum of squares of the residual data that is minimized during the regression. The parameter matrix β can be estimated through $\hat{\beta} = (X^T X)^{-1} X^T Y$. MLR is the most common and convenient way for estimating the regression parameters .

CCA CCA finds two sets of basis vectors, one for X and the other for Y, such that the correlations between the projections of the variables onto these basis vectors are mutually maximized. The canonical correlation basis $\mathbf{A}$ and $\mathbf{B}$ are calculated by solving the eigenvalue equations:

$$\begin{cases} \mathbf{C}_{XX}^{-1}\mathbf{C}_{XY}\mathbf{C}_{YY}^{-1}\mathbf{C}_{YX}\mathbf{A} &= \rho^2\mathbf{B} \\ \mathbf{C}_{YY}^{-1}\mathbf{C}_{YX}\mathbf{C}_{XX}^{-1}\mathbf{C}_{XY}\mathbf{A} &= \rho^2\mathbf{B}, \end{cases} \tag{3.9}$$

where $\mathbf{C}_{XX}$ and $\mathbf{C}_{YY}$ are the within-sets covariance matrices, $\mathbf{C}_{XY}$ and $\mathbf{C}_{YX}$ are the between-sets covariance matrices of X and Y. The eigenvalue ρ^2 is the squared canonical correlation defined as $\rho = \frac{\mathbf{A}^T\mathbf{C}_{XY}\mathbf{B}}{\sqrt{\mathbf{A}^T\mathbf{C}_{XX}\mathbf{A}\mathbf{B}^T\mathbf{C}_{YY}\mathbf{B}}}$. CCA has been recently used in biomedical image analysis fields, especially for predictive modeling of brain anatomical structures in neuroscience (Liu et al., 2004; Fillard et al., 2007; Gao et al., 2008).

PLS PLS searches for a set of components (called latent vectors) that perform a simultaneous decomposition of X and Y with the constraint that these components explain as much as possible of the covariance between X and Y. It is used to predict Y from X and to describe their common structure. Both X and Y are decomposed using factors extracted from the Y^TXX^TY matrix (Wold, 1976). PLS is often used in situations where the use of MLR is severely limited, such as when there are fewer observations than predictor variables.

In summary, MLR, CCA and PLS find the subspaces of minumum square error, maximum correlation and maximum covariance, respectively. The three methods, relating one group of data with the other, can be unified into the following generalized eigenproblem or two-matrix eigenproblem form (Borga et al., 1997):

$$\mathbf{M}w = \lambda\mathbf{N}w. \tag{3.10}$$

The corresponding covariance matrices used for $\mathbf{M}$ and $\mathbf{N}$ are shown in Table 3.1.

Table 3.1: Two-matrix problem foumula for MLR, CCA and PLS

	M	**N**
MLR	$\begin{pmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{pmatrix}$	$\begin{pmatrix} C_{xx} & 0 \\ 0 & \mathbf{I} \end{pmatrix}$
CCA	$\begin{pmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{pmatrix}$	$\begin{pmatrix} C_{xx} & 0 \\ 0 & C_{yy} \end{pmatrix}$
PLS	$\begin{pmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{pmatrix}$	$\begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{pmatrix}$

When dimension reductions via PCA are first carried out on both data sets (X and Y) before calculating the correlation, C_{xx} and C_{yy} are therefore both scaled to the identity matrix $\mathbf{I}$, which makes all the three methods deliver equivalent results.

The maximum correlation directions defined in CCA provide intuitions in understanding the correlations between the variation spaces of shape surrogate and the deformation fields. The maximum correlation coefficients also provide a direct measurement of the strength of the correlation. CCA is therefore used in the thesis for better interperating the correlations. The implementation details of the correlation analysis using CCA is presented as follows.

After the PCA dimension reduction, the canonical correlation (CC) between the two multidimensional variables X_t and Y_t is maximized via the following canonical transformation. Given data matrices $\mathbf{X} = [X_0 - \bar{X}, ..., X_t - \bar{X}, ..., X_1 - \bar{X}]$ and $\mathbf{Y} = [Y_0 - \bar{Y}, ..., Y_t - \bar{Y}, ..., Y_1 - \bar{Y}]$, where $\bar{X}$ and $\bar{Y}$ are the mean vectors of the training sample, the canonical variables α and β are calculated by projecting the data onto the basis matrices $\mathbf{A}$ and $\mathbf{B}$:

$$\alpha = \mathbf{X}^T\mathbf{A}, \beta = \mathbf{Y}^T\mathbf{B}. \tag{3.11}$$

The maximized correlation coefficient ρ indicates how strong the correlation between the surrogate shape deformation and the entire image deformation is. To use the statistics for predicting the dense image deformation $\hat{\mathbf{u}}^*$ of a target image at phase $*$, the following steps are carried out:

(1) Reparameterize the shape $\mathbf{p}^*$ via PCA: $X^* = (\mathbf{p}^* - \mu_p)\mathbf{E}_\mathbf{p}^T$;

(2) Calculate the CC variate of the shape: $\alpha^* = (X^* - \bar{X})\mathbf{A}$;

(3) Compute the CC variate of the image displacement vector field: $\beta^* = \alpha^*\rho$;

(4) Compute the PCA score of displacement vector field: $Y^* = \beta^*\mathbf{B}^{-1} + \bar{Y}$;

(5) Finally, calculate the predicted displacement vector field: $\hat{\mathbf{u}}^* = \mu_\mathbf{u} + Y^*\mathbf{E}_\mathbf{u}$.

3.3 Experimental Results

In our experiments the respiration-correlated CT data sets are provided by a 4-slice scanner (LightSpeed GX/i, GE Medical System), acquiring a time series of CT projections for a complete respiratory cycle at each couch position while recording patient respiration (Real-time Position Management System, Varian Medical Systems). The CT projections are retrospectively sorted (GE Advantage 4D) and reconstructed to produce a series of 3D images at 10 respiratory time points. The time resolution is 0.5 s. The CT slice thickness is 2.5 mm. For each time point, tumor contours by experts' delineation are used as the ground truth.

3.3.1 Intra-session study: leave-one-phase-out

A leave-one-phase-out (LOPO) study was carried out for each of the five patients. Namely, the statistics that were used on each target phase image were trained from the other 9 phases from the same breathing cycle.

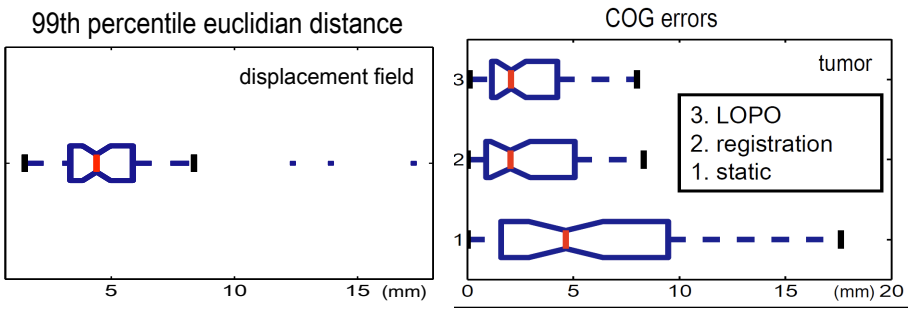


Figure 3.13: **Leave-one-phase-out study on 5 patients.** Left: The 99th percentile Euclidian distance between the predicted displacement field and the "true" displacement field; Right: The Euclidian distance on the center of gravity (COG) between the deformed GTV and the delineation truth. Each notched box contains 50 prediction results.

In Figure 3.13 I show the LOPO prediction results on five patients (50 predictions in total). Standard boxplots are used for statistical comparisons. The Euclidean distance between the predicted displacement fields and "true" displacement fields were measured. The "true" displacement fields were computed via diffeomorphic registration. The 90% quantile surface distance accounts for large discrepancies. Using clinicians' delineations as the ground truth, the predicted GTV is evaluated in terms of the center of gravity (COG) location errors. To estimate the mobility of the tumor region, the location errors without motion correction (static GTV) are calculated using the reference GTV for evaluating all phases. I also plot the estimation errors of directly applying the image deformations calculated by image registrations for comparison.

The LOPO studies show that the SCDS-prediction effectively corrects the tumor motion and achieves the same accuracy of the image registration. The maximum error of 0.25 cm surface distance is in fact within one voxel size, 0.1 cm $\times$ 0.1 cm $\times$ 0.25 cm. In all the patients the prediction substantially reduces the estimation error from the static (no motion correction) scenario. An example of predicted tumor volume overlapped with clinicians' manual segmentations is shown in Figure 3.14.

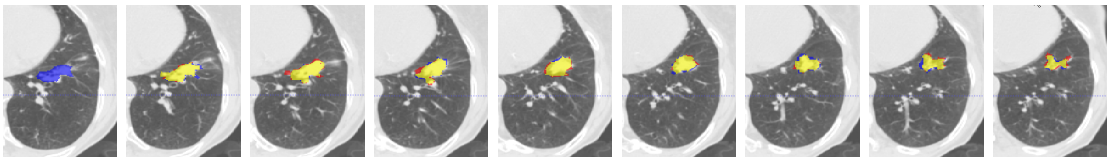


Figure 3.14: **An example of tumor evaluation of the leave-one-phase-out study.** A median case (from the 50 predictions in terms of the COG error) of predicted GTV in sequential 2D slices. Blue: predicted tumor volume; Red: manual delineation; Yellow: overlapping volume. This predicted GTV has a COG error of 1.9 mm and a Dice's coefficient of 0.86.

3.3.2 Inter-session study

The LOPO studies in the previous section were carried out using a single RCCT sequence where both shape and image deformation statistics were fairly tight. On the contrary, in this inter-session study a planning RCCT sequence obtained on one day was used to predict the motion from the images acquired at a different day. Although it is assumed that the breathing patterns of the same patient are repeatable to a certain degree, it is not surprising to observe noticeable variations between sequences taken days, months or years apart. Without extra knowledge introduced to the statistics, the current SCDS model is not designed to handle inconsistent motion patterns between the training and the test scan.

In the following experiments, a primary planning RCCT sequence and a secondary RCCT sequence of the same patient were taken with different imaging protocols. Besides lower image resolution, the secondary RCCT sequence has a much smaller imaging region and only four phase bins reconstructed. The coverage in the longitudinal direction of the second session is about 6 to 10 cm, centered around the tumor region, as shown in Figure 3.15. The four phases sample the whole breathing cycle with roughly equal intervals, with the EE and EI phases acquired. In predicting the motion for the secondary session, the statistics trained from all ten phases of the primary session were used.

Before applying SCDS, there are some pre-processing steps. A rigid registration based on the bony anatomy is necessary to align the secondary session images to the

primary session images. The pre-calculated displacement field and binary lungs at each phase from the primary session needed to be down-sampled and cropped to match the resolution and the size of the secondary scans.

The correlation statistics modeling process is not much different than previous intra-session study. With the same RCCT imaging modality, the same binary segmentation scheme can be used. The only notable difference is that the shape of the surrogate extracted from the secondary-session scan is no longer the full set of lungs due to the limited field of view. The secondary-session lung regions contain the bottom parts of the lungs and a cut plane on top. Since the entropy-based correspondence algorithm does not depend on topology, the points on the top plane are not treated differently and therefore forced to be matched across the phases, which introduces some artificial variations. Also, the real in-and-out-of-plane motion variations during the breathing will not be constructed into the statistics due to the cut-off. As a result, the statistical model in this setting is expected to be less robust than the intra-session study.

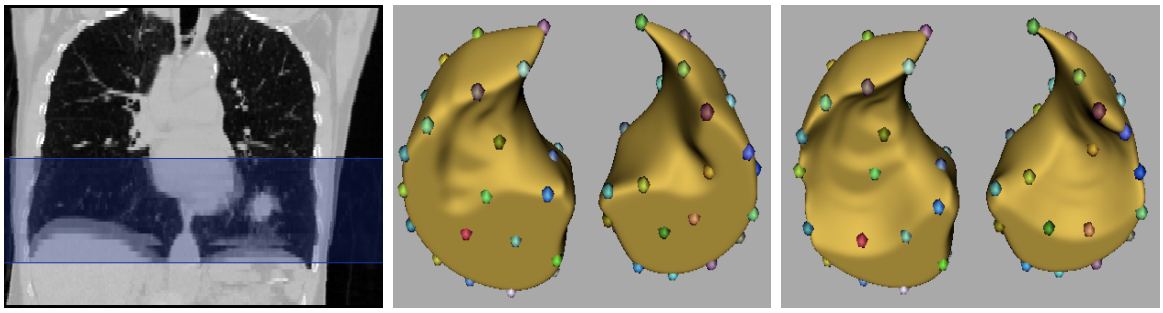


Figure 3.15: **Shape models used for inter-session study.** Left: The second session EI phase image is shown on top of the first session EI phase mage: a coronal view, the intersection region is the darker region on the bottom part of the lungs; Right: The PDM of the partial lung shape, composed of 64 surface points, shown with the surface of the lungs at EE phase and at EI phase with group-wise correspondence.

The preliminary tests were evaluated on 3 patients in terms of the tumor location errors (COG error), see Figure 3.16. Clinicians' delineations of the tumors were provided on the EI and EE phases of the secondary-session scans. The predicted image deformation propagated the tumor volume from the reference image (EE phase of the

primary scan). The prediction results were able to correct some of the tumor motions as large as 1.5 cm.

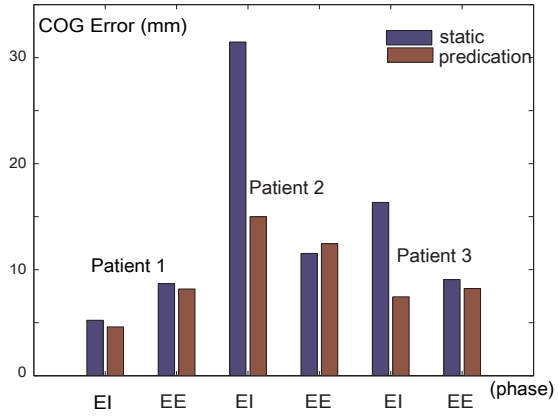


Figure 3.16: **Inter-session study on 3 patients** (each has 10 first-session images for training): The COG errors of the predicted GTV are compared with the clinician's delineations on EE and EI phases.

3.3.3 Surrogate comparison

In the earlier work of Zhang et al. (2007), the diaphragm position along the longitudinal axis was used as a distinctive one-dimension surrogate signal, which can be manually labeled in 3D. The one dimensional surrogate might be capable of navigating the major breathing motion, but it is overly simple for accurately estimating complicated respiratory motion. The farther away the region is to the diaphragm, the less accurate the deformation might become. Especially in cases where the tumors are located in the upper part of the lung, the deformation around a tumor might not correlate well with the diaphragm.

In SCDS a lot more features are used by utilizing the shape of the overall lung region, which therefore is designed to be more comprehensive and robust for the respiratory motion prediction. In order to quantitatively understand the advantages of engaging the more complicated surrogate, I compared the prediction results by replacing the shape model with the diaphragm position that were manually specified. Linear regression was

computed to map the diaphragm position to the dense image deformation field. I carried out the LOPO studies using two patients' 4D RCCT sequences in which the tumors are located in the neighborhood region of the diaphragms. The evaluations in terms of the prediction errors on the displacement field, lung boundaries and GTV are shown in Figure.3.17. As a local feature the diaphragm position is shown to be less capable of predicting the nonrigid deformation of the entire imaging field.

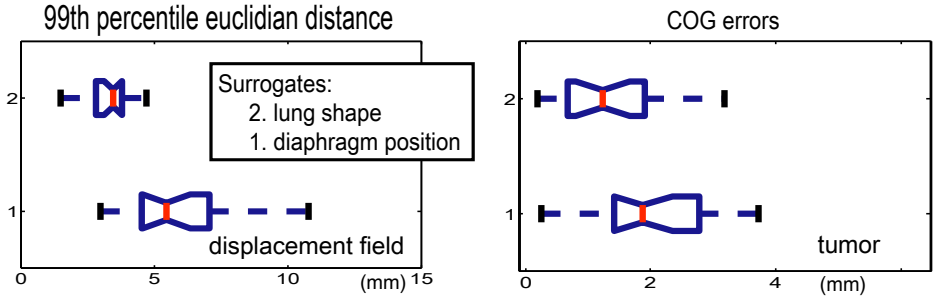


Figure 3.17: **Surrogates comparison evaluated on tumor.** Comparison of the surrogates on 2 patients (20 predictions in each notched box). Left: The 99th percentile Euclidian distance between the predicted and the true displacement field is plotted to show the large discrepancies.Right: The center of gravity of the predicted GTV is compared with the clinician's delineation.

3.4 Conclusion

The proposed shape-correlated deformation statistics (SCDS) was introduced to predict patient-specific respiratory motion efficiently. In training the SCDS, each phase in the planning 4D sequence is treated as an independent sample. No temporal correlations among the time-stamped phases are assumed, which makes the method applicable to irregular breathing patterns often seen in reality. The PCA dimension reduction in both image and shape deformation spaces not only retains anatomically relevant variation but also reduces sampling noise. The hypothesis of strong linear correlation between the shape and the image deformation space has been shown to be effective. The leave-one-phase-out study on both phantom data and patient data demonstrated that the SCDS is able to estimate the within-sequence image deformation accurately.

It is possible to extend the methodology to inter-patent studies to include more training samples in the hope of obtaining more robust statistics. However, the inter-patient variation can be very large due to various factors, including the different breathing patterns and different sizes of the anatomical structures etc. In terms of predicting individual patients' respiratory motion, a patient-specific model is more helpful.

Provided with the SCDS trained from planning respiration-correlated CTs, the next chapter presents the techniques used for the motion prediction from treatment cone-beam CTs with imaging artifacts.

Chapter 4

Respiratory Motion Prediction using SCDS

4.1 Introduction

In the previous chapter SCDS was introduced to model deformation correlations between lung shapes and dense image deformation fields from planning RCCT images. It was shown that linear statistics is able to effectively capture respiratory motion patterns.

In order to apply patient-specific statistics trained from planning time to image guidance at treatment time, consistent motion patterns between the planning time and the treatment time are assumed. This consistency assumption provides the basis for all surrogate-invovled motion prediction techniques (Dirk et al., 2003; Zhang et al., 2010; Gao et al., 2008). To be more specific for the SCDS-based prediction, the correlation parameters between the lung shape and the image deformation field are assumed to be the same at both times. The motion prediction accuracy relies on this consistency condition.

Another factor that determines the accuracy of the motion prediction is the surrogate extraction/measurement accuracy from the treatment images. When using the lung as the surrogate, severe artifacts in the treatment images certainly challenge the segmentation task. And the segmentation errors would directly result in errors in the predicted dense deformation fields.

To account for those image artifacts, a probabilistic deformable segmentation scheme has been developed to fit deformable surface models to the treatment images while staying in a trained shape space. The SCDS introduced in Section 3.2.3 uses PDMs generated from the entropy-based particle system. The PDMs are isolated point sets sampled on each surface object. The point-based representation is sufficient for statistical computations. However, a richer geometric description of the surface is needed for deformable model segmentation; a surface mesh structure from which the surface normals and curvatures can be computed serves the purpose.

In this chapter a surface interpolation method is introduced to construct surface meshes using the surface points generated from the planning images. The interpolation scheme is designed so that the resulting mesh vertices maintain good group-wise correspondences. After computing SCDS using the training meshes, the deformable segmentation optimization is initialized with the mean shape and optimized via a posterior probability energy function till convergence. The predicted motion is then calculated using the trained SCDS.

The rest of the chapter is organized as follows. Section 4.2.1 introduces the surface mesh interpolation method. Section 4.2.2 presents the probabilistic deformable segmentation method. Section 4.3 shows the experimental results on cone-beam CT images from computational phantom data and real patient data. Section 4.4 concludes this chapter.

4.2 Methodology

4.2.1 Surface mesh interpolation

In the inter-session prediction study of Section 3.3.2, both the target and the training data are high-quality RCCT images. The entropy-based particle system described in Section 3.2.1 produces PDMs from a sequence of binary segmented images. The points

with the same index on the resulting PDMs correspond to each other.

However, for treatment images such as CBCT and NST, the thresholding-based binary segmentation method is no longer applicable due to the interfering image artifacts. Therefore, the particle system can not be used to obtain PDMs from the treatment images. Instead, deformable model segmentations can be used to fit surface models into noisy images using the shape statistics trained from the planning images. Since the fitted models are all deformed from a common initial model, the group-wise correspondence is kept.

Theoretically it is possible and convenient to use a PDM as the deformable model representation. However, a typical PDM for a pair of lungs contains about 1000 points distributed on the surface. The sparse points set representation is lacking detailed geometrical descriptions of the shape, such as surface normal and surface curvatures, which are important features to use for accurate segmentation. The deformable model fitting accuracy is therefore limited. It is also hard to evaluate the credibility of the segmentation results with sparsely distributed points.

A detailed surface mesh representation that contains rich geometric attributes is desired for accurate surrogate segmentation. Oguz (2009) has proposed an interpolation scheme to reconstruct surface meshes from the PDMs generated from the entropy-based particle system. Given the PDMs and their corresponding-phase the binary lung images, there are five steps to obtain the surface meshes:

1. Generate surface meshes from the input binary image sequence using the marching cubes algorithm.

2. An arbitrary phase number is selected, and the surface mesh of that phase is used as the *template* mesh.

3. For each vertex of the template surface mesh, three nearest points from its corresponding-phase PDM are located. Using barycentric coordinates, the interpolation coeffi-

cients that interpolate the vertex position from the three points' positions are computed. Figure 4.1 illustrates the scenario.

4. For all other phases, the interpolation coefficients calculated from Step 3 are used to locate their interpolated mesh vertices from the corresponding-phase PDMs.

5. Finally, the interpolated vertices are projected onto the original surface for each phase to compensate for discrepancies.

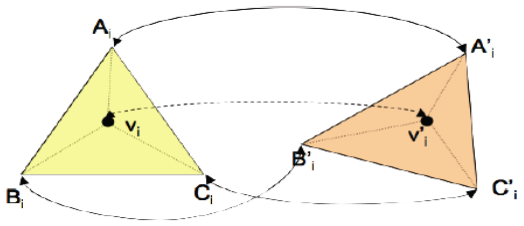
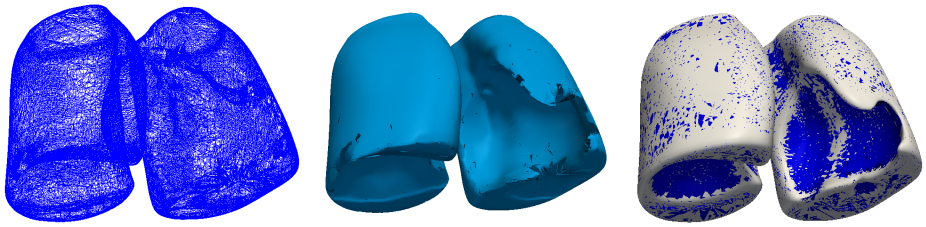
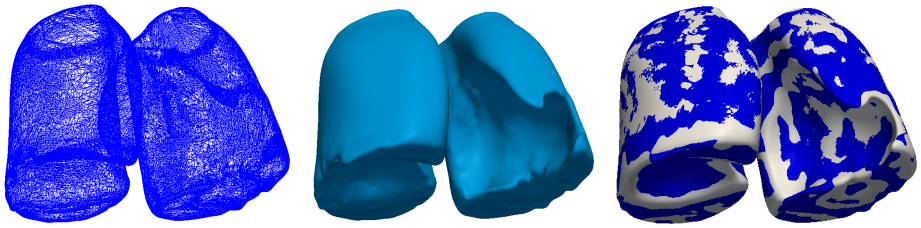


Figure 4.1: **Surface interpolation.** Surface interpolation from PDMs with group-wise correspondence. Given the desired surface sample set, the closest points A_i;B_i;C_i to each sample point v_i are identified, and the barycentric coordinates of v_i in this triangle are computed. Then, the corresponding sample location v_i on a different surface can be obtained by interpolating using these barycentric coordinates in the triangle $\triangle A'_i$,B'_i, C'_i, where A'_i;B'_i;C'_i are the points corresponding to A_i;B_i;C_i, respectively. Excerpt from Oguz (2009).

However, this method is limited to smooth surface shapes. In highly curved lung shape edge regions, the discrepancy can be significant, causing holes or spikes in the interpolated meshes as illustrated in Figure 4.2a. To improve the quality of the surface interpolation, a thin-plate splines (TPS) warping (Bookstein, 1989) scheme is used to achieve more smoother surface interpolation. In this TPS-based method, the particles are used as the control points of thin-plate splines to calculate a smooth transformation that warps each surface to the template surface. The resulting meshes are therefore much more smoother, as demonstrated in Figure 4.2b.



(a) Non-smoothing surface interpolation results.



(b) Improved surface interpolation results with TPS warping.

Figure 4.2: **Surface interpolation results with TPS warping.** The original surface interpolation results are shown in row a, and the improved results using the TPS warping are shown in row b. Three different visualizations are shown for comparison: (Left) The interpolated surfaces at the EE phase, shown as wireframes; (Middle) The interpolated surfaces at the EE phase, shown as rendered surfaces; (Right) The interpolated surfaces (in blue) overlapped with the original EE phase surfaces (in light gray). Overall, the TPS result is smooth and fits the original surface better.

Using this interpolation method, the mesh representation of the surrogate can be generated from the training/planning images. When computing the shape prior for treatment image segmentation, the coordinates of the mesh vertices are used as the input data of PCA. Since the group-wise correspondence is maintained during the interpolation process, there is no noticeable difference in the outcome of correlation statistics comparing to the PDM representation. During the deformable segmentation, the connectivity (triangulation) among the vertices is kept stable while the positions of the vertices are optimized.

4.2.2 Probabilistic deformable model segmentation

A posterior probabilistic optimization scheme has been adopted to calculate the deformable mesh models that fit into the target images while staying in the trained shape space. This probabilistic framework incorporates a prior distribution of the surface shape model and provides a measure of the uncertainty of the estimated shape after fitting the model to treatment image data (Szeliski, 1990).

In the following formulation J^i (in contrast to the planning image I^i) denotes the treatment image at phase i and $\mathbf{q}^i$ (in contrast to the planning shape model $\mathbf{p}^i$) denotes the shape model extracted from the image J^i. The posterior probability $p(\mathbf{q}^i|J^i)$ can be expressed by Bayes' rule:

$$p(\mathbf{q}^i|J^i) = \frac{p(J^i|\mathbf{q}^i)p(\mathbf{q}^i)}{p(J^i)}, \tag{4.1}$$

where $p(J^i|\mathbf{q}^i)$ is the conditional probability or the likelihood of producing image J^i for a given model $\mathbf{q}^i$. The true model $\mathbf{q}^i$ for an image can be inferred by maximizing the posterior probability, that is, by finding a model $\mathbf{q}^i$ that has the maximum posterior probability given the image data J^i:

$$p(\hat{\mathbf{q}^i}|J^i) = \max_{\mathbf{q}^i} p(\mathbf{q}^i|J^i) = \max_{\mathbf{q}^i} \frac{p(J^i|\mathbf{q}^i)p(\mathbf{q}^i)}{p(J^i)} \tag{4.2}$$

This kind of optimization is known as the maximum-a-posterior (MAP) estimation. The expression can be simplified by taking the **log** and leaving out the prior distribution of the image data $p(J^i)$ that is equal for all $\mathbf{q}^i$. Therefore, it is equivalent to the maximization expressed in

$$\hat{\mathbf{q}^i} = \arg\max_{\mathbf{q}^i} \left[\mathbf{log}\ p(J^i|\mathbf{q}^i) + \mathbf{log}\ p(\mathbf{q}^i)\right]. \tag{4.3}$$

The shape prior is measured by its Mahalanobis distance (refer to Section 2.4) in the

PCA space trained from the planning data, computed as

$$\mathbf{log}\ p(\mathbf{q}^i) = -\frac{1}{2}\sum_{j=1}^{N} x_j^2 = -\frac{1}{2}\sum_{j=1}^{N}\left(\frac{<\mathbf{q}^i - \mu_p, \mathbf{e}_{j;p}>}{\sigma_{j;p}}\right)^2, \tag{4.4}$$

where x_j is the score of the jth principal direction. The Mahalanobis distance evaluates how far the deformed model $\mathbf{q}^i$ is from the training mean in the trained shape space.

It is worth emphasizing that in real patient experiments the training (RCCT) shape variation space and the target (CBCT) shape variation space can be quite different even for the same patient, due to the different patient set-up positions during two different imaging procedures. However, after rigidly aligning the two image sequences based on the bone structures in a pair of images at the same reference phase (e.g., EE phase), I assume that the two spaces are approximately the same. As a result of using the shape prior, the CBCT shape models ($\mathbf{q}^i$) maintain the group-wise surface correspondence with the training models ($\mathbf{p}^i$).

The image-match term or the likelihood term is the driving force that pulls the surface to match the boundaries of the anatomical object. This force component is represented here as the sum of second-order gradient magnitudes, which is a common feature used for deformable surface model segmentations (chapter 9.4.4 of Ibanez et al. (2005)). The input image is passed into the filter that computes the magnitude of the image gradient. The resulting scalar image is then passed to another gradient filter. The output of this second gradient filter is a vector field, in which every vector points to its closest edge in the image, and it has a magnitude proportional to the second derivative of the image intensity (see Figure 4.3).

Since this vector field is computed using Gaussian derivatives, it is possible to adjust the smoothness of the vector field by the sigma assigned to the Gaussian. Large values of sigma result in large capture radii but with poor precision in locating the edges. A coarse-to-fine strategy may involve the use of large sigmas for the initial iterations of the

model and small sigmas to refine the model when it is close to the edges. When fitting lung models to a typical respiratory image sequence, the maximum surface distance between the initial model (the mean model of the training data) and the target model is less than 1.5 cm (rigid registration is carried out first to align the test data to the training data via bony structures). Sigmas of 5-10 mm usually perform equally well.

The whole pipeline of the probabilistic deformable model segmentation is summarized in Figure 4.3. The target RC-CBCT images are aligned with the training RCCT images first based on the bony structures in the image. Usually a pair of EE-phase images are used to calculate the rigid alignment parameters. Starting from the training mean model, the edge-enhanced gradient images are used for the deformable segmentations.

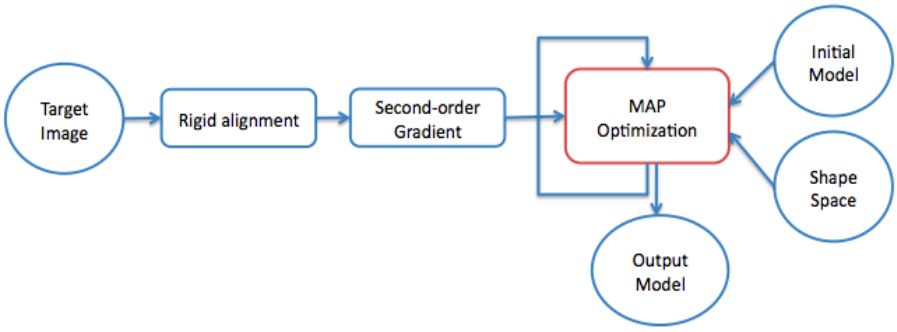


Figure 4.3: **Diagram of the probabilistic deformation segmentation pipeline.**

Figure 4.4 shows an example of the segmentation results on simulated CBCT images using Nurbs-based Cardiac-Torso (NCAT) computational phantoms.

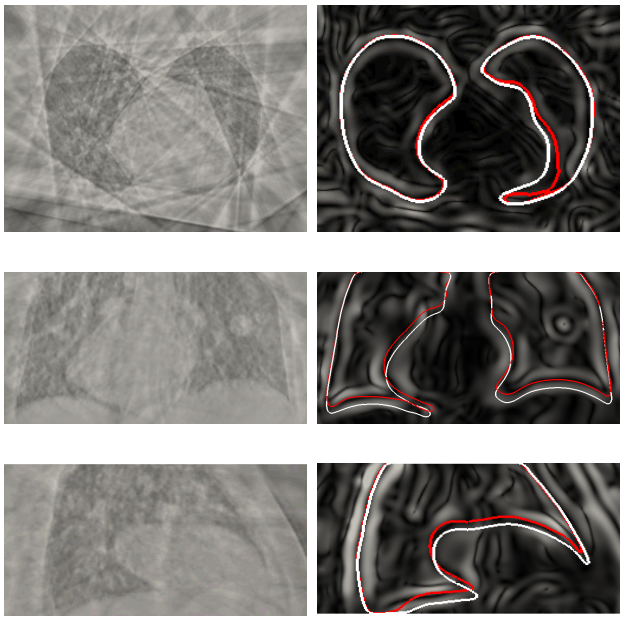


Figure 4.4: **An example of CBCT deformable model segmentation.** Left: The axial, coronal and sagittal slices of a NCAT CBCT (at EE phase) image. Right: The axial, coronal and sagittal slices of the second-order-gradient magnitude image are shown with the final segmentation contours (in red) and the initial mean shape model (in white).

4.2.3 Image deformation prediction and evaluation

The following three steps are carried out to estimate the image deformation represented by the dense vector field (DVF) $\mathbf{u}^i$ for CBCT image J^i at phase i:

1. **Training SCDS**: Obtain the image deformations in the planning CTs by the intensity-based Fréchet mean formation method (Section 3.2.2), and extract the surface mesh models of the lung from each CT phase images (Section 4.2.1). Then calculate the linear regression matrix $\mathbf{C}$ that maps the shape surrogate $\mathbf{p}$ to its corresponding dense image deformation $\mathbf{u}$, such that $\mathbf{u} = \mathbf{C} \cdot \mathbf{p} + \epsilon$, where ϵ is the regression error.

2. **Surrogate Segmentation**: Segment lung shapes ($\mathbf{q}^i$) from CBCT images (J^i) using the probabilistic deformable segmentation scheme (Section 4.2.2).

3. **Linear mapping**: Calculate the image deformation $\mathbf{u}^i$ of each phase of the CBCT sequence J^i by $\mathbf{u}^i = \mathbf{C} \cdot \mathbf{q}^i$.

With the predicted DVF $\mathbf{u}^i$, the corresponding diffeomorphic transformation φ^i for phase i is approximated by

$$\hat{\varphi^i} \approx x + \mathbf{u}^i. \tag{4.5}$$

Then the atlas image can be calculated by averaging all the intensity images after warping them using the transformations, expressed in

$$\hat{J} = \sum_{i}^{N} J^i \circ \hat{\varphi^i}. \tag{4.6}$$

The accuracy of the prediction can be measured on the contours of the tumor regions or other anatomical structures of interest by applying the predicted transformations. Usually the contour of the EE phase treatment image is manually segmented and used as the base/template contour O^b . The contour on the atlas image needs to be computed first, as $\hat{O} = O^b \circ (\varphi^b)^{-1}$. Then the predicted contour of the treatment image at phase i can be calculated by

$$O^i = \hat{O} \circ \varphi^i = (O^b \circ (\varphi^b)^{-1}) \circ \varphi^i. \tag{4.7}$$

4.3 Experimental Results on CBCTs

4.3.1 NCAT data

4D NURBS-based Cardiac-Torso (NCAT) phantom thorax CTs were produced (Segars et al., 2001) at 10 phases sampled in one breathing cycle. A corresponding CBCT sequence was reconstructed from the NCAT CTs using the protocol of a gantry-mounted

KV on-board imaging system (Varian Medical Systems) (Kriminski et al., 2008) used in one-minute-scan patient radiation therapy guidance. An example image pair is given in Figure 4.5. The streak artifacts in real patient CBCT data are realistically produced.

Since the CBCT is reconstructed from the CT images, both the training and the test sequences have the exact same underlying respiratory motion but at different phases. The CTs are generated at 10 time points, and the CBCTs are reconstructed at 6 time points. Time points in both sequences are evenly distributed within one breathing cycle.

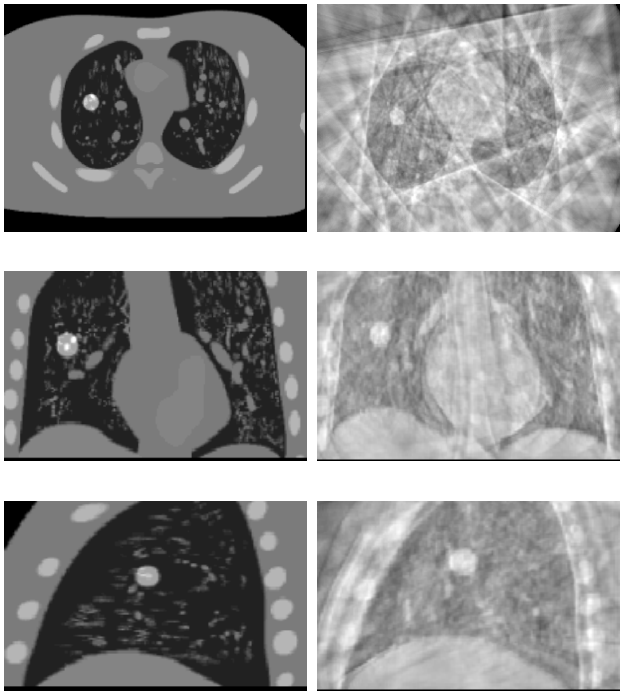


Figure 4.5: **CBCT images reconstructed from NCAT CTs.** Left: The axial, coronal and sagittal views of NCAT CT images. Right: The axial, coronal and sagittal views of NCAT CBCT images. The bright spherical object that can be seen from all three views is the tumor.

Two data sets were simulated with different physical parameters to vary the motion pattern. As introduced in Section 2.2.1, the diaphragm motion and the anterior-posterior motion are the major deformations that change the cavities of the lungs. The first NCAT data set has a maximum of 2.0 cm diaphragm motion and 0.5 cm anterior-posterior motion. The second NCAT data set has a maximum of 1.5 cm diaphragm motion and

1.5 cm anterior-posterior motion. All other parameters for the two phantom sets are the same. The tumor in the image moves along with its nearby tissues smoothly.

In both studies the prediction results were evaluated in terms of tumor location errors and entire image deformation fields' estimation errors. The tumor location estimations were calculated by propagating a template contour (usually provided at EE phase) to all the phases using the predicted image deformations. In order to evaluate the estimation errors of the overall image deformation fields, the ground truth of the deformations were obtained by linear interpolations from planning CT deformations according to their phase numbers in the breathing cycle.

The evaluations are shown in Figure 4.6 and Figure 4.7 for the two NCAT data sets separately. The accumulated PC variance in percentage is shown for both the shape space and the image deformation space from training. Depending on how many PCs it took to capture more than 90% total variance, usually one or two PCs were kept for SCDS computation. If only one PC was kept there was no need to carry out multivariate canonical correlation analysis (CCA). A simple linear regression is used instead.

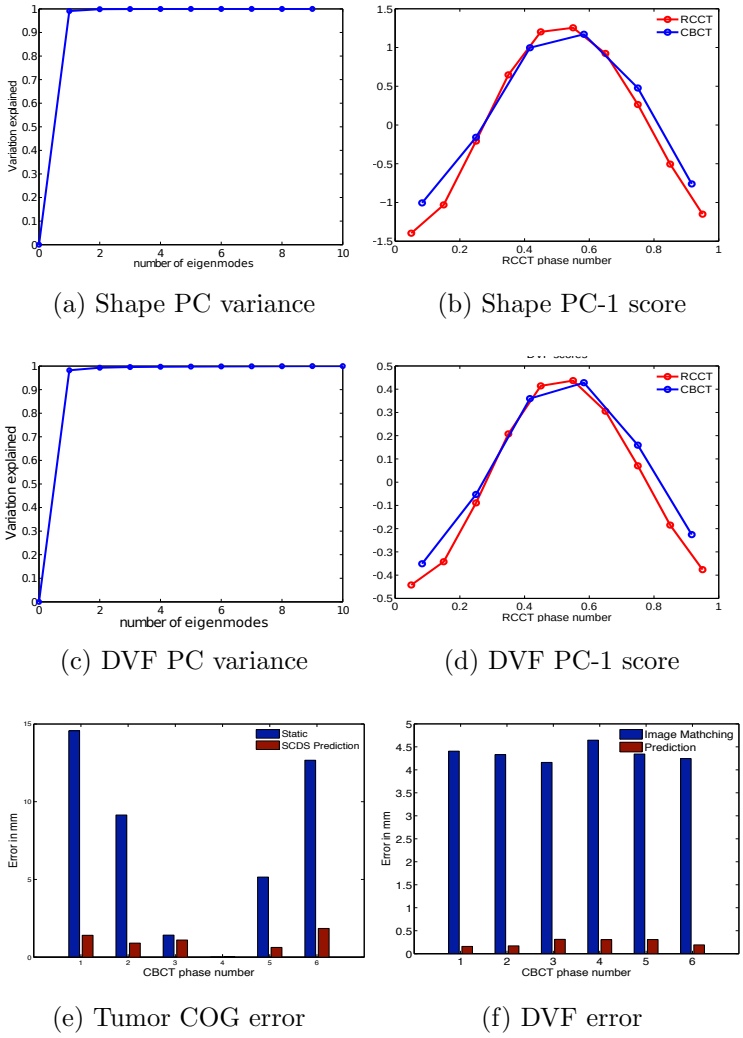
The "shape PC-1 score" plot shows the first PC scores of the extracted surrogate shapes at each phase (of RCCT and CBCT) in the training PCA space. Since the underlying deformation spaces are the same for the computational phantom data, the small discrepancies between the two curves reflect the segmentation errors of the CBCT due to image artifacts. The DVF discrepancies are the direct results of the shape discrepancies via the linear correlation mapping (that maps each shape to its corresponding DVF). The "shape CV-1" plot shows the first canonical variable values (the scores on the maximum correlation direction) of the shapes. Since the correlations between the shape and the deformation fields are close to 1 in the studies, the corresponding "DVF CV-1" plot look almost identical.

For the convenience of understanding the statistics plots in the following experiments, I list the common measurements and their meanings in Table 4.1.

Table 4.1: Statistical measurements reference.

shape PC variance	The cumulative percentage variation of the PCs for the shape.
shape PC-1 score	The first PC score of the shape.
shape PC-2 score	The second PC score of the shape.
DVF PC variance	The cumulative percentage variation of the PCs for the image deformation.
DVF PC-1 score	The first PC score of the image deformation.
DVF PC-2 score	The second PC score of the image deformation.
shape CV-1	The first canonical variable of the shape.
DVF CV-1	The first canonical variable of the image deformation.

Error analysis In both experiments the tumor location errors are reduced significantly. In terms of the overall image deformation errors, the SCDS prediction is shown to be much more effective than the pure-intensity-based image matching approach in the presence of image artifacts. In comparison, the NCAT test 2 in Figure 4.7 contains more remaining error than the NCAT test 1 in Figure 4.6. The main reason is the segmentation error showing in the "shape PCA-1 score" and the "DVF PC-1 score" plots. The maximum of discrepancy of the "DVF PC-1 score" between the RCCT phase and CBCT phase is about 0.1 standard deviation for the NCAT test 2. Another factor is the tightness of the linear statistics. The deformation spaces are more complicated in NCAT test 2 due to two distinctive motions (large mobility in both AP and diaphragm motions). Therefore the PCA statistics is less tight, and it takes two PCs to cover the majority of the variation. Errors or noise kept in the statistics directly affect the prediction results.

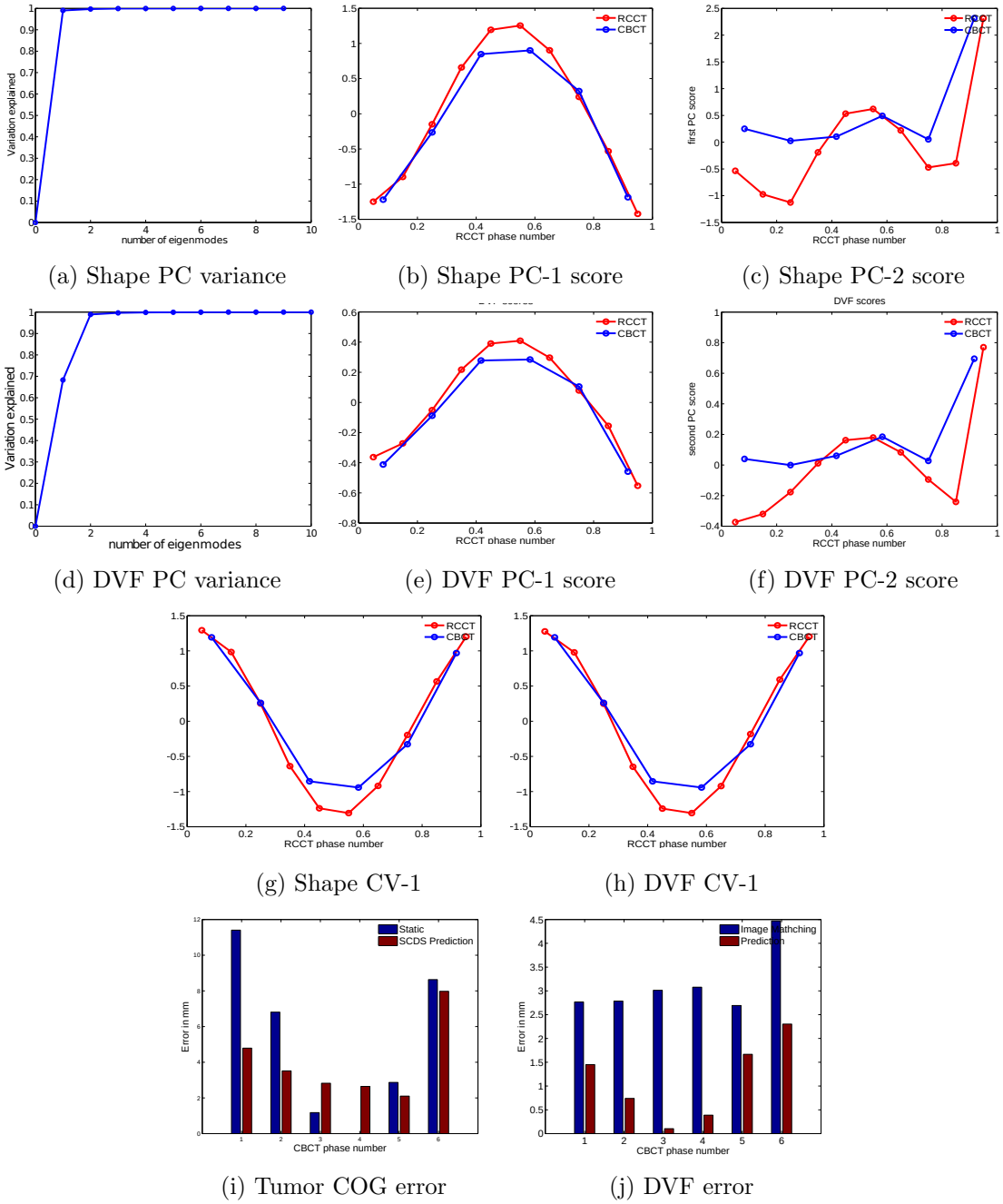


(a) Shape PC variance (b) Shape PC-1 score

(c) DVF PC variance (d) DVF PC-1 score

(e) Tumor COG error (f) DVF error

Figure 4.6: **NCAT CBCT SCDS-prediction error evaluation 1.** This NCAT data set contains 0.5 cm anterior-posterior motion and 2.0 cm diaphragm motion. Only the first PC is used for SCDS calculation. a)-d) Please refer to List 4.1. e) The tumor center of gravity (COG) prediction errors are shown for all CBCT phases. The fourth-phase contour is used as the template contour. The *static* measurement illustrates the mobility of the tumor by using the template contours for all phases. f) The magnitude of the dense deformation prediction error averaged over the voxels in the image. The results from pure intensity-based image matching are measured for comparison.



(a) Shape PC variance (b) Shape PC-1 score (c) Shape PC-2 score

(d) DVF PC variance (e) DVF PC-1 score (f) DVF PC-2 score

(g) Shape CV-1 (h) DVF CV-1

(i) Tumor COG error (j) DVF error

Figure 4.7: **NCAT CBCT SCDS-prediction error evaluation 2.** This NCAT data set contains 1.5 cm anterior-posterior motion and 1.5 cm diaphragm motion. Two PCs are used for the SCDS calculation. a)-h) Please refer to List 4.1. i) The tumor COG prediction errors are shown for all CBCT phases. The fourth-phase contour is used as the template contour. The *static* measurement illustrates the mobility of the tumor by using the template contours for all phases. j) The magnitude of the dense deformation prediction error averaged over the voxels in the image. The results from pure intensity-based image matching are measured for comparison.

4.3.2 Patient data

Respiration-correlated CT (RC-CT) data sets were provided by a 4-slice scanner (LightSpeed GX/i, GE Medical System), acquiring CT images for a complete respiratory cycle at each couch position while recording patient respiration (Real-time Position Management System, Varian Medical Systems). The CT images were retrospectively sorted (GE Advantage 4D) to produce a series of 3D images at different respiratory time points. The CBCT scans were five-minute scans acquired using a gantry-mounted KV on-board imaging system (Varian Medical Systems).

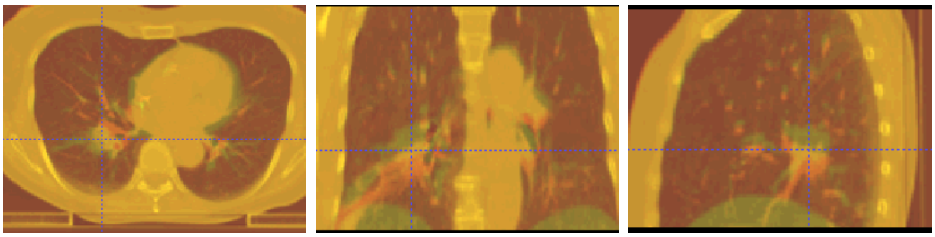
The following experiments evaluated the motion estimation results in terms of tumor location errors. The manual tumor contour of the fourth phase image (EE phase) was propagated to all phases for validation. The tumor contours of all the other phases of the CBCT image sequence were also provided for error measurements. Figure 4.8 shows the results from a patient with a diaphragm mobility up to about 1.5 cm and little anterior-posterior motion. The three views in Figure 4.8a illustrate the maximum discrepancy caused by respiratory motion. The *static* measurement illustrates the mobility of the tumor by using the template contour for all phases. The fourth-phase contour is used as the template contour. Figure 4.9 shows the results from a data set with a tumor region in the mediumstinum. The diaphragm region is out of the imaging field.

Error analysis Both patients show obvious respiratory motions, especially near the diaphragm region. The evaluated tumor region in patient 1 (Figure 4.8) is inside the bottom lung region and moves along the diaphragm motion up to 1.5 cm. The average tumor location errors are decreased by half, from 7.2 mm to 3.1 mm. The tumor region of patient 2 (Figure 4.9) is in the mediastinum region and thus only has about 3 mm maximum mobility. The correction of the location error is therefore marginal.

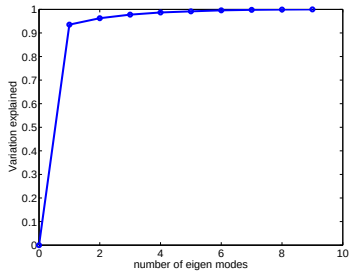
Unlike the NCAT data, in which the underlying motion between the training and testing data are exactly the same, the real patient data contains motion pattern varia-

tion. The discrepancies shown in the "shape PC scores" and the "DVF PC scores" are mainly the true motion variations. For example, patient 1 has a shallower breathing in the CBCT imaging time than the planning RCCT time, which are reflected in the discrepancies of about 0.5 standard deviations in the "DVF PC-1 score".

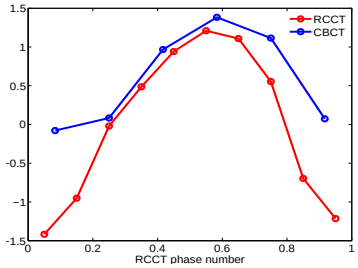
In fact, the correlation parameters between the shape surrogate and the deformation fields can not be exactly the same between the planning time and the treatment time. For example, when the patient breathing with different amplitudes, the interactions between the lung and its surrounding tissues also undergo certain variation accordingly. The non-avoidable correlation variations will therefore contribute to the overall estimation errors.



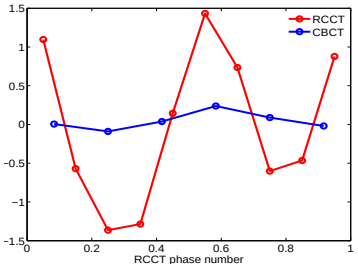
(a) Overlapping RCCTs at EE and EI phase



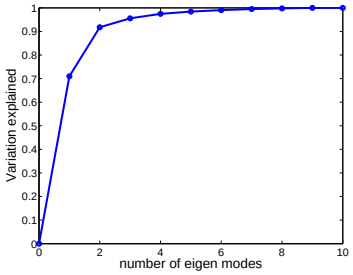
(b) Shape PC variance



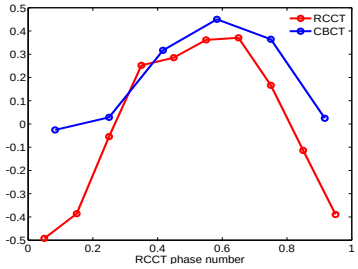
(c) Shape PC-1 score



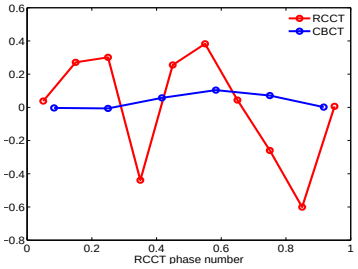
(d) Shape PC-2 score



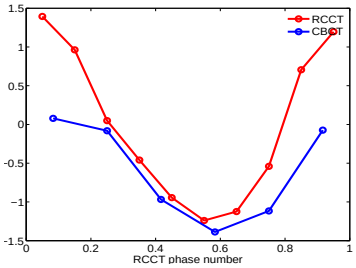
(e) DVF PC variance



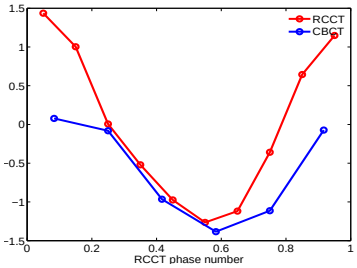
(f) DVF PC-1 score



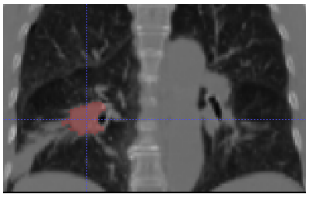
(g) DVF PC-2 score



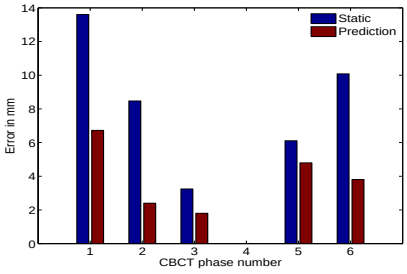
(h) Shape CV-1



(i) DVF CV-1

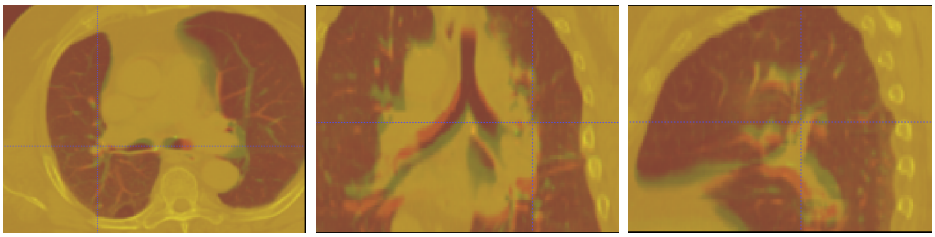


(j) Tumor location

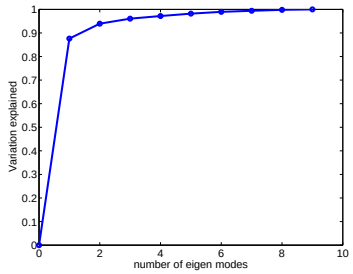


(k) Tumor COG error

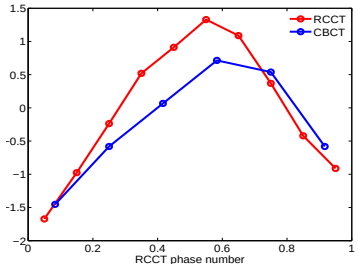
Figure 4.8: **Real patient CBCT SCDS-prediction error evaluation 1.** Two PCs are used for SCDS calculation. a) Three views of the overlapping image of the EE (green) and the EI (red) phases of the training CT. b)-i) Please refer to List 4.1. j) The tumor is located near the diaphragm region of the lung. k) The tumor COG prediction errors are shown at all 6 CBCT phases.



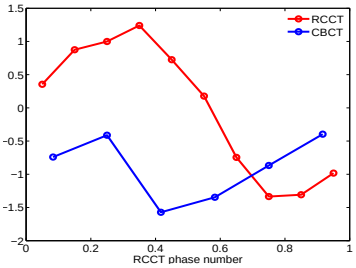
(a) Overlapping CTs of EE and EI phase



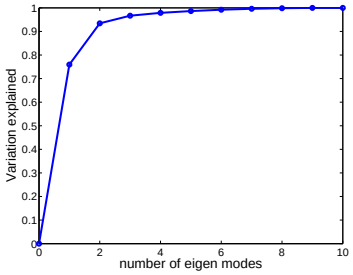
(b) Shape PC variance



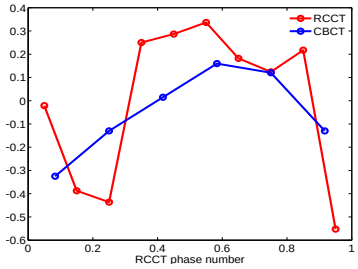
(c) Shape PC-1 score



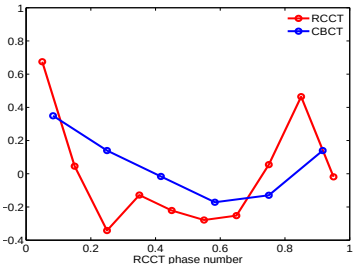
(d) Shape PC-2 score



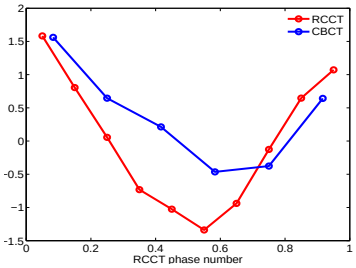
(e) DVF PC variance



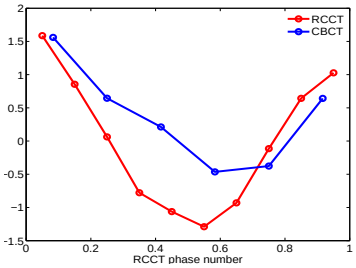
(f) DVF PC-1 score



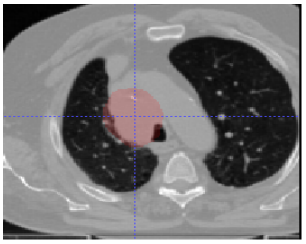
(g) DVF PC-2 score



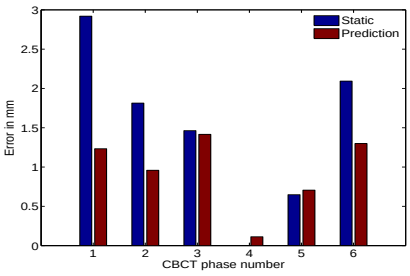
(h) Shape CV-1



(i) DVF CV-1



(j) Tumor location



(k) Tumor COG error

Figure 4.9: **Real patient CBCT SCDS-prediction error evaluation (2).** Two PCs are used for SCDS calculation. a) Three views of the overlapping image of the EE (green) and the EI (red) phases of the training CT. b)-i) Please refer to List 4.1. j) The tumor is located near the diaphragm region of the lung. k) The tumor COG prediction errors are shown at all 6 CBCT phases.

4.4 Conclusion

In this chapter the SCDS trained from the planning CT sequence was applied for respiratory motion prediction for the CBCT image sequence acquired at treatment time. The posterior deformable segmentation method was used to segment the lungs from the noisy images with satisfactory accuracy, which enabled the learning-based strategy to effectively estimate the respiratory motion.

For improving the error margin, the trained SCDS could be made tighter with a better surface correspondence condition of the shape models and better image registration results, and the segmentation errors could be minimized by carefully tuning the segmentation parameters. But the unavoidable breathing pattern variation (the inconsistency between the planning time and the treatment time) is inherently determined by the breathing mechanisms of specific patients. In the following chapter a new motion estimation method that is less sensitive to the inconsistency between the planning data and the treatment data is presented for improved robustness.

Chapter 5

Prediction-driven Respiratory Motion Atlas Formation

5.1 Introduction

In previous chapters (Chapter 3 & 4) the learning-based SCDS model has been introduced and applied for 4D CBCT respiratory motion prediction. The underlying assumption of the model is that the correlations between the surrogate and the image deformation are the same for both planning time and treatment time. The assumption provides the foundation for directly incorporating the training information into the target prediction. However, the correlations can not be exactly the same between the planning time and the treatment time, especially for lung cancer patients who have difficulties in stabilizing their breathing over time. Besides, noise present in the SCDS trained from HDLSS data might result in predictions that are not consistent with the actual anatomical conformations in some local regions. Improvement of the SCDS-based motion estimation can be made by including more intensity information from the target images besides lung boundaries.

Recall that a pure-intensity-based image matching or a non-learning based method has the over-fitting problem when image artifacts are present. However, RC-CBCT images still contain many intensity features that are informative and can be utilized.

Beyond the lung boundary features (used in the surrogate object extraction), other high-contrast intensity features such as the bony rib cage, bronchial structures and the tumor region itself are helpful in estimating the respiratory motion.

To fully utilize both the learning-based prediction method and the intensity-based matching method while avoiding their limitations, a respiratory motion atlas formation method driven by a combination of prediction matching forces and image matching forces is developed in this chapter. On the one hand, the motion prediction can help regularize the intensity matching from over-fitting. On the other hand, meaningful image features can be utilized to reduce prediction errors due to the moderate inconsistency condition. The deformations predicted from the SCDS model are used as soft constraints during the optimization. The balancing force between the prediction force and the intensity force can be adjusted via a weighting factor that is selected upon the credibility of the training statistics and the quality of the treatment images.

The rest of the chapter is organized as follows. Section 5.2 introduces the framework of the proposed prediction-driven deformation atlas formation. Section 5.3 presents the experimental results on simulation data, NCAT data and real patient CBCT data. Section 5.4 concludes this chapter with discussions.

5.2 Methodology

5.2.1 Prediction-driven respiratory motion atlas formation

A prediction-driven deformation atlas formation method has been designed to combine SCDS-prediction constraints and intensity matching forces into a unified framework.

Recall that the original intensity-based Fréchet mean atlas formation (Section 3.2.2), which is used to match the individual planning phase-images to their Fréchet mean atlas,

has the energy function:

$$\hat{J}, \hat{\varphi^i} = \underset{J,\hat{\varphi^i} \in \mathbb{I} \times \mathrm{Diff}_V(\Omega)^N}{\operatorname{argmin}} \sum_1^N \left[\int_0^1 ||v_t^i||_V^2 dt + \frac{1}{\sigma^2} ||I - J^i \circ \varphi^i||_{L2}^2 \right],$$
$$\text{subject to } \varphi_0^i = Id,\ \varphi^i(x) = x + \int_0^1 v_t^i(\varphi_t^i(x)) dt, \tag{5.1}$$

where J^i denotes the noisy treatment image at phase i, $\hat{J}$ is the atlas image and φ^i refers to the image deformation that matches the individual CBCT phase image J^i to the atlas image $\hat{J}$.

Based on this formula, the prediction-driven deformation atlas formation method appends SCDS-prediction term as a soft constraint in this iterative Fréchet mean image optimization as follows:

$$\hat{J}, \hat{\varphi}^i = \underset{J,\hat{\varphi^i} \in \mathbb{I} \times \mathrm{Diff}_V(\Omega)^N}{\operatorname{argmin}} \sum_1^N \left[\int_0^1 ||v_t^i||_V^2 dt + \frac{1}{\sigma_1^2} ||J - J^i \circ \varphi^i||_{L2}^2 + \frac{1}{\sigma_2^2} d_R^2(\varphi^i, \phi(q^i)) \right],$$
$$\text{subject to } \varphi^i = x + \int_0^1 v_t^i dt, \tag{5.2}$$

where $\mathbf{q}^i$ is the lung shape segmented from J^i using the deformable segmentation method introduced in Section 4.2.2. $d_R(\varphi, \phi(q^i))$ is the distance between the varying deformation φ and the prediction $\phi(q^i)$. It is measured via the Riemannian manifold metric d_R, which is defined by

$$d_R(\psi_1, \psi_2) = \underset{\gamma:[0,1]->\mathcal{M}, \gamma(0)=\psi_1, \gamma(1)=\psi_2}{inf} \int_0^1 \sqrt{< \dot{\gamma(t)}, \dot{\gamma(t)} >_V \, dt}, \tag{5.3}$$

where the Riemannian distance between two points ψ_1 and ψ_2 on $\mathcal{M}$ is defined as the infimum of this integral over all piecewise smooth curves γ that connect ψ_1 and ψ_2. This distance can be alternatively computed by $d_R(\psi_1 \circ \psi_2^{-1}, id)$, where id is the identity transformation.

To directly use the resulting deformation represented by DVFs computed from the

SCDS-prediction, a Euclidean approximation of the Riemannian distance is given by

$$d_R^2(\varphi^i, \phi(q^i)) \approx ||\mathbf{u}_{\varphi^i} - \mathbf{u}_{\phi(\mathbf{q}^i)}||_{L^2}^2, \tag{5.4}$$

where $\mathbf{u}_\psi$ denotes the dense DVFs resulting from the diffeomorphic transformation ψ, and the SCDS-predicted deformation $\mathbf{u}_{\phi(\mathbf{q}^i)}$ is computed by linear mapping $\mathbf{u}_{\phi(\mathbf{q}^i)} = \mathbf{C} \cdot \mathbf{q}^i$. The Euclidean deformation space is only an approximation of the Riemannian deformation manifold. However, when deformations are not very large, the Euclidean space can be thought of as the tangent plane of the Riemannian manifold at the Fréchet mean, so the linear approximation is sensible.

The balancing force between the prediction and the noisy intensity profile can be adjusted via the weighting factors σ_1 and σ_2, selected upon the credibility of the training statistics and the quality of the treatment images. Normally the weighting factors should make the two forces have the same order of magnitude. Built on top of the original intensity-based atlas formation method, the new energy term of the prediction is treated as an extra feature channel.

The prediction-driven motion atlas optimization is summarized in Algorithm 5.1.

In the implementation of integrating the SCDS-prediction constraints to the original intensity-based atlas formation framework, a DVF vector-channel is simply appended to the intensity channel of each phase image. Since this DVF constraint itself is a three-dimensional vector that measures the displacements in the three coordinate directions, it takes three times the memory space as the intensity-based atlas formation. The optimization steps are exactly the same.

Algorithm 5.1 Prediction-driven atlas formation algorithm.

Input: A sequence of N-phase CBCT respiratory images: $\{J_i\}_{i=1}^N$ and their corresponding predicted deformations $\{\mathbf{u}_{\phi(\mathbf{q}^i)}\}_{i=1}^N$ (predicted from the SCDS of planning CT images)

Output: A CBCT atlas (Fréchet mean) image $\hat{J}$ and a collection of diffeomorphic deformations $\varphi_{i=1}^N$ that transform J_i to $\hat{J}$.

Initialize deformation with identity

1: **for** i=1:N **do**
2: $\varphi^i \leftarrow Id_{\mathrm{Diff}_V(\Omega)}$
3: **end for**
4: **repeat**
5: $\hat{J} = \frac{1}{N}\sum_{i=1}^N J^i_{\varphi^i}$
6: **for** i=1:N **do**
7: $v^{int}_{inc} \leftarrow K(\frac{2}{\sigma_1^2}\nabla J^i_{\varphi^i_{0,1}}(J^i_{\varphi^i_{0,1}} - \hat{J}))$
8: $v^{pred}_{inc} \leftarrow K(\frac{2}{\sigma_2^2}(\nabla \mathbf{u}^x_{\varphi^i}(\mathbf{u}^x_{\varphi^i} - \mathbf{u}^x_{\phi(\mathbf{q}^i)}) + \nabla \mathbf{u}^y_{\varphi^i}(\mathbf{u}^y_{\varphi^i} - \mathbf{u}^y_{\phi(\mathbf{q}^i)}) + \nabla \mathbf{u}^z_{\varphi^i}(\mathbf{u}^z_{\varphi^i} - \mathbf{u}^z_{\phi(\mathbf{q}^i)})))$
9: $v_{inc} \leftarrow v^{int}_{inc} + v^{pred}_{inc}$
10: $\varphi^i \leftarrow Exp_{\varphi_i}(\epsilon v_{inc})$
11: **end for**
12: **until** convergence

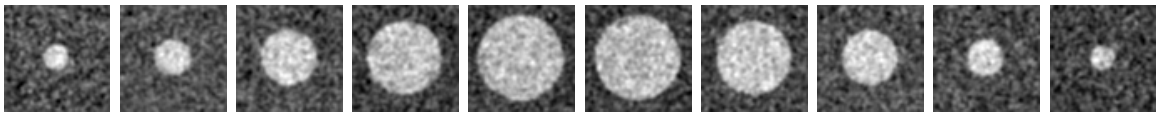
5.3 Experimental Results

5.3.1 Breathing spheres

To test the prediction-driven atlas formation method, I started with the same simulated breathing spheres data used in Chapter 3 (see Figure 3.8). The sequence of sphere images with varying radius were designed to mimic the breathing scenario. The radii follow a sinusoidal curve to simulate the breathing pattern of a volume enlarging process followed by a volume shrinking (see Figure 3.8 for illustrations of the Euclidean mean image and the Fréchet mean image). To simulate image artifacts, independent Gaussian noise with 0.0 mean and 0.25 standard deviation was added to each voxel of the binary image, which has 0.0 as the background intensity value and 1.0 as the foreground intensity value. The results are shown in Figure 5.1a.

Given the noisy breathing sphere images, I first compared the effectiveness of the

SCDS-prediction method to the intensity-based atlas formation method. Since the only difference between the training data and test data is the added Gaussian noise, the underlying linear correlations of the two are identical. Figure 5.1c shows that the intensity-based atlas formation method tends to over-fit the noise, while the SCDS prediction (Figure 5.1d) is only determined by the training statistics.



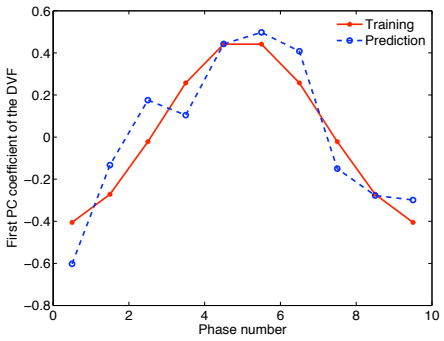
(a) noisy input images

(b) Euclidean mean (c) Fréchet mean (d) SCDS-prediction atlas

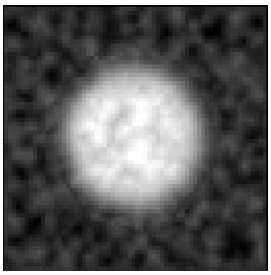
Figure 5.1: **Noisy breathing spheres test**: a) 2D slices of the ten input noisy breathing sphere images. b) The Euclidean intensity mean of the noisy test sequence image. c) The Fréchet mean atlas image of the test data using pure-intensity-based atlas formation. d) The resulting atlas image of the test data using the SCDS-prediction.

To simulate inconsistent correlation patterns between the training set and the test set, the correlation regression parameters obtained from the training data were perturbed slightly before they were used for predicting the motion in the test data. The perturbation amount added to each regression parameter was randomly generated by uniform distribution selected between -20% and 20% of the parameter value by a random number generator. As shown in Figure 5.2a, the resulting "DVF-PC1 score" has a maximum 0.2 standard deviation difference. The results of three methods are shown in Figure 5.2. The SCDS prediction is affected by artificial correlation perturbation and produces errors mostly visible at the edge of the atlas sphere. The intensity-only Fréchet

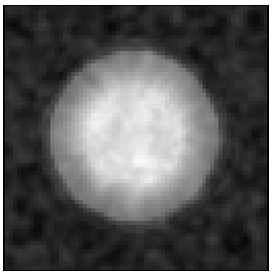
mean atlas (the same image shown in Figure 5.1c) is placed beside for convenience of comparisons. It is shown that the prediction-driven deformation has the best visual results among the three. The signal-to-noise ratio both inside and outside the sphere objects appear to be higher than the intensity atlas, and the edge of the sphere does not contain the thick-ring artifact shown in the SCDS-prediction atlas.



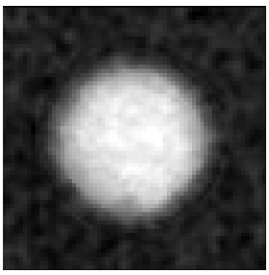
(a) Perturbed DVF PC-1 score



(b) Intensity



(c) SCDS-prediction

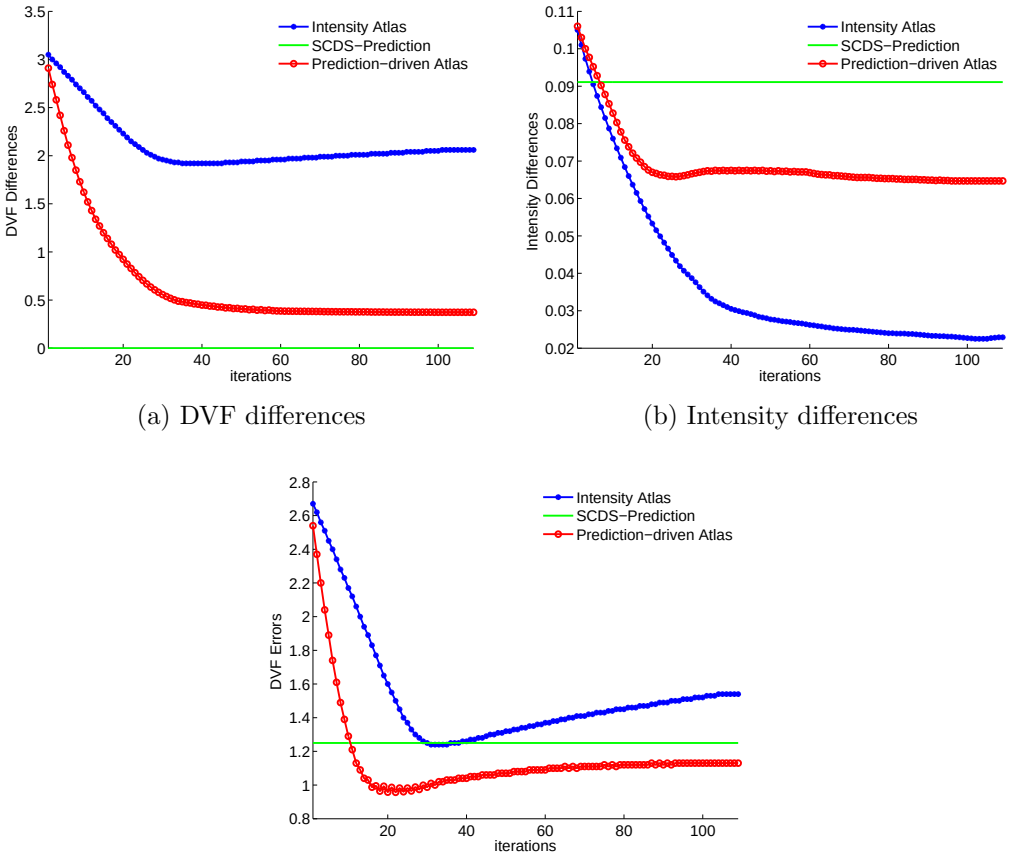


(d) Prediction-driven

Figure 5.2: **Noisy breathing spheres test results with perturbed correlation.** a) The resulting perturbation on the predicted deformations measured in terms of the first principal component coefficients in its training deformation space. b) The intensity-based Fréchet mean atlas image. c) The atlas image from the SCDS-predicted deformations. d) The prediction-driven atlas image.

The SCDS-prediction forces, the intensity-matching forces and the deformation estimation errors for the same data are shown at each iteration step in Figure 5.3 for detailed investigation. Since the SCDS-prediction is directly computed without iterative optimizations, it is shown as the constant value line for comparison. Figure 5.3a and Figure 5.3b show that the prediction-driven atlas converges to an intermediate level

between the intensity atlas and the SCDS-prediction in both energy measurements. The best estimation result is achieved by the prediction-driven atlas formation method as shown in Figure 5.3a.



(a) DVF differences

(b) Intensity differences

(c) DVF Errors

Figure 5.3: **Noisy breathing spheres test with perturbed correlation.** Comparison curves of the optimization iterations. a) Average DVF differences between the optimizing DVF and the predicted DVF. b) Average image intensity force at each iteration. The intensity-based matching method is shown to over-fit the noise. c) Average DVF errors at each iteration, with the ground truth deformations calculated from the training images.

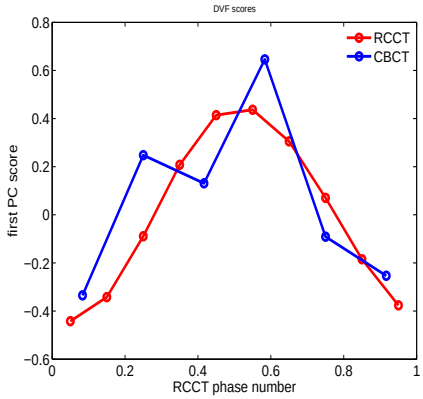
5.3.2 NCAT data

The same NCAT data sets used in Section 4.3.1 were used to test the prediction-driven atlas formation method. Since the NCAT CBCT data was reconstructed from the

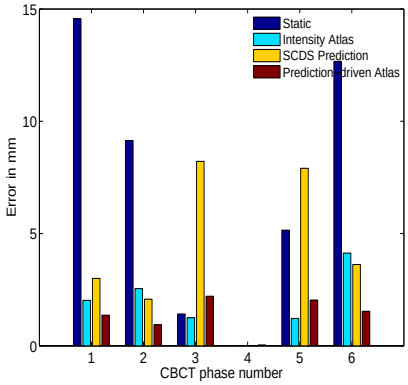
RCCT, the correlation consistency between the CBCT and RCCT is guaranteed. As shown in the results of Section 4.3.1, the SCDS-prediction results are satisfactory. In order to test the prediction-driven atlas formation method in this chapter, inconsistent breathing correlation patterns between the training data and the test data were simulated by adding perturbations to the correlation regression parameters with a random change in value up to 20% (the same perturbation scheme used for the breathing sphere data in Figure 5.2) .

The NCAT data in Figure 5.4 has a maximum of 2.0 cm diaphragm motion and 0.5 cm anterior-posterior motion. The intensity-based atlas formation results, the SCDS motion prediction results and the prediction-driven atlas formation results are compared in terms of the center of gravity (COG) location errors of the tumor region, see Figure 5.4b. Also the overall DVF errors of the three methods are compared in Figure 5.4c. Figure 5.5 illustrates the resulting atlas images from all three methods.

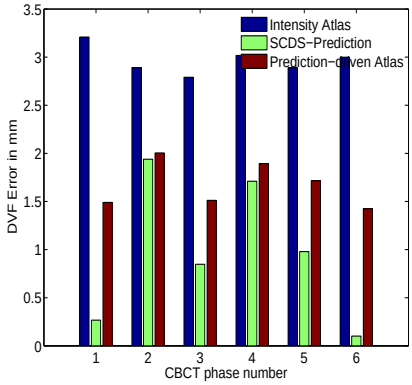
In Figure 5.4b, the intensity-based atlas does a much better job than the SCDS-prediction in terms of the tumor region estimation. The reason is that the tumor region after the CBCT reconstruction still has a quite strong contrast respect to its surrounding tissue despite the global streak artifacts. In real patient CBCT images, less contrast on the tumors is expected. On the other hand, in terms of the overall deformation field estimation the SCDS-prediction performs better than the intensity-based atlas method, as shown in Figure 5.4c. The prediction-driven atlas method shows a balanced overall DVF estimation between the two as a result of the combination. These measurements are confirmed from the visual comparison of the atlas images in Figure 5.5. The prediction-driven atlas (Figure 5.5d) has less global signal-to-noise ratio (SNR) compared to the intensity atlas (Figure 5.5b), and it has a sharper tumor boundary than the SCDS-prediction atlas (Figure 5.5c).



(a) Perturbed DVF PC-1 score



(b) Tumor COG error



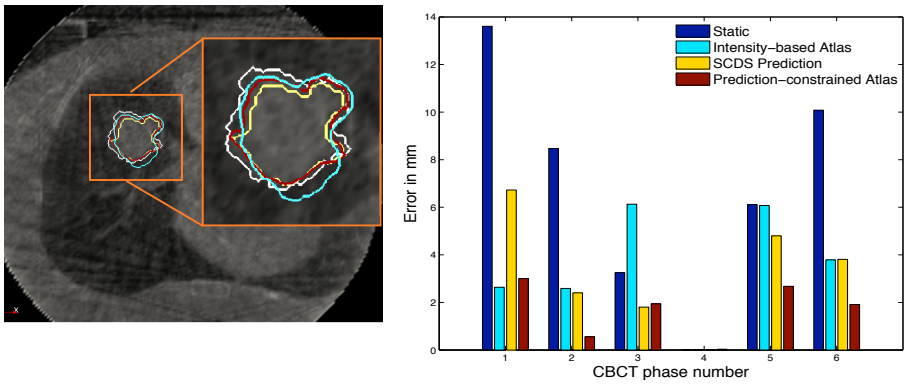
(c) DVF error

Figure 5.4: **Tumor evaluation of prediction-driven atlas formation method on NCAT (1).** a) CBCT prediction contains artificial perturbations to simulate the inconsistency. b) Three methods are compared in terms of the tumor COG estimation error, with the static measurement indicating the mobility of the tumor. The fourth phase is used as the base phase to propagate the tumor contour to other phases. c) Three methods are compared in terms of the average DVF estimation error per voxel; the ground truth DVF is obtained by linear interpolation from the RCCT DVFs.

5.3.3 Patient data

The same patient data sets used in Chapter 4 were used to test the prediction-driven atlas formation method, with comparisons to the intensity-based atlas formation method and the SCDS-prediction method. Motion estimation results were evaluated on a mock tumor region, as shown in Figure 5.6a. Manual segmentations were provided for each CBCT image for error measurements. The manual tumor contour of the fourth phase image (the end-expiration phase) was propagated to all the other phases. The three

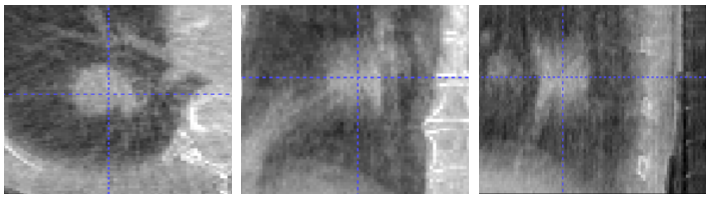
methods are compared in terms of the COG location errors in Figure 5.6b. The average COG errors of the 5 phases for the three approaches are 3.5 mm (Intensity-based atlas), 2.3 mm (SCDS prediction) and 1.7 mm (prediction-driven atlas).



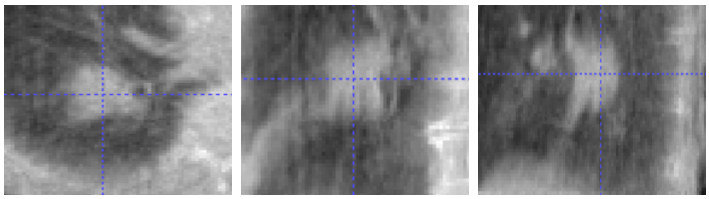
(a) Tumor illustration (b) Tumor location error

Figure 5.6: **Tumor evaluation of prediction-driven atlas formation method on real patient images (1).** Three methods are compared in terms of the COG estimation error, with the static measurement indicating the mobility of the tumor. a) Axial slices of the estimated tumor contours at the fifth phase from the three methods, with the same colors tagged in the bar figure on the right. The manual segmentation of the tumor contours is shown in white. b) The tumor COG errors for the four methods. The fourth phase is used as the base phase.

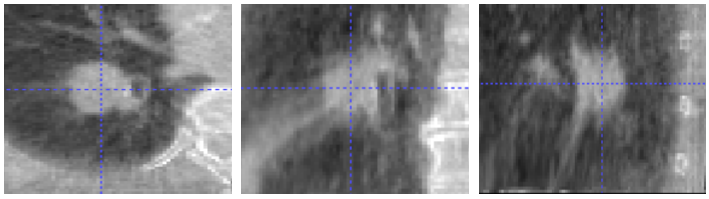
The prediction-driven atlas method outperforms both the other two approaches in half of the phases. Figure 5.7 illustrates the resulting atlas image near the tumor region from the three methods.



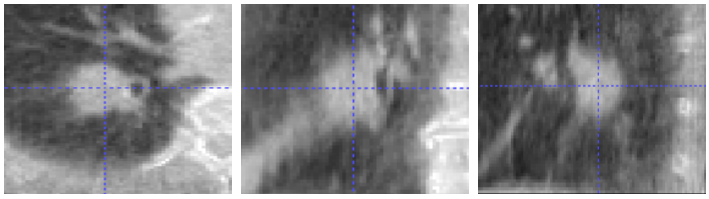
(a) Intensity Euclidean Mean



(b) Intensity Atlas



(c) SCDS-prediction Atlas



(d) Prediction-driven Atlas

Figure 5.7: **Real patient CBCT tumor region in atlas image: a comparison.** a) The Euclidean mean of the CBCT sequence. b) The Fréchet mean atlas image of the CBCT sequence. c) The atlas image from the SCDS-prediction results. d) The atlas image from the prediction-driven atlas formation method.

Another patient's results are given in Figure 5.8. This patient has relatively stable tumor motion with a maximum of 3 mm mobility. The SCDS-prediction reasonably predicted the small motion, while the intensity-based atlas method was greatly affected by the intensity artifacts. The intensity-prediction-combined optimization provided the best estimation results for 4 out of the 6 phases. For the voxel size of 1.52 mm $\times$ 1.52 mm $\times$ 1.52 mm, the SCDS-prediction achieved sub-voxel accuracy. The prediction-

driven atlas is shown to be stable with marginal improvement. The tumor contour in the fourth phase was used as the base contour to propagate to the other phases through Fréchet mean. The errors showed in the fourth phase was caused by the interpolation errors during the inverse transform.

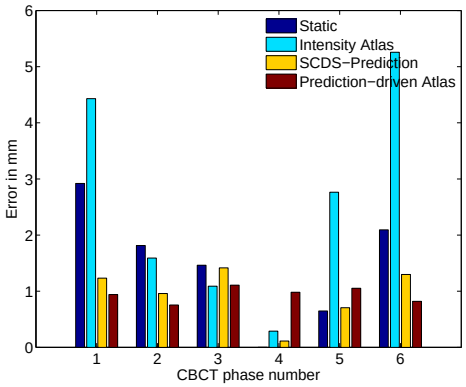
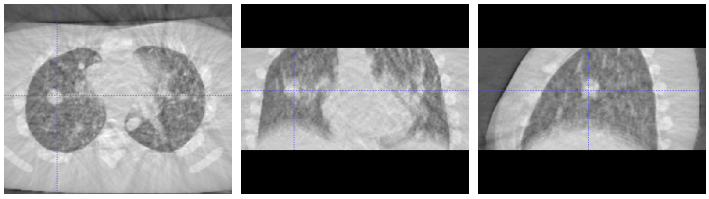


Figure 5.8: **Tumor evaluation of prediction-driven atlas formation method on real patient images (2).** Three methods are compared in terms of the COG estimation errors, with the static measurement indicating the mobility of the tumor. The tumor COG errors for the four methods. The fourth phase is used as the base phase.

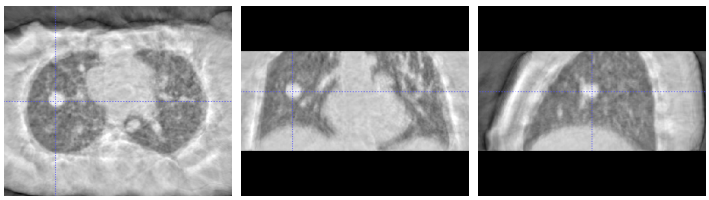
5.4 Conclusion

The prediction-driven atlas formation framework has the advantage of utilizing high-contrast intensity information from the target images while constrained by the shape-correlated prediction results. The overall image deformation is balanced between the intensity-driven deformation and the SCDS-predicted deformation. The anatomical structures that have high intensity contrasts contribute to a more accurate local motion estimations.

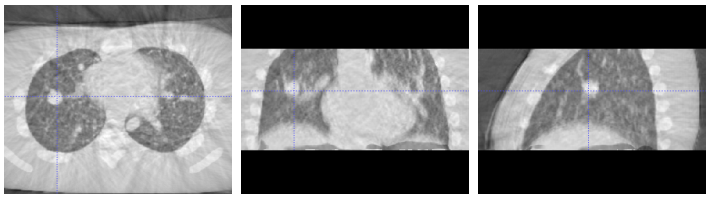
Based on the preliminary studies on both simulated phantom data and RC-CBCT data, the prediction-driven atlas method was shown to be more robust for modeling and estimating sophisticated respiratory motion in lung than both the intensity-based atlas method and the learning-based SCDS-prediction method. More patient studies are needed to fully validate and quantify the robustness of the method.



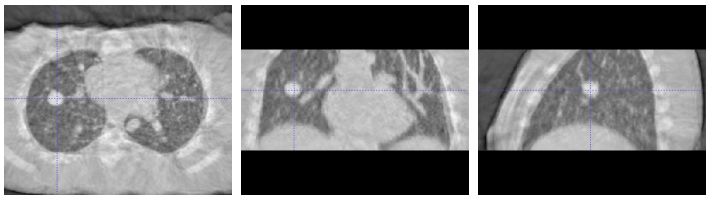
(a) Intensity Mean (Euclidean mean)



(b) Intensity Atlas (Fréchet mean)



(c) SCDS-Prediction Atlas



(d) Prediction-driven Atlas

Figure 5.5: **Atlas image comparison on NCAT data set.** a) The Euclidean mean of the CBCT sequence. b) The Fréchet mean atlas image of the CBCT sequence. c) The atlas image from the SCDS-prediction results. d) The atlas image from the prediction-driven atlas formation method.

Chapter 6

Discussion

6.1 Summary of Contributions

In this section the thesis claims presented in Chapter 1 are revisited. Following each claim is a discussion that relates to the claim and summarizes how that claim is addressed in this dissertation.

The contributions of this dissertation are as follows:

1. *I use shape of the lung as the surrogate for respiratory motion estimation in the thorax. The high contrast of the air-filled lung regions provides comprehensive motion evidence.*

The notion of "surrogate" has been used in two different contexts of 4D IGRT applications. In retrospective imaging (introduced in Section 2.2.2), surrogates such as the marker position in the RPM system (Figure 2.4) are used to provide the respiration-phase gating information to sort the projection images for 3D reconstruction. In surrogate-correlated respiratory motion modeling discussed throughout the dissertation, surrogates are the group of parameters used to model the respiratory motion. Given a 4D (phase-stamped) image sequence, the modeling surrogates are usually a few internal or external landmarks, such as the diaphragm apex points used in Zhang et al. (2007) and the skin

markers used in Gao et al. (2008). The positional changes of the landmarks are correlated with the whole image deformation field. In this dissertation the lung boundary is used as the surrogate object that is shown to be more effective in modeling the respiratory motion than sparse local landmarks. A surrogate comparison study is given in Section 3.3.3.

2. *I apply an entropy-based particle system to 4D lung CT images to obtain surface lung models with group-wise correspondence. The surface correspondence across the 4D sequence is crucial for statistical shape modeling.*

In order to extract tight statistics from the shape variation during a respiratory cycle, good group-wise geometrical correspondence is required. A state-of-the-art entropy-based particle system has been adopted for this purpose in Section 3.2.1. For complicated anatomical motion variations, the non-parameterized point distribution model (PDM) representation is desired. To be more specific, when modeling lung surfaces, it is often useful to include the bronchial structure that connects the left lung and the right lung due to its high-contrast intensity profile. Especially when the diaphragm region is not included in the imaging field, spatial changes of the bronchial structure usually provide the strongest motion evidence. The particle system can easily incorporate arbitrary topology of the shape such as the highly-curved tubular surfaces of the bronchial structures and can provide credible point sampling on the bronchial surface without reparamterization.

Another advantage of the surface particle system is its ability to incorporate new features or attributes of the sampling points. The common features used to distribute the surface sampling points are the x, y and z coordinates. In the application of cerebral cortical surface modeling (Oguz et al., 2008), local curvature features such as Gaussian and mean curvatures have been shown to be useful for improving the group-wise correspondence. In the experiments with the lung models, the positional features seem to be sufficient to produce good correspondence results; this sufficiency is validated by the

tightness of the principal component analysis (PCA) results. Usually the first principal component captures more than 90% of the total variation (see Figure 3.9). Also, by visual inspection the principal component intuitively reflects the respiratory motion (see Figure 3.10). Nevertheless, it is an interesting research topic to explore other geometrical features that might further improve the tightness of the statistics.

3. *I apply the Fréchet mean image formation method to generate the respiratory motion atlas from respiration-correlated CT images. The motion atlas contains an atlas image and the transformations that match each individual phase image to the atlas.*

A respiratory motion atlas for a sequence of respiratory phase images contains an atlas/reference image and the transformations that match each individual phase image to the atlas image. Usually the end expiration (EE) phase image has been chosen to be the atlas image due to its stability in the breathing cycle. However, from the statistical point of view, an atlas image with "averaged" geometry gives the tightest probability distribution of the deformations. The Fréchet mean atlas formation method introduced in Section 3.2.2 has the property of requiring the least amount of total deformations. A comparison in Figure 3.12 demonstrates the improved statistics by using the Fréchet mean atlas image [???] . Since the statistics of the deformation fields is a crucial part of the statistical modeling of the respiratory motion, this improvement in the tightness of the statistics matters.

Computationally the Fréchet mean atlas formation process is intensive both in memory consumption and computation time. The dominating computation of each optimization iteration is a Fast Fourier Transform. The order of the algorithm is $MNnlogn$, where M is the number of iterations, N is the number of images, and n is the number of voxels along the largest dimension of the images. A multi-threaded C++ implementation on an eight-processor system takes about 20 minutes to compute a sequence of 10 images of the size $196 \times 196 \times 100$ and requires at least about 1 GB memory. A graphics

programming unit (GPU) implementation could potentially accelerate the computation by a factor of 50 (Muyan-Ozcelik et al., 2008; Shams et al., 2010).

The Fréchet mean atlas formation is used for quantifying the dense spatial variations in planning RCCT images. The resulting displacement vector fields are used as the input to compute the shape-correlated deformation statistics SCDS (that are discussed in the next list item). The distance metric used in the Fréchet mean image formation framework is measured on the Riemannian manifold of the diffeomorphic transformations. A fluid-flow based diffeomorphic image matching method (Section 2.3) is used to bring input images to the atlas while the atlas is updated at each iteration.

It is ideal to save the whole geodesic path of the deformations to characterize the diffeomorphisms completely. However, the need for memory is too high for that to be practical. Therefore, only the displacement vector fields (DVF) are recorded as the final image deformation results. For typical breathing motion, the maximum spatial variation for a voxel is normally less than 2 cm or 10 voxel units. Therefore, the DVF is typically a good approximation of the actual geodesic path. To pursue real-time performance, it might be sufficient to use an elastic image matching scheme (Holden, 2008) rather than the diffeomorphic one.

4. *I present shape-correlated deformation statistics (SCDS) to capture correlations between the shape deformation and the dense image deformation fields from the RCCT image sequence.*

Shape-correlated deformation statistics (SCDS) is a novel concept (introduced in Section 3.2.3) that contains the statistics of shape variation, the statistics of dense deformation field variation, and the correlation between the two. To predict motions from artifact-laden treatment images, SCDS is extracted from the planning RCCT images that have high image quality in terms of the contrast to noise ratio (CNT). In computing the SCDS, the shape surrogate is represented by the surface point sets resulting from the entropy-based particle system, and the dense deformation field is represented by the DVF

resulting from the Fréchet mean formation. Both representations are high dimensional, on the order of thousands and millions, respectively. In order to carry out correlation analysis, PCA is used to reduce their dimensions (Section 3.2.3). In the experiments carried out in the thesis, two PCs are shown to be sufficient to capture the majority of the variation for both spaces.

After the PCA reparameterization, linear correlation analysis is carried out to model the low-dimensional image deformation as a function of the low-dimensional shape surrogate. It was shown in Section 3.2.3 that after removing the multicollinearity via PCA, all three correlation methods: canonical correlation analysis (CCA), multilinear regression (MLR) and partial least squares (PLS), deliver the same correlation results under a unified eigenproblem formulation. The CCA implementation has been used due to its direct interpretation of the maximum correlations. All experiments in this thesis reported maximum correlation coefficients that are close to 1.0, indicating a strong linear correlation between the two spaces.

In the SCDS-prediction, phase-stamped training samples for both the shape and the DVF are assumed to have Gaussian distributions in order to use PCA dimension reduction. Also, the correlation between the two variation spaces is assumed to be linear. The linear assumption greatly simplifies the motion modeling. However, the complicated interactions among the anatomical organs are non-linear in nature. The Gaussian distribution assumption and linear approximations therefore may eventually limit the motion estimation accuracy.

5. *I apply patient-specific SCDS trained from planning CT images to predict the respiratory motion from cone-beam CT images that contain artifacts due to image reconstructions with sparse projections.*

In order to apply SCDS trained from planning images to treatment-guidance images of the same patient, accurately extracting the surrogate shape from the treatment images is crucial. To account for serious RC-CBCT image artifacts, a probabilistic

deformable segmentation scheme (maximum-a-posteriori probability estimation) is developed in Section 4.2.2. The shape prior is provided by the PCA on the shape models from the planning images. The model-to-image-match is evaluated on edge-enhanced second-order gradient images. The shape prior offers a good initial model (the mean model from training) and prevents the segmentation process from overfitting the image artifacts. Depending on how artifacts interfere with the lung boundary, segmentation errors could occur in local regions where artifacts dominate. Those segmentation errors will directly result in prediction errors via the linear mapping, which maps the segmented shape to its corresponding image deformation.

A series of experiments on respiration-correlated CBCT images from both computational phantoms (NCAT) and patients were carried out in Section 4.3. The results demonstrated the effectiveness of the SCDS-prediction in terms of tumor location accuracy and overall deformation field estimation accuracy. In the NCAT experiments, the RC-CBCT sequence was reconstructed from the simulated RCCT projection images. Therefore, the target RC-CBCT images and the training RCCT images have the exact same respiratory motion inherently. The prediction errors are therefore mainly caused by CBCT segmentation errors due to the artifacts. In the patient experiments, another error source besides the CBCT segmentation errors is the inconsistent motion patterns between the training (RCCT) data and the target (RC-CBCT) data. To be more specific, the correlation patterns between the shape and the image deformation might change. As a result, the correlation parameters trained from the RCCT data may not fit RC-CBCT well enough for predictions.

6. *I present a prediction-driven atlas formation method that combines SCDS-prediction constraints and intensity-matching forces into a unified framework for robust respiratory motion estimation.*

In order to reduce the SCDS-prediction method's sensitivity to the correlation inconsistency, as well as to compensate statistical errors and segmentation errors, a prediction-

driven atlas formation method is presented in Chapter 5 towards robust respiratory motion estimation. In this method the SCDS-prediction results are used as soft constraints in a Fréchet mean atlas formation framework. By including intensity information (besides the shape surrogate) from the target images, useful high contrast features such as the vascular structures, the rib cage and the tumor itself are utilized for improved local estimation accuracy. Based on the relative credibility of the image quality and the correlation consistency, the weights between the intensity forces and the prediction forces can be adjusted accordingly.

However, it is difficult to quantify the credibility, which makes the weight tuning mostly empirical. Here are some heuristics to follows. Assuming the signal-to-noise ratio of the image intensity is stable, if the motion patterns of the RCCT and the RC-CBCT data observed from their principal component scores (after projection onto the PCA space) are highly consistent, the weight on the SCDS-prediction force should be higher than the intensity-matching forces. On the other hand, if large discrepancies in motion patterns are observed, the weight on the SCDS-prediction force should be lower.

Experimental results on both the NCAT data and the patient data (Section 5.3) showed improved performance using the prediction-driven atlas formation method, compared to the pure intensity-based atlas method and the SCDS-prediction method. Computationally, in comparison to the intensity-based atlas method, the prediction-driven atlas method takes more time to compute the gradient for the extra prediction constraints. On the other hand, because of the prediction constraints fewer iterations are typically required for convergence.

Thesis: *To account for the motion effects on the radiation dose to tumors and organs at risk in image-guided radiotherapy for lung cancer, shape-correlated deformation statistics (SCDS) trained from planning CT images can be used to effectively estimate the respiratory motion in artifact-laden treatment images, when breathing patterns are consistent between planning time and treatment time. A more robust motion estimation*

can be achieved by combining the SCDS-prediction constraints and intensity-matching forces into a unified atlas formation framework.

Respiratory motion challenges image-guided radiation therapy (IGRT) with location uncertainties of important anatomical structures in the thorax. Effective and accurate respiration estimation is crucial to account for the motion effects on the radiation dose to tumors and organs at risk. Moreover, severe image artifacts in treatment-guidance images such as 4D cone-beam CT cause difficulties in estimating the deformation. Commonly used non-linear dense image matching methods easily fail in regions where artifacts interfere.

Learning-based linear motion modeling techniques have the advantage that training knowledge can be introduced to predict the motion. This dissertation has proposed a mathematical system for effectively extracting the shape-correlated deformation statistics (SCDS) from 4D images. SCDS obtained from the RCCT planning images efficiently captures the variation correlations between the lung shape and the overall image deformation (Chapter 3). It is then used to predict the respiratory motion for treatment time via the lung shape segmented from the RC-CBCT treatment images (Chapter 4). To increase the robustness of the SCDS-prediction, an SCDS-prediction-driven atlas formation framework has been presented. The strategy of balancing between the prediction constraints and the intensity-matching forces makes the method less sensitive to variation in the correlation and utilizes intensity information besides the lung boundaries (Chapter 5).

6.2 Extended Applications

Validation with more patient data

In this thesis the RC-CBCT patient images used for experiments are five-minute-duration scans that are acquired using a slowly rotating gantry-mounted KV on-board imag-

ing system, which are used mainly for research purposes. One-minute scans are more practical within patient treatments for reduced scan time and radiation dose. Those one-minute scans contain more streak artifacts and more blurring due to fewer projections used for 3D reconstructions (Jaffray et al., 2002; Li & Xing, 2007). Validations of the SCDS-based methods on one-minute RC-CBCTs are going to be carried out in the future.

Real-time respiratory motion prediction

The ultimate goal for this respiratory motion estimation is to apply it to on-line treatment guidance immediately prior to dose delivery. Therefore, real-time computation is required. A great computational advantage of learning-based SCDS-prediction (Chapter 4) is that all the statistics calculations are carried out off-line. Once the prior statistics is ready, the on-line prediction during the treatment can be done through a correlation regression linear mapping (matrix multiplication), which is $O(1)$ in computation complexity when computed for each voxel in parallel using a GPU implementation. As for the prediction-driven atlas formation method (Chapter 5), a GPU-accelerated implementation could reduce the computation time by a factor of 50. For 10 input images with image size of $196 \times 196 \times 100$, the atlas optimization therefore could presumably converge within a minute.

SCDS for other applications

Although SCDS is designed for CBCT image-guidance radiation therapy in this thesis, the mathematical framework itself is widely applicable to respiratory motion estimations using other image modalities. For example, MR imaging has been used for respiratory motion monitoring in the lung (Gao et al., 2008). Without the imaging radiation dose limit, a longer imaging time of MRI is allowed to acquire 4D data of multiple respiratory cycles. The increased number of training samples should be able to increase the accuracy

and the robustness of the SCDS. Similar to CBCT, nanotube stationary tomosynthesis (NST) images will have limited usability due to the blurring artifacts (see Figure 1.2) caused by limited angle reconstruction (Chang et al., 2009; Chou et al., 2010). SCDS trained from planning CT images can directly apply to NST-guidance images in the lung for respiratory motion estimation. The lung boundaries could be segmented from the NST images using the same probabilistic deformable segmentation scheme. The model-to-image-match measurements other than the second-order gradients might need to be designed according to the intensity patterns in NST.

In this thesis SCDS-prediction is applied to thorax images, and the lung is the shape surrogate object. Respiratory motion estimation in the abdomen is also an important and active research field (Rohlfing et al., 2004; Guckenberger et al., 2008). Due to the relatively low contrast of the liver tissue (compared with lung), it could be challenging to accurately segment the liver surface from the images. An easier surrogate object could be the surface of the diaphragm on top of the upper liver part. Since the entropy-based particle system (Section 3.2.1) supports arbitrary shape topology, it can be directly used to extract the open surface of the diaphragm with group-wise correspondence.

Beyond respiratory motion, SCDS has the potential to be used for studying correlations among deformable objects in other medical image analysis or computer vision applications. SCDS can be used to discover how deformable objects are interacting with each other during certain motions or over time. For example, brain MR images collected over time provide evidence of the aging development of various structures in the brain (Xu et al., 2008). Whether certain diseases are caused by shape variation in a single structure or several structures, and how those structures might relate to each other require a correlation study on the longitudinal data. SCDS could provide statistical quantifications for these correlation studies and could be further applied for predictive studies.

Prediction-driven atlas formation for other applications

The prediction-driven atlas formation method (Chapter 6) combines the prediction information and the intensity information into a unified framework. The SCDS-prediction is used as a soft constraint that prevents the optimization from deviating too much from the prior knowledge. In fact, any other kind of deformation constraints could be plugged into the atlas formation framework in the same way. For example, a landmark-matching constraint defined on the image grid could be used to enforce landmark matching among the input images. In applications where landmark features are easily detectable, the landmark constraints can drive the deformations towards matching locations. The weights between the intensity forces and the landmark forces could be adjusted accordingly.

Bibliography

Balter, J. M., Brock, K. K., Litzenberg, D. W., Mcshan, D. L., Lawrence, T. S., Haken, R. T., Mcginn, C. J., Lam, K. L., & A., D. L. (2002). Daily targeting of intrahepatic tumors for radiotherapy. *Int J Radiat Oncol Biol Phys*, 53, 266–271. 16

Balter, J. M., Ten Haken, R. K., Lawrence, T. S., Lam, K. L., & Robertson, J. M. (1996). Uncertainties in CT-based radiation therapy treatment planning associated with patient breathing. *Int J Radiat Oncol Biol Phys*, 36(1), 167–74. 2

Barker, Jr, J. L., Garden, A. S., Ang, K. K., O'Daniel, J. C., Wang, H., Court, L. E., Morrison, W. H., Rosenthal, D. I., Chao, K. S. C., Tucker, S. L., Mohan, R., & Dong, L. (2004). Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated ct/linear accelerator system. *Int J Radiat Oncol Biol Phys*, 59(4), 960–70. 16

Beg, M. F., Miller, M. I., Trouvé, A., & Younes, L. (2005). Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. J. Comput. Vision*, 61(2), 139–157. 18

Bookstein, F. (1989). Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), 567 – 585. 57

Borga, M., Landelius, T., & Knutsson, H. (1997). *A Unified Approach to PCA, PLS, MLR and CCA*. Report LiTH-ISY-R-1992, ISY, Sweden. 44, 45

Cates, J., Fletcher, P., Styner, M., Shenton, M., & Whitaker, R. (2007). Shape modeling and analysis with Entropy-Based particle systems. *Information Processing in Medical Imaging*, (pp. 333–345). 33, 34

Chang, J., Sillanpaa, J., Ling, C. C., Seppi, E., Yorke, E., Mageras, G., & Amols, H. (2006). Integrating respiratory gating into a megavoltage cone-beam ct system. *Med Phys*, 33(7), 2354–61. 16

Chang, S., Frederick, B., Liu, X., Tracton, G., Lawrence, M., Lalush, D., & Pizer, S. (2009). Image-guided radiotherapy using nanotube stantionary tomosynthesis technology. *AAPM annual meeting*. 97

Chou, C.-R., Frederick, C. B., Chang, S., & Pizer, S. M. (2010). A learning-based patient repositioning method from limited-angle projections. *AAPM annual meeting*. 97

Christensen, G. E., Rabbitt, R. D., & Miller, M. I. (1996). Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10), 1435–1447. 18, 19

Davis, B. C. (2008). Medical image analysis via fréchet means of diffeomorphism. *Ph.D. Thesis, Computer Science, University of North Carolina at Chapel Hill*. 18, 38, 39

Davis, B. C., Fletcher, P. T., Bullitt, E., & Joshi, S. C. (2007). Population shape regression from random design data. *Eleventh IEEE International Conference of Computer Vision (ICCV)*, (pp. 1–7). 36, 37

Dirk, M., Kay, N., & Peter, B. (2003). Novel prospective respiratory motion correction approach for free-breathing coronary MR angiography using a patient-adapted affine motion model. *Magnetic Resonance in Medicine*, 50, 122 – 131. 3, 4, 27, 44, 54

Dupuis, P., Grenander, U., & Miller, M. (1998). *Variational problems on flows of diffeomorphisms for image matching*, volume 56, (pp. 587–600). 18, 19

Ehrhardt, J., Werner, R., Schmidt-Richberg, A., Schulz, B., & Handels, H. (2008). Generation of a mean motion model of the lung using 4D CT image data. *Eurographics Workshop on Visual Computing for Biomedicine*, (pp. 69–76). 27, 36

Endo, M., Tsunoo, T., Kandatsu, S., Tanada, S., Aradate, H., & Saito, Y. (2003). Four-dimensional computed tomography(4D CT). concept and preliminary dvelopment. *Radiat Med*, 21(1), 17–22. 12

Fillard, P., Pennec, X., Thompson, P., & Ayache, N. (2007). Evaluating brain anatomical correlations via canonical correlation analysis of sulcal lines. In *MICCAI 2007 workshop: statistical registration*: HAL - CCSD. 45

Foskey, M., Davis, B., Goyal, L., Chang, S., Chaney, E., Strehl, N., Tomei, S., Rosenman, J., & Joshi, S. (2005). Large deformation three-dimensional image registration in image-guided radiation therapy. *Phys Med Biol*, 50(24), 5869–5892. 18

Gao, G., McClelland, J., Tarte, S., Blackall, J., & Hawkes, D. (2008). Modelling the respiratory motion of the internal organs by using Canonical Correlation Analysis and dynamic MRI. *The First International Workshop on Pulmonary Image Analysis, MICCAI 2008, London, UK*, (pp. 145–155). 28, 45, 54, 89, 96

Grenander, U. (1996). *Elements of Pattern Theory*. The Johns Hopkins University Press. 18

Grenander, U. & Miller, M. I. (1998). Computational anatomy: An emerging discipline. *Quarterly of Applied Mathematics,*, 56(4), 617–694. 18

Guckenberger, M., Sweeney, R. A., Wilbert, J., Krieger, T., Richter, A., Baier, K., Mueller, G., Sauer, O., & Flentje, M. (2008). Image-guided radiotherapy for liver cancer using respiratory-correlated computed tomography and cone-beam computed tomography. *Int J Radiat Oncol Biol Phys*, 71(1), 297–304. 97

Guckenberger, M., Wilbert, J., Meyer, J., Baier, K., Richter, A., & Flentje, M. (2007). Is a single respiratory correlated 4D-CT study sufficient for evaluation of breathing motion? *Int J Radiat Oncol Biol Phys*, 67(5), 1352–9. 15

Hammoud, R., Yin, F., Yoo, S., Guan, H., Pradhan, D., Ajlouni, M., & Movsas, B. (2005). Clinical procedures for daily patient setup verification using on-board imager (obi). *Medical Physics*, 32(6), 1938–1938. 16

Hart, G., Zach, C., & Niethammer, M. (2009). An optimal control approach for deformable registration. *Computer Vision and Pattern Recognition Workshop, 2009s*, (pp. 9–16). 18

Heuberger, J., Geissbuhler, A., & Muller, H. (2005). Lung CT segmentation for image retrieval using the insight tookit(ITK). *Medical Imaging and Telemedicine*. 30

Holden, M. (2008). A review of geometric transformations for nonrigid body registration. *IEEE Trans Med Imaging*, 27(1), 111–128. 36, 91

Ibanez, L., Schroeder, W., Ng, L., & Cates, J. (2005). *ITK Software Guide for ITK 2.4*. http://www.itk.org, second edition. 60

Jaffray, D., Siewerdsen, J., Wong, J., & Martinez, A. (2002). Flat-panel cone-beam computed tomography for image-guided radiation therapy. *Int J Radiat Oncol Biol Phys*, 53, 1337– 1349. 6, 16, 96

Keall, P. (2004). 4-dimensional computed tomography imaging and treatment planning. *Semin Radiat Oncol*, 14(1), 81–90. 13

Keall, P. J., Mageras, G. S., Balter, J. M., Emery, R. S., Forster, K. M., Jiang, S. B., Kapatoes, J. M., Low, D. A., Murphy, M. J., & et al. (2006). The management of respiratory motion in radiation oncology report of aapm task group 76. *Med Phys*, 33(12), 3874–3900. 2

Köhler, T., Proksa, R., Bontus, C., Grass, M., & Timmer, J. (2002). Artifact analysis of approximate helical cone-beam ct reconstruction algorithms. *Med Phys*, 29(1), 51–64. 12

Kriminski, S. A., Lovelock, D. M., Seshan, V. E., Ali, I., Munro, P., Amols, H. I., Fuks, Z., Bilsky, M., & Yamada, Y. (2008). Comparison of kilovoltage cone-beam computed tomography with megavoltage projection pairs for paraspinal radiosurgery patient alignment and position verification. *Int J Radiat Oncol Biol Phys*, 71(5), 1572–80. 26, 64

Lee, M. H. (2007). Continuum direction vectors in high dimensional low sample size data. *Ph.D. Thesis,Unveristy of North Carolina at Chapel hill*. 40

Li, G., Citrin, D., Camphausen, K., Mueller, B., Burman, C., Mychalczak, B., Miller, R. W., & Song, Y. (2008). Advances in 4D medical imaging and 4D radiation therapy. *Technology in Cancer Research & Treatment*, 7(1), 67–81. 15

Li, T., Koong, A., & Xing, L. (2007). Enhanced 4D cone-beam CT with inter-phase motion model. *Med Phys*, 34(9), 3688–95. 17

Li, T. & Xing, L. (2007). Optimizing4D cone-beam CT acquisition protocol for external beam radiotherapy. *Int J Radiat Oncol Biol Phys*, 67(4), 1211–1219. 96

Liu, T., Shen, D., & Davatzikos, C. (2004). Predictive modeling of anatomic structures using canonical correlation analysis. *IEEE International Symposium on Biomedical Imaging: Nano to Macro*, 2, 1279–1282. 45

Low, D., Parikh, P., Lu, W., Dempsey, J., Wahab, S., Hubenschmidt, H., Nystrom, M., Handoko, M., & Bradley, J. (2005). Novel breathing motion model for radiotherapy. *Int J Radiat Oncol Biol Phys*, 63, 921–929. 3, 4, 27

Lu, W., Chen, M., Olivera, G. H., Ruchala, K. J., & Mackie, T. R. (2004). Fast free-form deformable registration via calculus of variations. *Physics in Medicine and Biology*, 49(14), 3067–87. 18

Lu, W., Parikh, P. J., Naqa, I. M. E., Nystrom, M. M., Hubenschmidt, J. P., Wahab, S. H., Mutic, S., Singh, A. K., Christensen, G. E., Bradley, J. D., & Low, D. A. (2005). Quantitation of the reconstruction quality of a four-dimensional computed tomography process for lung cancer patients. *Medical Physics*, 32, 890–901. 15

Lujan, A. E., Balter, J. M., & Ten Haken, R. K. (2003). A method for incorporating organ motion due to breathing into 3D dose calculations in the liver: sensitivity to variations in motion. *Med Phys*, 30(10), 2643–9. 14

Mageras, G. S., Pevsner, A., Yorke, E. D., Rosenzweig, K. E., Ford, E. C., Hertanto, A., Larson, S. M., Lovelock, D. M., Erdi, Y. E., Nehmeh, S. A., Humm, J. L., & Ling, C. C. (2004). Measurement of lung tumor motion using respiration-correlated ct. *Int J Radiat Oncol Biol Phys*, 60(3), 933–941. 15

Mageras, G. S., Yorke, E., Rosenzweig, K., Braban, L., Keatley, E., Ford, E., Leibel, S. A., & Ling, C. C. (2001). Fluoroscopic evaluation of diaphragmatic motion reduction with a respiratory gated radiotherapy system. *J Appl Clin Med Phys*, 2(4), 191–200. 17

Maltz, J., Fuerst, J., Paidi, A., Fadler, F., A.R., B.-H., & Sprenger, F. (2009). Fixed gantry tomosynthesis system for radiation therapy image guidance based on a multiple source x-ray tube with carbon nanotube cathodes. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009. ISBI '09.*, (pp. 1239–1242). 7

Miller, M. I., Trouvé, A., & Younes, L. (2006). Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, 24(2), 209–228. 18

Miller, M. I. & Younes, L. (2001). Group actions, homeomorphisms, and matching: A general framework. *Int. J. Comput. Vision*, 41(1/2), 61–84. 18

Minohara, S., Kanai, T., Endo, M., Noda, K., & Kanazawa, M. (2000). Respiratory gated irradiation system for heavy-ion radiotherapy. *Int J Radiat Oncol Biol Phys*, 47(4), 1097–103. 2

Muyan-Ozcelik, P., Owens, J., Xia, J., & Samant, S. (2008). Fast deformable registration on the gpu: A cuda implementation of demons. *International Conference on Computational Sciences and Its Applications, 2008. ICCSA '08.*, (pp. 223–233). 91

Naqa, I. E., Low, D., Deasy, J., Amini, A., Parikh, P., & Nystrom, M. (2003). Automated breathing motion tracking for 4D computed tomography. *IEEE Nuclear Science Symposium Conference*, 5, 3219–3222. 18

Nehmeh, S. A., Erdi, Y. E., Pan, T., Yorke, E. D., Mageras, G., Rosenzweig, K., Schoder, H., Mostafavi, H., Squire, O., Pevsner, A., Larson, S., & Humm, J. (2004). Quantitation of respiratory motion during 4D-PET/CT acquisition. *Medical Physics*, 31, 1333. 3

Neicu, T., Shirato, H., Seppenwoolde, Y., & Jiang, S. B. (2003). Synchronized moving aperture radiation therapy (smart): average tumour trajectory for lung patients. *Phys Med Biol*, 48(5), 587–98. 3

Oguz, I. (2009). Groupwise shape correspondence with local features. *Ph.D. Thesis, Computer Science, University of North Carolina at Chapel Hill*. 35, 56, 57

Oguz, I., Cates, J., Fletcher, T., Whitaker, R., Cool, D., Aylward, S., & Styner, M. (2008). Cortical correspondence using entropy-based particle systems and local features. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, (pp. 1637–1640). 33, 89

Oldham, M., Letourneau, D., Watt, L., Hugo, G., Yan, D., Lockman, D., Kim, L., Chen, P., Martinez, A., & Wong, J. (2005). Cone-beam-ct guided radiation therapy: A model for on-line application. *Radiotherapy and Oncology*, 75(3), 271–278. 16

Pizer, S. M., Chaney, E. L., & Broadhurst, R. E. (2006). Segmentation of kidneys and pelvic organs from ct by posterior optimization of M-reps. *Technical report, Computer Science Dept, Unversity of North Carolina at Chapel Hill*. 32

Pizer, S. M., Fletcher, P. T., & Joshi, S. C. e. a. (2003). Deformable m-reps for 3d medical image segmentation. *Int. J. Comput. Vision*, 55(2-3), 85–106. 32

Reinhardt, J., Christensen, G., Hoffman, E., Ding, K., & Cao, K. (2007). Registration-Derived estimates of local lung expansion as surrogates for regional ventilation. *Information Processing in Medical Imaging*, (pp. 763–774). 27

Rit, S., Wolthaus, J., van Herk, M., & Sonke, J. (2008). On-the-fly motion-compensated cone-beam CT using an a priori motion model. *1th international conference on Medical Image Computing and Computer-Assisted Intervention*, (pp. 729–736). 17

Rohlfing, T., Maurer, Jr, C. R., O'Dell, W. G., & Zhong, J. (2004). Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated mr images. *Med Phys*, 31(3), 427–32. 97

Rueckert, D., Aljabar, P., Heckemann, R. A., Hajnal, J. V., & Hammers, A. (2006). Diffeomorphic registration using B-splines. *Med Image Comput Comput Assist Interv*, 9, 702–9. 18

Sanghani, M. & Mignano, J. (2006). Intensity modulated radiation therapy: a review of current practice and future directions. *Technol Cancer Res Treat*, 5(5), 447–50. 13

Segars, W., Lalush, D., & Tsui, B. (2001). Modeling respiratory mechanics in the MCAT and spline-based MCAT phantoms. *IEEE Transactions on Nuclear Science*, 48(1), 89–97. 63

Segars, W. P. (2001). *Development and application of the new dynamic development and application of the new dynamic nurbs-based cardiac-torso(NCAT) Phantom*. PhD thesis, Unveristy of North Carolina at Chapel hill. 13

Shams, R., Sadeghi, P., Kennedy, R., & Hartley, R. (2010). A survey of medical image registration on multicore and the gpu. *IEEE Signal Processing Magazine*, 27, 50–60. 91

Siddiqi, K. & Pizer, S. M. (2008). *Medial Representations*. Springer. 32

Siebenthal, M. V. (2008). *Analysis and Modelling of Respiratory Liver Motion using 4DMRI*. PhD thesis, Electrical Engineering and ETH Zurich. 13, 14

Sonke, J., Zijp, L., Remeijer, P., & Herk, M. (2005). Respiratory correlated cone beam CT. *Medical Physics*, 32(4), 1176–1186. 7

Styner, M. & Gerig, G. (2001). Three-dimensional medial shape representation incorporating object variability. *Computer Vision and Pattern Recognition*, (pp. 651–656). 32

Szeliski, R. (1990). Bayesian modeling of uncertainty in low-level vision. *Int. J. Comput. Vision*, 5(3), 271–301. 59

Takuya, U., Kensaku, M., Manabu, M., Ken, M., & Ito, H. (2006). Trends in oncological ct imaging: clinical application of multidetector-row ct and 3D-CT imaging. *International Journal of Clinical Oncology*, 11(4), 268–277. 12

Uematsu, M., Shioda, A., Suda, A., Tahara, K., Kojima, T., Hama, Y., Kono, M., Wong, J. R., Fukui, T., & Kusano, S. (2000). Intrafractional tumor position stability during computed tomography (CT)-guided frameless stereotactic radiation therapy for lung or liver cancers with a fusion of ct and linear accelerator (focal) unit. *Int J Radiat Oncol Biol Phys*, 48(2), 443–8. 16

Vedam, S. S., Keall, P. J., Kini, V. R., Mostafavi, H., Shukla, H. P., & Mohan, R. (2003). Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Phys Med Biol*, 48(1), 45–62. 3

Vercauteren, T., Pennec, X., Perchant, A., & Ayache, N. (2007). Non-parametric diffeomorphic image registration with the demons algorithm. *Med Image Comput Comput Assist Interv*, 10(Pt 2), 319–26. 18

WHO (2006). *World Health Organization fact sheet 297*. Technical report. 1

Wold, H. (1976). Soft modelling by latent variables: The non-linear iterative partial least squares (nipals) approach. *Perspectives in Probability and Statistics*, (pp. 177–144). 45

Xing, L., Thorndyke, B., Schreibmann, E., Yang, Y., Li, T.-F., Kim, G.-Y., Luxton, G., & Koong, A. (2006). Overview of image-guided radiation therapy. *Med Dosim*, 31(2), 91–112. 13

Xu, S., Styner, M., Gilmore, J., & Gerig, G. (2008). Multivariate nonlinear mixed model to analyze longitudinal image data: MRI study of early brain developmentri study of early brain development. *IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*. 97

Yang, D., Lu, W., Low, D. A., Deasy, J. O., Hope, A. J., & El Naqa, I. (2008). 4D-CT motion estimation using deformable image registration and 5D respiratory motion modeling. *Med Phys*, 35(10), 4577–90. 3

Zeng, R., Fessler, J. A., & Balter, J. M. (2007). Estimating 3D respiratory motion from orbiting views by tomographic image registration. *IEEE Trans Med Imaging*, 26, 153–163. 3

Zerda, A., Armbruster, B., & Lei, X. (2007). Formulating adaptive radiation therapy (ART) treatment planning into a closed-loop control framework. *Phys Med Biol*, 52(14), 4137–53. 13

Zhang, Q., Hu, Y. C., Liu, F., Goodman, K., Rosenzweig, K. E., & Mageras, G. S. (2010). Correction of motion artifacts in cone-beam CT using a patient-specific respiratory motion model. *Medical physics*, 37(6), 2901–2910. 27, 36, 54

Zhang, Q., Pevsner, A., Hertanto, A., Hu, Y., Rosenzweig, K., Ling, C., & Mageras, G. (2007). A patient-specific respiratory model of anatomical motion for radiation treatment planning. *Medical Physics*, 34, 4772. 3, 4, 27, 44, 51, 88