2015

Spatio-Temporal Modeling Of Anatomic Motion For Radiation Therapy

Elizabeth Zachariah
zachariahes@mymail.vcu.edu

Spatio-Temporal Modeling Of Anatomic Motion For Radiation Therapy

A dissertation submitted in partial fulfillment of the requirements for the degree Doctor of Philosophy at Virginia Commonwealth University.

by

Elizabeth Shobha Zachariah

Master of Science
Virginia Commonwealth University
December 2010

Bachelor of Science
Virginia Commonwealth University
May 2008

Director: Alen Docef, Ph.D.
Associate Professor
Department of Electrical and Computer Engineering

Virginia Commonwealth University
Richmond, VA
August 2015

ACKNOWLEDGMENT

First, I would like to thank my dissertation advisor, Dr. Alen Docef. During my years in graduate school, he has provided me with endless support and guidance. I am very grateful for the knowledge that he has bestowed on me. I am also thankful for the patience, kind words, and advice that he has given me. Thank you Dr. Docef for being my PhD advisor.

I would also like to thank all of my committee members for the guidance you have given me throughout my academic career at VCU. Thank you to Dr. Ding-Yu Fei, Dr. A. Vennie Fillipas, Dr. Ou Bai, and Dr. Rosalyn Hobson Hargraves. I have known Dr. Fei the longest, and when I first met him during Freshman Orientation, I knew that I wanted to do what he did. He was my inspiration to begin learning about bioinstrumentation and medical imaging. Next I would like to thank Dr. Filippas for her guidance. I consider her my mentor. She has done so much for me in the past nine years I have known her. I know that without her guidance I would not be where I am today. She is also one of the best human beings I know, and I want her to know that I truly appreciate her. I want to thank Dr. Bai for his guidance and especially his enthusiasm towards teaching me. Thank you to Dr. Hobson Hargraves for being a great role model. She has always given me good advice and support when I needed it.

I would also like to thank all of the staff at the Office of Student Services and the Office of Career Services at VCU Engineering. It has been a pleasure working with you. Thank you to Ms. Rivers, Ms. Anita, Laura, Sarah, Scott, Micki, and Rebecca. You have given me lots of moral support, and I am always grateful. Next, I want to thank Ms. Leena Joseph. She has been a guiding light and a great support.

Finally and most importantly, I want to thank my family. My parents and my sister have been so supportive during this process. They have gone through the highs and lows with me, and they are most deserving of my thanks. I am very proud and thankful for my family, and I love you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CT	Computed Tomography
IGRT	Image Guided Radiation Therapy
4D CT	Four Dimensional Computed Tomography
PCA	Principal Component Analysis
DVF	Deformation Vector Field
rMSE	Root Mean Square Error
RMS Error	Root Mean Square Error
3D CRT	Three dimensional conformal radiation therapy
IMRT	Intensity modulated radiation therapy
MRI	Magnetic Resonance Imaging
ABC	Active Breath Control
NN	Neural Network
FF Network	Feed-Forward Network

Abstract

Spatio-Temporal Modeling Of Anatomic Motion For Radiation Therapy

By Elizabeth Zachariah, M.S.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2015.

Major Director: Alen Docef, Ph.D., Associate Professor, Department of Electrical and Computer Engineering

In radiation therapy, it is imperative to deliver high doses of radiation to the tumor while reducing radiation to the healthy tissue. Respiratory motion is the most significant source of errors during treatment. Therefore, it is essential to accurately model respiratory motion for precise and effective radiation delivery. Many approaches exist to account for respiratory motion, such as controlled breath hold and respiratory gating, and they have been relatively successful. They still present many drawbacks. Thus, research has been expanded to tumor tracking.

The overall goal of 4D-CT is to predict tumor motion in real time, and this work attempts to move in that direction. The following work addresses both the temporal and the spatial aspects of four-dimensional CT reconstruction. The aims of the paper are to (1) estimate the temporal parameters of 4D models for anatomy deformation using a novel neural network approach and (2) to use intelligently

chosen non-uniform, non-separable splines to improve the spatial resolution of the deformation models in image registration.

CHAPTER 1

INTRODUCTION

Radiation therapy is the primary method used for eradicating tumors in the thorax and abdomen. A major problem in radiation therapy is that normal healthy tissue in the path of radiation is affected and incorrectly irradiated. Another problem is identifying the tumor or tissues to be irradiated and delivering the dose to the appropriate tissues once they have been identified. For this reason, there is a need to minimize the amount of radiation received by the healthy cells adjacent to the tumor [1].

To understand why this collateral damage might occur even with a priori knowledge of the position and size of the tumor, one must consider that the tumor can move in unpredictable ways, either because of patient motion or, as in the case of a lung tumor, because of a patient's unique breathing patterns. Thus, it is important to track the location and changing geometry of the tumor during therapy [2]; in fact, because of the latency of the available medical devices, one must predict the location and geometry of the tumor through the study of its behavior in previous breathing cycles. This technique, known as image guided radiation therapy (IGRT), is a prominent method used to increase accuracy in dose targeting. IGRT uses numerous and repetitive imaging to improve the localization of the target tissue by tracking the changes of the target during the radiation delivery [3,4].

A critical tool in the acquisition of data for IGRT is time-dependent computed tomography (ordinarily denoted as 4DCT). 4DCT is able to image the movement of the anatomy [5]. It is principally used to account for respiratory motion. As such, 4DCT requires a large amount of data to reconstruct a time sequence of CTs from scratch [6]. Numerous methods have been implemented to construct a 4DCT model of the moving anatomy. Some methods include modeling respiratory motion with cosine models, or using multiple breathing periods and averaging them to create one model period. In actuality, respiratory motion is more complex, and studies show breathing changes during treatment [6].

While the major goal of 4D-CT is to predict tumor motion in real time, or tumor tracking, this research seeks to take a step in that direction by modeling breathing induced motion and tumor deformation, the change in its location and geometry. This research models the respiratory motion using 4D-CT model that uses a real breathing trace, a static high-resolution CT, and CT projections.

This paper presents a method to create a more realistic and versatile spatio-temporal model of the moving anatomy for use in 4D cone beam CT reconstruction. The aims of this paper are two-fold: to (1) estimate the temporal parameters of 4D models for anatomy deformation and (2) to use non-uniform, non-separable splines to improve the spatial resolution of the deformation models. For aim 1, it is hypothesized that temporal parameter estimation using neural networks will result in clinically accurate DVF estimation. For aim 2, it is hypothesized that using non-uniform, non-separable splines in deformable image registration will

result in greater registration accuracy than the conventional method of using uniform splines.

CONTEXT OF THE PROBLEM

In this section we briefly clarify the role that the techniques developed in our research have in the image-guided radiation therapy cycle shown in Figure 1. Radiation therapy begins with a treatment planning CT scan, where the cancerous tumor is delineated and the treatment parameters of the radiation field are determined for optimal coverage of a static tumor. The planning scan is a one-time process, and is not acquired during radiation treatment; therefore, patient comfort is less of a concern. Thus the scan can be as long as needed and can employ controlled breathing techniques in order to obtain a high quality image of the tumor.

Radiation is then delivered over several fractions (days) as determined by the particular type, size, and location of the tumor. The patient breathes freely during treatment. Each treatment fraction may start with a calibration or training phase where additional imaging is performed but no radiation is delivered.

Tumor displacement occurs both between planning and treatment and within each treatment fraction. An ideal image guided system would allow perfect tracking of the tumor by adjusting the radiation parameters in real time based on both a priori information (the planning scan, images acquired during training) and real time information (such as a breathing trace and x-ray images acquired in real time). This is, as of now, an unsolved problem and research carried out in 4D-CT,

including the research presented in this dissertation, aims to take steps towards the goal of perfect tumor tracking.

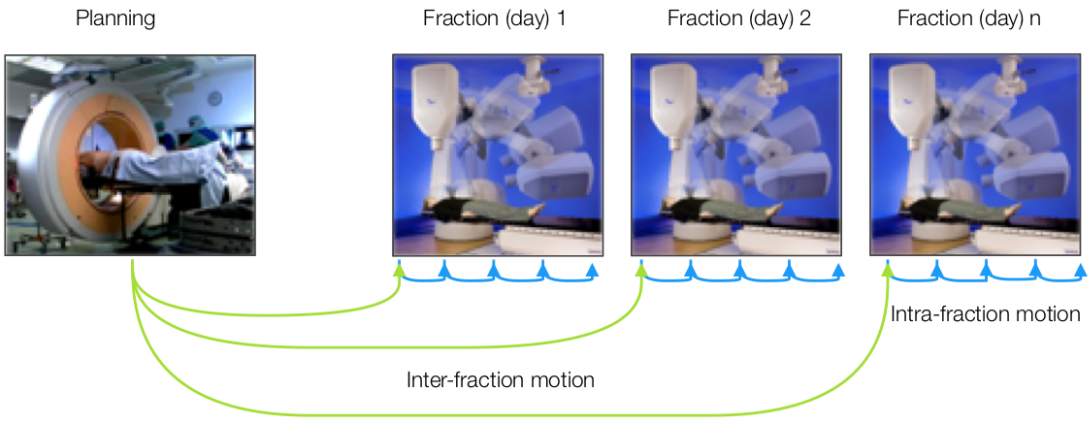


Figure 1: The radiation therapy cycle

While the ideal tracking system needs to predict the tumor location in real time, most of the research so far has focused on modeling the breathing-induced motion and tumor deformation. The intention has been to understand this motion to eventually be able to predict it. In conventional binning-based 4D-CT, motion is reconstructed at 8-10 time instants within a breathing period, as depicted in Figure 3. Model-based 4D-CTs, such as the one employed in this research, attempt instead to produce a much higher temporal resolution by reconstructing the motion at an arbitrary time *t*. The scope of this research fits into this context by:

Aim 1 (Chapter 3): A 4D-CT technique based on a PCA representation of anatomical motion was introduced in Staub et al. [9]. I proposed, in a previous work, an efficient method for estimating the basis functions that does not require a full PCA analysis [10]. The method requires prior knowledge of a few temporal parameters. This thesis developed a method for estimating these parameters using neural networks.

Aim 2 (Chapter 4): Image registration is a critical component of image-guided therapy. For inter-fraction motion, it is used to map the planning CT to the daily anatomical geometry. For intra-fraction motion it is used to generate a small number of DVFs, which in turn can be used to obtain the PCA basis functions. This thesis outlines an innovative method for image registration which can model discontinuous anatomical motion, such as sliding motion.

IMPACT OF THE RESEARCH

The significance of this research has both short-term and long-term implications. In the short term, the results presented in this work, in Chapters 3 and 4, are of interest mostly to other researchers working to improve image-guided radiation therapy for cancers in the chest and abdomen. The temporal estimation method discussed in Chapter 3 can be used with the 4D-CT model to accurately

model realistic breathing motion. The image registration method is useful in estimating a deformation vector field modeling the motion between consecutive bins in a conventional 4D-CT. The techniques developed here are additional tools, which will prove to be useful in optimizing tumor tracking. While these techniques are not yet ready for integration into clinical systems, they are at the stage where they can be disseminated to therapists and equipment manufacturers in order to ensure clinical relevance, to identify potential applications in other clinical procedures, and to determine synergies with existing equipment functionalities. In the long term, these techniques can potentially be useful in creating a 4D-CT model that can predict tumor motion and deformation in real-time. Finally, although these techniques are specific to radiation therapy in the thorax, these techniques, especially image registration using splines, are applicable for and have potential value for any medical imaging modality.

NOVELTY AND CONTRIBUTIONS

The focus of this dissertation is two-fold; (1) it is to estimate the temporal parameters of the 4D-CT model using statistical features of CT projections, and (2) to a rough DVF of the deformation, from splines, between consecutive bins during a conventional 4D-CT. The most common method for model-based 4D-CT is using an a priori deformation model with a corresponding respiratory trace taken from the patient [11,12,13,14,15,2,3,16]. The model in this research is based on the one introduced by Docef and Murphy [15]. This 4D-CT DVF model that uses basis

vectors to model the spatial variation and a breathing (respiratory) trace to model temporal variations. In this method the breathing trace and CT projections are collected and a 4D-CT image is reconstructed using filtered back-projection. Staub et al. have elaborated on this model and have used it with Principal Component Analysis, where the PCA method is used to determine the basis vectors [9]. I have, in a previous work, developed a way to estimate the basis vectors that does not require a full PCA analysis [10]. However, this method requires previous knowledge of temporal parameters of the model. Figure 2 shows the contributions this work makes to the overall modeling of the 4D-CT.

In the 4D-CT process, initially a patient scan is conducted, and from the scan projections, a breathing trace, and a conventional filtered back projection 4D CT is constructed. An estimated rough DVF of the deformation between consecutive time bins is modeled using non-uniform, non-separable spline based image registration (aim 2). A detailed schematic diagram of the image registration is shown in *Figure 19* and the methodology is further explained in Chapter 4. The DVF is used to compute the basis vectors for the 4D spatio-temporal model. In parallel, the temporal delays are estimated using neural networks (aim 1). The temporal parameter estimation method is explained in Chapter 3, and a schematic diagram of the process is shown in Figure 6. The temporal parameters are used to compute the DVF basis vectors using the methodology developed by the author's previous work. The DVF basis vectors, temporal parameters, projections and breathing trace form the patient scan are used to model a DVF using the spatio-temporal model. This results in a reconstructed 4D-CT of the patient's anatomy.

The novelties of this approach are in two main objectives, shown in the green blocks on the schematic diagram in Figure 2. One of the objectives of this research is to model these temporal parameters using a novel neural network based method. The second objective of this research is to create a new image registration technique using splines that models the deformation vector field between consecutive reconstructed CT images acquired from the filtered back-projection method. The proposed methodology, which addresses these objectives, is shown in two main components:

1. A new neural network method, described in Chapter 3, is developed to estimate the temporal parameters for the 4D-CT reconstruction model. The network is evaluated using a set of statistical features from a set of CT projections.
    a. The proposed method uses a set of statistical features instead of a full set of CT projections to determine the phase delays (temporal parameters) of the 4D-CT deformation model. This methodology reduces the amount of computation time and memory needed to estimate the temporal parameters.
    b. The methodology also estimates the delay values in the *x* and *y* directions independently on separate neural networks, thereby reducing the computation time and allowing for a smaller number of neural network neurons.

2. A novel non-uniform, non-separable spline based image registration technique, described in Chapter 4, is used to model deformation vector fields.
    a. The proposed method uses a set of intelligently chosen spline knots to accurately model a deformation vector field given a reference image, deformed image, and an original deformation vector field.
    b. The novel non-uniform, non-separable spline method is used to model a deformation vector field representing the change in motion between two images.
    c. The method reduces the modeling error significantly when compared to standard uniform spline based methods.

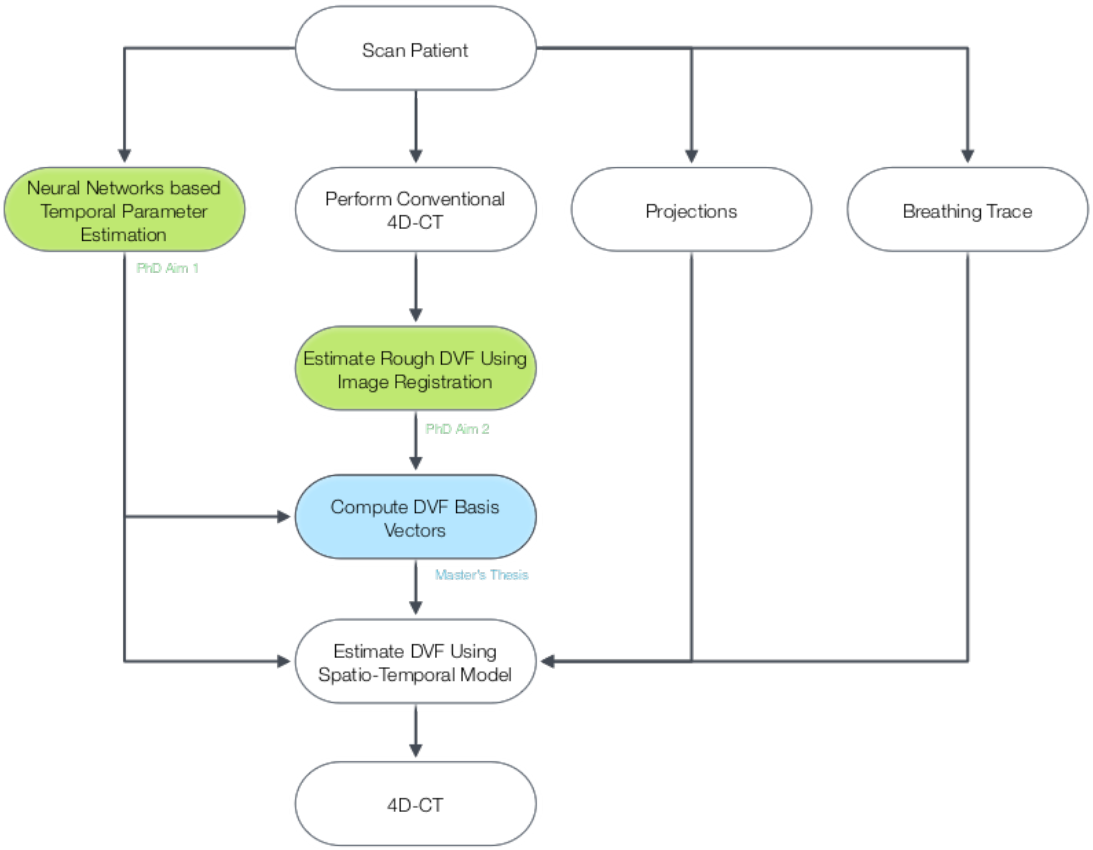


Figure 2: Schematic diagram of the overall 4D-CT process

CHAPTER 2

BACKGROUND

RADIATION THERAPY

Radiation therapy is the preferred method of eradicating tumors in the thorax and abdomen. The radiation is deposited as energy into the patient's body. When human body cells acquire too much energy at one time, they are damaged. They lose their ability to reproduce. Because a tumor is an abnormal growth of cells within the human body, it is necessary to inhibit their reproductive ability. An unfortunate side effect, however, of radiation is that normal cells in the path of the radiation are also affected. Other uncertainties in radiation therapy planning and treatment include distinguishing the tissues to be radiated and delivering the radiation dose once the tissue is identified [2,7,4]. For this reason, there is a need to locate and differentiate between normal cells and cancer cells during radiation therapy treatment [4,6,8,17,18].

Imaging is used in all stages of the radiation therapy process from patient staging to simulation, treatment planning, and radiation delivery; it is an image-guided intervention [6]. While all steps of radiation therapy are image guided, traditionally imaging technologies have only been used in identifying the location of a tumor prior to radiation treatment. With the invention of the x-ray, the scientists have been able to see two-

dimensional (2D) representations of the interior of the human body. These x-rays, at the time, provided unique images of the body structures of the human body. By using the location of the bony structures, radiologists were able to locate the position of internal organs [6]. As such, initial cancer treatments were planned by collimating rectangular fields that bounded the presumed locations of tumors based on x-ray data.

COMPUTED TOMOGRAPHY

Current practices use computed tomography (CT) for planning and treatment in radiation therapy. Computed tomography, also known as digital radiography, is an imaging technique that produces a three-dimensional representation of the anatomy through a number of radiologic images from different viewpoints [19]. CT is advantageous because it allows for the acquisition of volumetric data that outputs high quality images [20]. Benefits include the capability to process and obtain thin cross-sectional images with superior spatial resolution and spatial integrity than other imaging modalities, such as ultrasound, magnetic resonance imaging (MRI), and nuclear medicine [1,21]. Other advantages include depicting bony structure with very good accuracy, as well as providing relative electron density information. Electron density information becomes useful in calculations for dose distribution [6]. CT also displays precise information of the human anatomy by avoiding the overlay of three-dimensional information onto a single two-dimensional image, thus allowing for the use of methods such as tumor localization and treatment in radiation therapy [21].

IMAGE GUIDED RADIATION THERAPY

With the advent of computed tomography (CT), image data was used to build three-dimensional (3D) models of the patient anatomy and 3D conformal radiation therapy was developed. 3D conformal radiation therapy (3D CRT) is an image guided radiation therapy (IGRT) method where a 3DCT image of the anatomy is taken, and visualized using computer software. Mathematical algorithms are used to calculate a conformed or focused target area for radiation, corresponding to the tumor target volume. Another technique commonly used in image-guided radiation therapy is intensity modulated radiation therapy (IMRT). IMRT, like 3D CRT, uses a 3D image and delivers a high dose of radiation fit tightly around the target volume, with greater dosage to the center and diminished dosage to the outer boundaries of the tumor volume [6,4].

3D CRT and IMRT technologies, allow for the optimization of both accuracy and precision of radiation therapy planning and treatment. This is done by allowing the radiation beam to be adjusted based on the position of the tumor and the critical organs. These advances, 3D CRT and IMRT, provide an unparalleled method where high-dose radiation conforms closely around the tumor, while at the same time doses are reduced to susceptible healthy tissue [4]. This diminishes the toxic effect to the healthy tissues.

A health risk is associated with ionizing radiation. McCollough, et al. states that the radiation dose of one CT scan, 1-14mSv (milli-Sieverts), is equivalent to the annual dose of background radiation (1-10mSv), as with radon and cosmic radiation. Further statistics show that 0.4 percent of all cancers in the United States are caused from CT radiation [1]. Improvement of precision from 3D CRT and IMRT has been shown to shorten the duration

of radiation therapy, thereby reducing the number of treatment sessions for some types of cancer, which also reduces the amount of radiation a patient receives [6]. While these technologies provide a means for delivering radiation within the target volume dimensions and sparing healthy tissue, they so only when the patient is completely immobile [6].

RESPIRATORY MOTION IN RADIATION THERAPY

In practice, considerable ambiguities still exist in tumor volume delineation, patient immobilization, and the patient's breathing motion that makes it difficult to administer the high dose required to irradiate the tumor [6]. In these cases, the dose delivered to the tumor might be lower than projected, whereas the dose to the healthy tissue will be higher than what was planned [4,22]. As such, accounting for anatomical motion can help improve radiation planning, thereby improving coverage and reducing excessive radiation dose [23,24,25].

Organs in the upper abdomen can move up to 4cm during the breathing cycle. Consequently, 3D CRT and IMRT cannot be fully utilized until temporal motion is accounted for during therapy planning and radiation. Time-dependent imaging is necessary to account for the tumor motion due to breathing [5].

Respiration-induced motion is the most significant source of positional errors in the thorax and abdomen [11]. This motion affects the reproducibility of the target volume during radiation therapy by distorting the target volume. This results in the positioning errors where parts of the tumor move in and out of dosing image window. Internal motion of the body can be classified into inter-fraction and intra-fraction components. Inter-

fraction motion occurs through the change in tumor size, either growth or shrinkage, as well as the daily filling and emptying of the bladder and bowel [26]. Motion can also be the result of weight loss or gain between radiation therapy treatment sessions. Intra-fraction motion occurs through organ processes, such as motions related to respiration, the cardiac cycle, or peristalsis of the digestive system [6,26].

Intra-fraction motion, in the thorax and abdomen, is the primary cause for tumor movement. Several studies have determined that the range of motion in the superior-inferior direction of the diaphragm due to normal respiration is approximately ~0.5 - 4.0 cm [4,6,27,28]. Motion distorts the target volume to move in and out of the dose-targeting window. Subsequently, the area to dose the tumor is increased to compensate for the organ motion. As a result, large amounts of normal tissue are radiated.

FOUR DIMENSIONAL COMPUTED TOMOGRAPHY (4DCT)

The goal of four-dimensional radiation therapy is to deliver high doses of radiation to the tumor, while minimizing the dose to the surrounding healthy tissue [29]. Respiratory motion should be considered and compensated for beginning with the simulation of radiation therapy treatment planning. 4D imaging focuses on reconstructing a volume that shows the changes in tumor shape, size, and position, as well as accounts for respiratory motion with minimized motion artifacts [30,31].

Current practices of 4D CT attempt to include temporal aspects and have been reasonably successful, yet areas that need improvement exist. The standard practice requires the acquisition of large amounts of CT data to reconstruct a complete temporal

sequence. A typical three-dimensional CT data set generally contains approximately 100 axial slices, where each slice contains 512x512 pixels. With 16 bits per pixel, a single data set takes over 50 megabytes of storage [6]. Therefore, large data sets increase processing time. Also, as stated previously, larger amounts of data require a greater all around radiation increase to the patient. It is imperative that the amount of data is decreased so patient radiation dose and processing time is decreased.

4D CT is used as verification imaging; it is obtained before, during, and/or after radiation therapy treatment. With such numerous imaging, large setup errors can be detected. Imaging at the time of treatment can also help identify and increase awareness of the range of motion, size, and shape of the tumor [4].

A 4D CT is in either of two cases: prospective 4D CT or retrospective 4D CT. I prospective 4D CT is collected by collecting CT images at one static breathing phase versus collecting at all phases. Whereas, the retrospective 4D CT collects information corresponding to multiple breathing phases by recording respiratory signals, acquiring time-dependent CT data, and constructing a 4D image from the two previous sets of data [32,33,34].

The breathing signal, in conventional approaches, is assumed by the location of internal anatomy from some set of surface marks [8]. A substitute of respiration motion is measured by using a reflective maker, which is tracked by a camera, on the abdomen, or by measuring the tidal volume with a spirometer [32,14]. When using the substitute signal for respiration, the signal is recorded only when the CT is also simultaneously collecting imaging data, so that the temporal data correlates with the CT imaging data [35]. Once the breathing pattern has been established, time dependent CT projection data is acquired in

the helical or cine mode [11,14,36,33], and CT slices are constructed correlating to the full respiratory cycle [34,37].

The 4D CT patient setup is similar to that of a standard 3D CT exam. The patient is positioned as for treatment delivery with the aid of optical lasers and is immobilized [4,18]. An initial CT scan is conducted to obtain structural information on the patient's anatomy [18]. The information is contained in a three-dimensional data set. The information collected in this data set is used to guide any radiation therapy treatment as well as to position the patient during multiple treatment sessions [4]. Once the CT data set is obtained and the physical examination is conducted, the images and information are transferred to the radiation treatment-planning phase. In the radiation therapy treatment-planning phase, tumor extension and the organs at risk are determined and assessed [4]. The target volume to treat is also defined. The treatment parameters and the target volume are defined with the information taken from the previous stage. The plan for treatment is calculated with the parameters and is transferred to the next phase, the set-up verification phase.

The set-up verification phase occurs each time the patient is administered radiation. The patient is positioned in the exact manner as in the patient staging and simulation phase. Typically a reflecting block is attached just inferior to the xiphoid process on the patient's abdomen, and a camera captures and analyzes the anterioposterior motion of the moving block in real-time [6]. The breathing pattern is recorded from the rise and fall of the reflecting block and is referred to as the "respiratory trace" [6]. 4D CT is designed to create image data from the respiratory trace, captured by a reflecting block or other acquisition device, and with images obtained from the CT scan [14,35,37,34].

CONVENTIONAL 4D CT

In the conventional method, once the breathing pattern has been established, time dependent CT projection data is acquired in the helical or cine mode [4,14,33,37]. The data, projections or images from the CT scan for each anatomical section is partitioned into time bins, based on amplitude or phase values of the breathing trace, corresponding to a user specific time interval [38,39]. Generally, a time cycle is divided into 10 time bins [6]. CT slices are then generated from these bins. Amplitude-based methods take the raw image data or projections and label it based on the relative magnitude with respect to the respiratory trace [40]. The conventional method is to use phase-based binning, where corresponding CT images are assigned to the respiratory phase, calculated from the recorded breathing trace for the corresponding point in time [34,40]. The end-respiration peaks are detected by software, and a linear interpolation method is used to assign percentages, based on the end-respiration peak, to the remaining points on the trace [18]. The sorted images on the time sequence form the basis for 4D CT treatment. This is depicted in Figure 3.

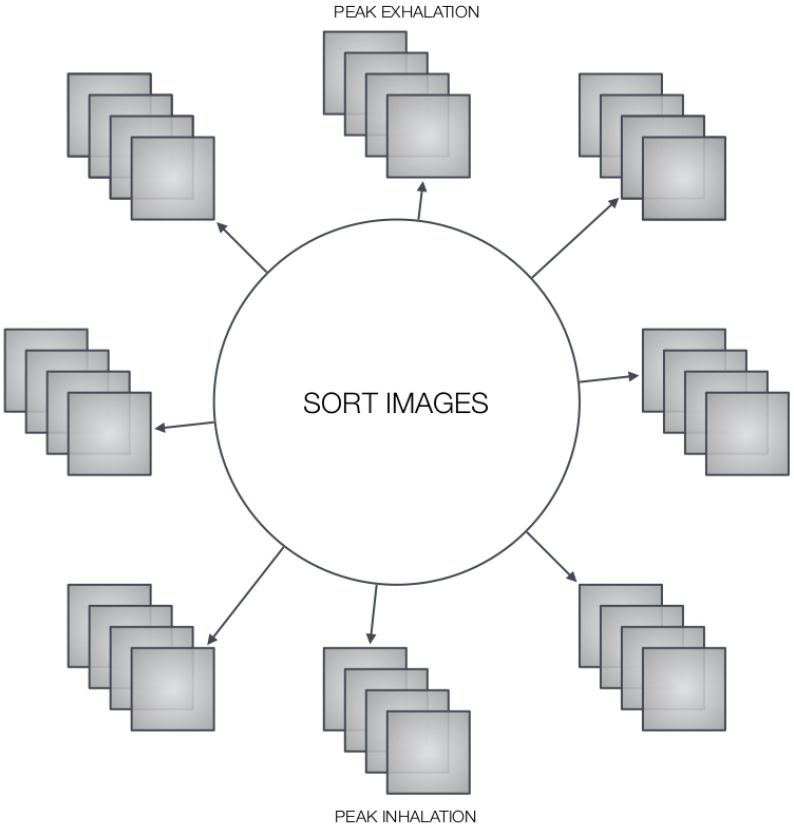


Figure 3: Conventional 4D-CT sorting method: CT images acquired and sorted into time bins by phase

Breathing must be regular when acquiring the respiratory trace; therefore the patient is encouraged to breath calmly and consistently. The patient's compliance is necessary to achieve a steady breathing trace. Variations between the breathing cycles will affect the consistency of the data set, as the 3D data sets representing each phase are collected from multiple breathing cycles [32]. Breathing training techniques are introduced in the form of audio coaching and active breath control (ABC) [23,41,42]. In the former case, audio and/or video coaching is provided by a "breathe in, breathe out" recording. Audio and video coaching has shown to help stabilize the respiratory period,

amplitude, and base line; however for patients who have compromised respiratory function, it is very difficult and nearly impossible for the patient to keep a regular rhythm [6,43,44].

For the latter case, active breath control, either actively or passively suspends the patient's breathing. The patient will either hold his or her breath, or a modified ventilator is used to control the patient's airflow [41]. While the respiration is halted, and thereby tumor motion is negligible, the treatment is administered in the interval [4,41]. This method, while effective, is at the expense of patient comfort [41]. Many patients have difficulty holding breath, especially those with lung cancer [26,13]. For this reason, there is need for technologies that allow the patient to breathe freely.

Respiratory gating is a technique that allows the patient to breathe freely. The patient's natural breathing pattern is observed, and treatment is delivered periodically when breathing reaches a particular phase [6,18,42]. Usually, the dose is delivered at end-inhale or end-exhale [6]. The patient's respiration is observed and the thoracic wall displacement generally triggers the activation of the radiation dose [45]. Like the controlled breath hold procedure, the initial CT scan taken during simulation must represent the CT sets taken during treatment. A disadvantage of respiratory gating is the active participation of the patient. The patient is required to control breathing to a steady pattern. This is not always possible as there are patients with respiratory complications who are unable to breathe steadily [13]. Another disadvantage is that treatment is longer, since radiation is delivered only over a fraction of the breathing cycle.

Tumor tracking, however, allows for the patient to breathe freely. The delivery field follows the displacement of the target volume during the entire breathing cycle. The field

follows the 4DCT set acquired during simulation. Multiple methods have been proposed to increase the accuracy of tumor tracking, but the subject still requires more research.

An accurate method of four-dimensional computed tomography reconstruction is necessary to conduct the previous treatment procedures. 4D imaging focuses on reconstructing a volume that accounts for respiratory motion with minimized motion artifacts [30]. Respiration motion must be accounted for during imaging of moving objects, else the motion artifacts will lead to errors in the detection of the shape and position of the target structures as well as any critical structures [32,5,31]. It has been observed that image quality and information from CTs have been degraded with respiratory motion, and these temporal inaccuracies have lead up to 40% delineation of target sizes and volumes [46,5]. This makes it such that the planned dose is delivered inaccurately [22].

CHAPTER 3

ESTIMATION OF TEMPORAL PARAMETERS USING NEURAL NETWORKS

ESTIMATION OF TEMPORAL PARAMETERS OF 4D MODELS FOR ANATOMY DEFORMATION

The purpose of this work is to create a network that can determine phases for a spatio-temporal model of the moving anatomy. When modeling anatomic motion, the most challenging task is to achieve a suitable representation of displacement variation in time.

A critical requirement for tumor tracking is a time-dependent CT study, ordinarily denoted as 4D CT. Numerous methods have been implemented to construct a 4D-CT model of the moving anatomy. Some methods include modeling respiratory motion with cosine models, or averaging multiple breathing periods to create one model period [12]. In actuality, respiratory motion is more complex, and breathing periods change during treatment [6]. The more common approach is to combine CT data with a respiratory trace and assume that anatomical motion occurs in synchrony with the trace [11,41,47]. The trace can be obtained by monitoring a

physical marker (such as a reflecting block attached to the patient's abdomen) or by measuring the tidal volume using a spirometer [48].

In the conventional 4D-CT method, the breathing trace is collected as CT projections are acquired [4,47,33,49]. Based on their temporal locations within a breathing cycle, the projections are partitioned into several bins [48]. Commonly, 10 time bins are used. A 3D CT image is reconstructed for each time bin using filtered back-projection, resulting in ten CTs covering a typical breathing period.

Model-based 4D-CT methods do not limit the 4D-CT to one typical period, thus addressing inter-period variations. They also provide much better temporal resolution, potentially providing images of the anatomy at arbitrary time instants. These methods use a reference 3D image $S(\boldsymbol{r})$ which is deformed by a time-dependent Displacement Vector Field (DVF) $\boldsymbol{d}(\boldsymbol{r}, t)$ such that the anatomy at any time t is

$$S'(\boldsymbol{r}, t) = S\big(\boldsymbol{r} + \boldsymbol{d}(\boldsymbol{r}, t)\big) \qquad (1)$$

Here, $\boldsymbol{r}$ is the 3D spatial variable. In our notation all bold-faced variables are vectors in space.

The spatio-temporal DVF model employed in this research is the one introduced in [15,3,9]. This model employs DVF basis vectors $\boldsymbol{u}_i(\boldsymbol{r})$ to model spatial variation in the DVF and the breathing trace $b(t)$ to model temporal variation:

$$\boldsymbol{d}(\boldsymbol{r}, t) = \sum_{\boldsymbol{i}} a_i \boldsymbol{u}_i(\boldsymbol{r}) b(t - \theta_i) \qquad (2)$$

Each spatial basis vector $\boldsymbol{u}_i(\boldsymbol{r})$ is scaled by the amplitude a_i and modulated by the term $b(t - \theta_i)$, resulting in a DVF component that varies proportionally with

breathing, with a delay θ_i. It has been shown that three basis vectors are sufficient for accurate representation of DVFs [9]. The basis vectors are obtained as follows: First, conventional 4D-CT is performed on the CT projection set. Pairs of consecutive projections are then registered, resulting in a set of 3D DVFs. Lastly, three basis vectors are computed from this set by means of Principal Component Analysis (PCA).

The model parameters a_i and θ_i could be computed using a curve-fitting approach for anatomical landmarks (either fiducials or manually identified features); however, these are not available in realistic scenarios. Instead, projection-matching methods have been proposed in [3,9]. These methods deform the reference image using a DVF model, compute simulated projections, and iteratively adjust the DVF model parameters to minimize the difference between simulated and actual projections. Each iteration involves a large number of computations, and the algorithm may converge to a local minimum.

A method is proposed for computationally efficient marker-less estimation of the temporal parameters. A neural network is trained to estimate the parameters based on simple statistical features of the CT projections. For projection-based parameter estimation, this method reduces the number of computations and the likelihood of convergence to local minima. For binning-based DVF interpolation, the method becomes a two-step non-iterative algorithm.

NEURAL NETWORKS

Neural networks belong to a field of study that models and simulates the characteristics of the human brain [50]. A neural network is a data processing unit that is modeled on the cerebral cortex of the brain; as such artificial neurons are the basic unit of the neural network [51]. The purpose of each neuron is to relay information and make connections. These neurons have multiple inputs, and the sum of these inputs is passed through a transfer function, usually a nonlinear filter. Each neural network consists of processing elements whose outputs are summed with connection weights.

The processing elements are organized such that they form a sequence of layers (Figure 4). A neural network consists of a minimum of three layers: an input layer, one or more hidden layers, and an output layer. These layers are created as multiple processing elements (neurons) joined together [51]. The network is parameterized by weights associated to the connections between neurons in the various layers.

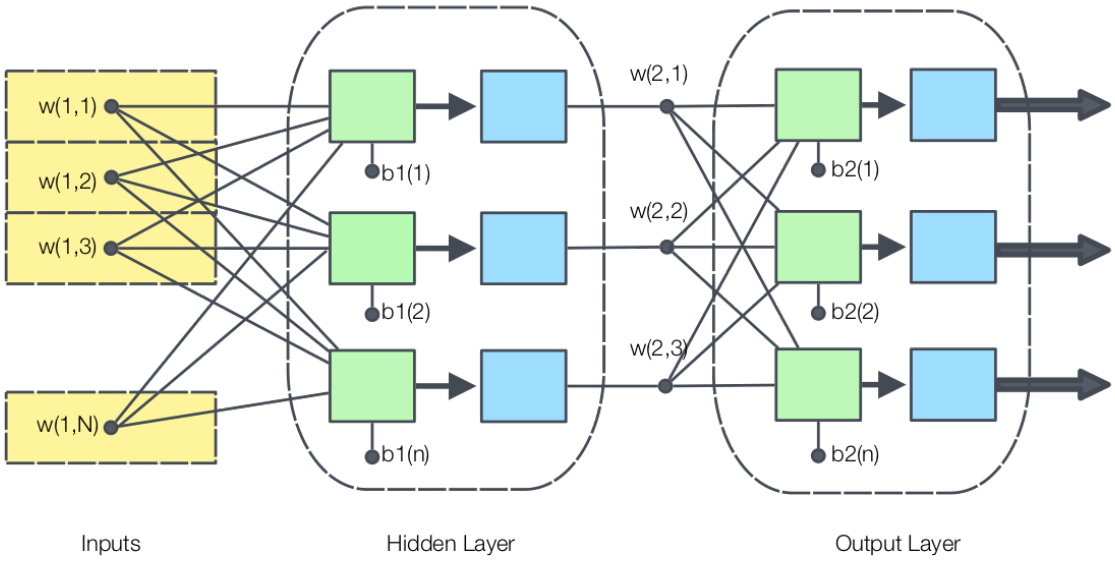


Figure 4: Neural network processing elements for the Feed-forward Neural Network

The network operates two processes, learning and recall. As with the human brain, the neurons learn and adapt weights as inputs or stimuli are presented to it. Learning is conducted through multiple learning processes; these include supervised learning, unsupervised learning, random learning, and graded learning. In all the learning algorithms, an input is presented to the network and an output is generated. The weights are then adjusted accordingly until a desired output is reached. These learning processes are governed by a learning algorithm [50].

Once the network has been trained, it recalls the information when new inputs are introduced to the network. As such, neural networks are valuable for

their ability to learn by example, their ability to recognize patterns, their fault tolerance, and for their distributed associative memory [50].

The most widely used neural network is the Feed-Forward neural network (multilayer perceptron), shown in Figure 4. It is the network used in this work. In the Feed-Forward network, as the name implies, information is fed forward through the input, hidden, and output layers without any feedback loops. The input layer consists of the inputs to the network is defined by the number of input data. The hidden layer has neurons placed in parallel, with weights and biases. The number of hidden layers and the number of neurons in each hidden layer is determined when the neural network is defined and is irrespective of the number of input and output values. The output of each neuron is defined by Eq. 3. The output layer is defined by the number of output values and is formed by a weighted summation of the outputs of the hidden layer [52].

$$y = \sigma\left(\sum_{i=1}^{n} x_i w_{ij} + b_j\right) \tag{3}$$

where y is the output of each neuron,

$\sigma()$ is the activation function,

x_i is the input to each neuron,

b_j is the bias weight to the neuron,

and w_{ij} is the weight between the each neuron and input.

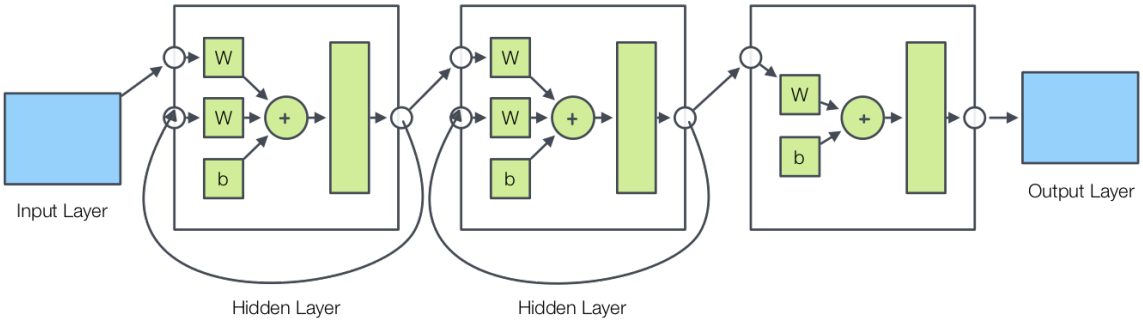


Figure 5: Schematic diagram of the Elman Recurrent Neural Network

The Elman recurrent neural network is specifically designed for the detection of time-varying patterns [53,54]. The Elman neural network consists of a two-layer back-propagation network that is connected as feedback from the outputs of the first layer to the inputs of the first layer [53,54]. This connection allows for the network to detect temporal and spatial patterns. The hidden layers all have a hyperbolic tangent sigmoid transfer function as the learning algorithm, while the output layer has a linear transfer function [54]. The Elman network is unique because every output of a hidden layer in the network is fed back to the input layer of the very same hidden layer.

ESTIMATION METHODS

In a typical clinical scenario, a 4D-CT image set must be reconstructed from the following available data: a complete set of CT projections collected as the patient

breathes freely, a digitized breathing trace, and an a priori 3D reference CT scan acquired for treatment planning. From these inputs, a complete 4D volumetric image must be computed. Usually, the reference is a fan-beam scan while the free-breathing scan is a cone-beam CT. To assess the accuracy of the reconstructed images, these must be compared to the true state of the anatomy during the scan. Since this is not available, for the purpose of algorithm development a combination of synthetic data and simulated CT scanning is used.

The temporal parameters for the model are estimated in this work using neural networks. Neural networks have been used successfully for prediction of breathing amplitudes and irregular breathing patterns [55,56]. Given their proven ability to model temporal behavior, neural networks are a good candidate for temporal parameter estimation. The use of the method in two different scenarios is described.

CASE 1: USE IN PROJECTION MATCHING 4D-CT

In this scenario, the spatial basis vectors $\boldsymbol{u}_i$ in Eq. 2 are known; a_i and θ_i are sought. A feature vector is first computed from the projection set $\{P_i, i = 1, \dots, N\}$. The features are then fed to a neural network trained to estimate the delay parameters θ_i. Once θ_i are known, the amplitudes a_i are found using the projection matching approach. An initial guess of a_i is assumed and a DVF is computed using Eq. 2. The DVF is used to deform the reference image at time instants $t_1, \dots, t_N$. Simulated CT projections $\{P_i'\}$ are computed through the moving anatomy and

compared to the actual projections $\{P_i\}$. The parameters a_i are iteratively adjusted to minimize the difference between the simulated and actual projections.

The features used to estimate θ_i are simple statistical features of the CT projections. In this paper, the mean, standard deviation, and centroid of each projection are computed and grouped in a vector of size $3N$. The neural network is trained using a large number of simulated projection sets corresponding to known θ_i values.

In contrast with existing projection matching methods [3,15], the proposed method estimates a_i and θ_i separately. Since the number of free parameters is reduced by half, projection matching is less likely to converge to a local minimum, requires fewer iterations, and fewer computations in each iteration. The single-step estimation of θ_i requires a comparatively insignificant amount of computation.

CASE 2: USE IN BASIS DVF COMPUTATION

The proposed method can be used to simplify the process by which the DVF basis functions are obtained. The PCA method of [9] assumes no knowledge of the delays θ_i. The computations can be significantly simplified if θ_i are first estimated using the method described in Case 1. We next describe the method for computing $\boldsymbol{u}_i$ and a_i, when θ_i are known.

The method assumes that N 3D DVFs are known at time instants $t_1, \ldots, t_N$. These have been computed from a conventional 4D-CT. Eq. 2 is applied N times, resulting in a set of equations that can be written in matrix form as:

$$\boldsymbol{D} = \mathbf{BAU}, \tag{4}$$

where each row of matrix □ is a DVF, each row of matrix $\boldsymbol{U}$ is a row basis vector, $\boldsymbol{A} = \mathrm{diag}(a_1, a_2, a_3)$ is a diagonal matrix of amplitude parameters, the matrix

$$\mathbf{B} = \begin{bmatrix} \mathrm{b}(\mathrm{t}_1 - \theta_1) & \mathrm{b}(\mathrm{t}_1 - \theta_2) & \mathrm{b}(\mathrm{t}_1 - \theta_3) \\ \mathrm{b}(\mathrm{t}_2 - \theta_1) & \mathrm{b}(\mathrm{t}_2 - \theta_2) & \mathrm{b}(\mathrm{t}_2 - \theta_3) \\ \vdots & \vdots & \vdots \\ \mathrm{b}(\mathrm{t}_\mathrm{N} - \theta_1) & \mathrm{b}(\mathrm{t}_\mathrm{N} - \theta_2) & \mathrm{b}(\mathrm{t}_\mathrm{N} - \theta_3) \end{bmatrix} \tag{5}$$

contains values of the breathing trace, and $\theta_1, \theta_2, \theta_3$ are the temporal parameters. In Eq. 4, the unknowns are the matrices $\boldsymbol{A}$ and $\boldsymbol{U}$. This is an over-determined system, and the least-square solution for the product $\boldsymbol{AU}$ is given by

$$\mathbf{AU} = \boldsymbol{B}^{+}\boldsymbol{D}\,, \tag{6}$$

where $\boldsymbol{B}^{+}$ is the Moore-Penrose pseudo-inverse of $\boldsymbol{B}$ [57]. Since $\boldsymbol{u}_i$ are unit vectors, they, together with a_i, can be computed by normalization of $a_i\boldsymbol{u}_i$, the rows of $\boldsymbol{AU}$:

$$\boldsymbol{u}_i = \frac{a_i\boldsymbol{u}_i}{\|a_i u_i\|}, \quad a_i = \frac{a_i\boldsymbol{u}_i}{\|u_i\|} \tag{7}$$

The method is non-iterative and involves computation of the pseudo-inverse of a small matrix (typically 10×3). By comparison, the PCA method of [9] processes vectors of size equal to the number of voxels in the image. More details about the pseudo-inverse method and its performance can be found in [10]. Note that in this context the proposed method can also be seen as a model-based DVF interpolation method.

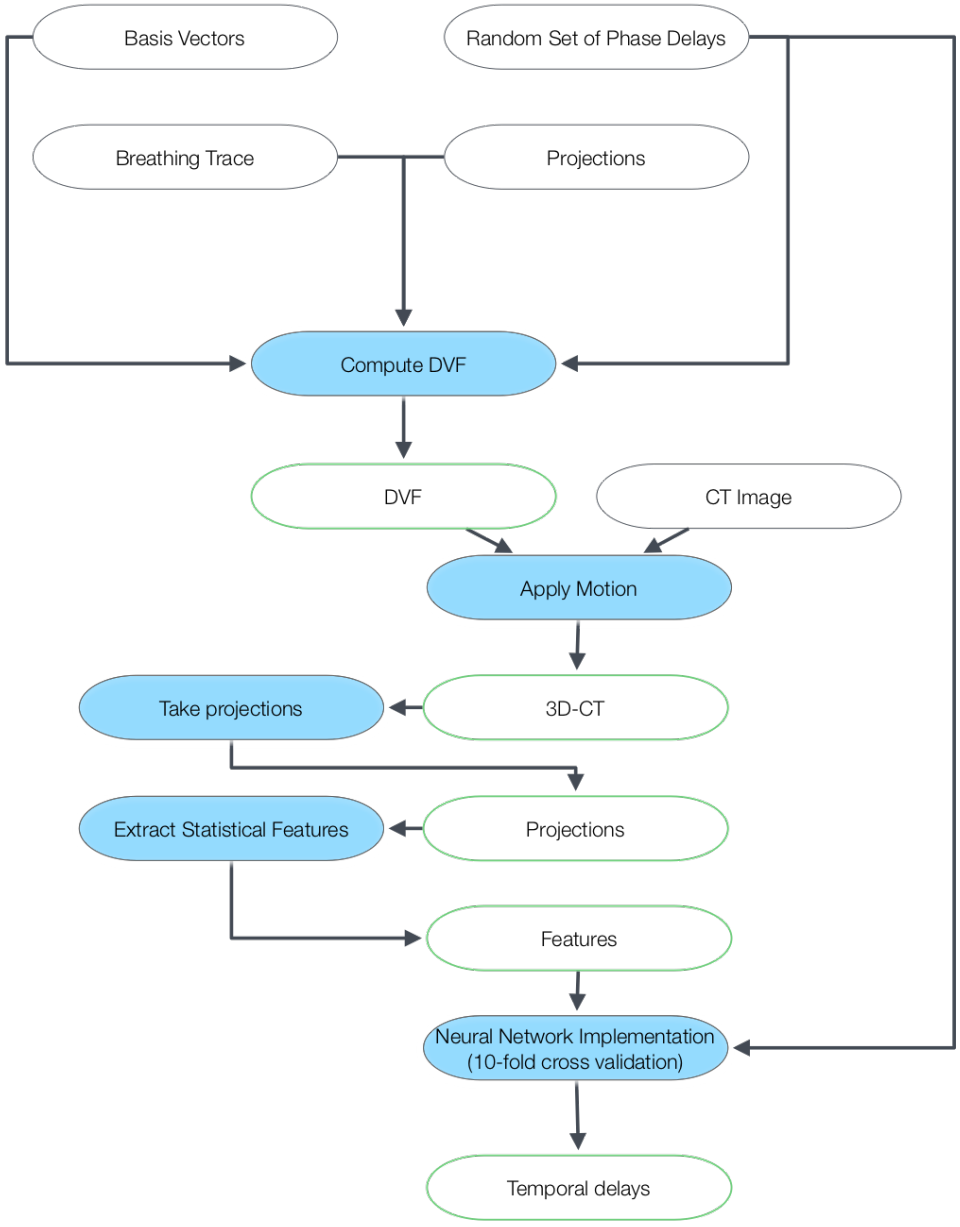


Figure 6: Schematic diagram of the Temporal Parameter Estimation Process

The proposed method is evaluated in the context of Case 1 described in the previous section. The methodology is depicted as a block diagram in Figure 6. The CT images in real clinical settings are 3D, and cone-beam CT projections are 2D. In this paper, the feasibility of the proposed method is assessed in a simplified setting using 2D CT images. The projections are therefore 1D and the DVF in Eq. 2 is 3D (2D + time).

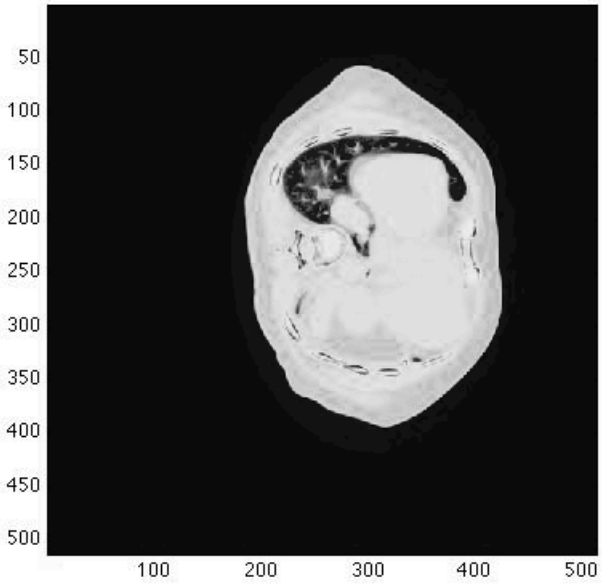
Since ground truth θ_i values are not available for real clinical data, a combination of real anatomy, real breathing, synthetic motion, and simulated projections was used, as follows:

- The reference static CT image S used in all experiments is an axial fan-beam CT slice of size 256×256. A cropped version of the image is shown in Figure 7.
- The real trace $b(t)$ previously used in [55] was chosen to model breathing and is shown in Figure 8. The trace was normalized to have zero mean and a maximum absolute value of 1.0. The scan time instants $t_1, \ldots, t_N$ were chosen to uniformly cover two full breathing periods. The period was estimated to be $T = 3.1$ s based on the average interval between peaks.
- The synthetic motion corresponds to a DVF computed using Eq. 2 with two basis vectors: $\boldsymbol{u}_1$ being an antero-posterior expansion and $\boldsymbol{u}_2$ a lateral contraction of the anatomy. In both directions the maximum displacement was chosen to be 3.5 cm. For clarity, we

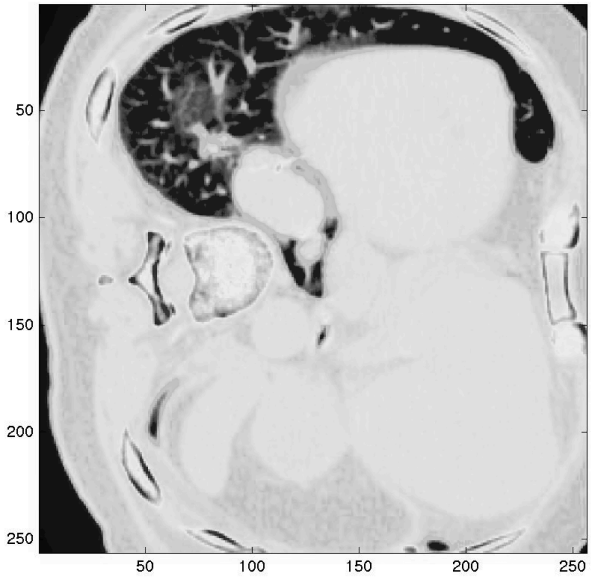
hereafter denote the two vectors by $\boldsymbol{u}_x$ and $\boldsymbol{u}_y$ and the corresponding delays by θ_x and θ_y.

- For a given (θ_x, θ_y) pair, the DVF is computed and used to animate the reference image. A set of N simulated parallel projections is then computed at time instants $t_1, \ldots, t_N$. The N projection angles uniformly cover a full gantry rotation. The algorithm is evaluated for N ranging from 8 to 64.
- For each projection, the features (mean, standard deviation, and centroid) are computed. The N triplets are merged into a single feature vector, then fed to the neural network, which estimates the two delays.

This approach made possible the comparison between the estimated and actual temporal parameters.



(a)



(b)

Figure 7: Reference high-resolution CT image (a) Full Image (b) Modified image

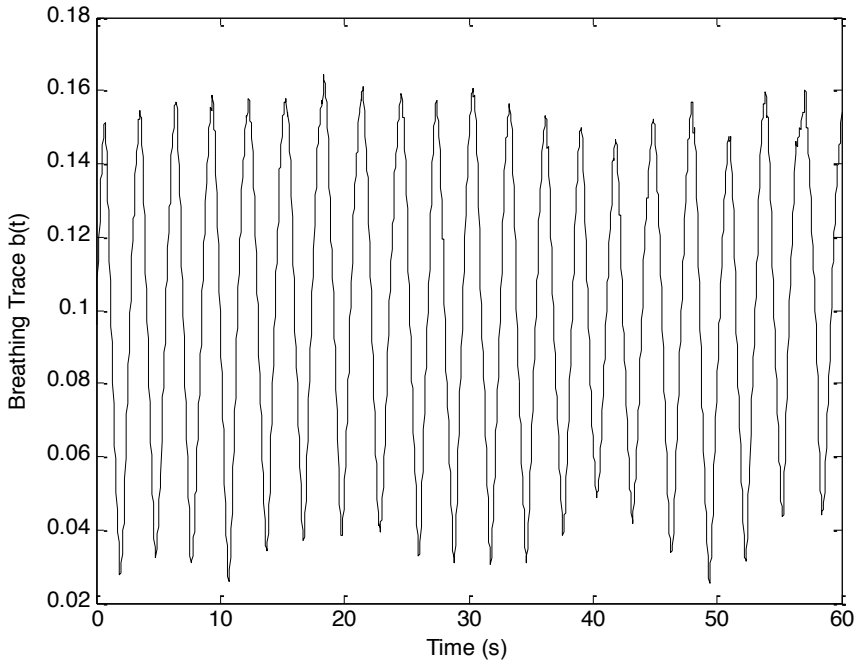


Figure 8: Breathing trace, in arbitrary dimensionless units.

A feed-forward neural network with one hidden layer of 20 neurons was used. The network takes a vector of projection features as input and outputs the estimated delay parameters. The network is trained to minimize the estimation error in the Mean Square Error (MSE) sense. The testing data sets, consisting of representative input feature vectors and target delay parameters, were created as follows: A set of random (θ_x, θ_y) delays is first generated with values ranging from $-T/2$ to $+T/2$ where T is the period of the breathing trace. For each of these, a DVF is computed and used to animate the reference image, simulated projections are computed, and a feature vector is calculated. A total of 2200 samples were used to train and test the network, with 10-fold cross-validation. In all the experiments

described below, the neural network is trained for 50 Levenberg-Marquardt iterations.

To evaluate the accuracy and robustness of the proposed method, multiple scenarios have been considered, with varying numbers of projections, projection resolutions, and levels of projection noise. The method's accuracy was quantified by means of the root MSE (rMSE) error between the actual output delays and the target delays.

RESULTS

The neural network model is evaluated by comparing the estimated temporal delays with the actual temporal delays by means of the root mean square error. Initially the model is evaluated with the feed-forward network and the Elman network. The results show the feed-forward network estimates the temporal delays more accurately than the Elman.

The neural network model is tested and trained with the Elman network and the feed-forward network. The model is evaluated under ideal conditions where there is no noise present and the projections have not been downsampled. The time delays in the *x* and *y* directions are trained together on a single neural network implementation and then trained separately on two separate neural networks. When training on a single network, the rMSE values for θ_x and θ_y are 0.2499 and 0.2186 seconds, respectively. On the other hand training in separate networks gives rMSE values for θ_x and θ_y to be 0.2136 and 0.2140 seconds, respectively. While the

error values for the separate networks are only slightly less than the single network, one advantage is the computation time is reduced when training and testing with the separate networks by half of the single network. For this reason, all experiments carried out have been conducted on separate networks.

Next the suitability of the networks is compared. Both the Elman and feed-forward networks are tested on the same data set with ideal condition where there is zero noise present and full projection resolution. When training with separate networks, the rMSE values for θ_x and θ_y to be 0.2136 and 0.2140 seconds, respectively for the Elman network. For the feed-forward network, the rMSE values are 0.0062 and 0.0159 seconds for θ_x and θ_y, respectively. The feed-forward network implementation gives an average decrease in error by 95% compared to the Elman network. This difference in this error is seen clearly in Figure 9, which shows the regression analysis for the temporal delays for the Elman and feed-forward networks. Accordingly, all experiments shown are conducted on the feed-forward network.

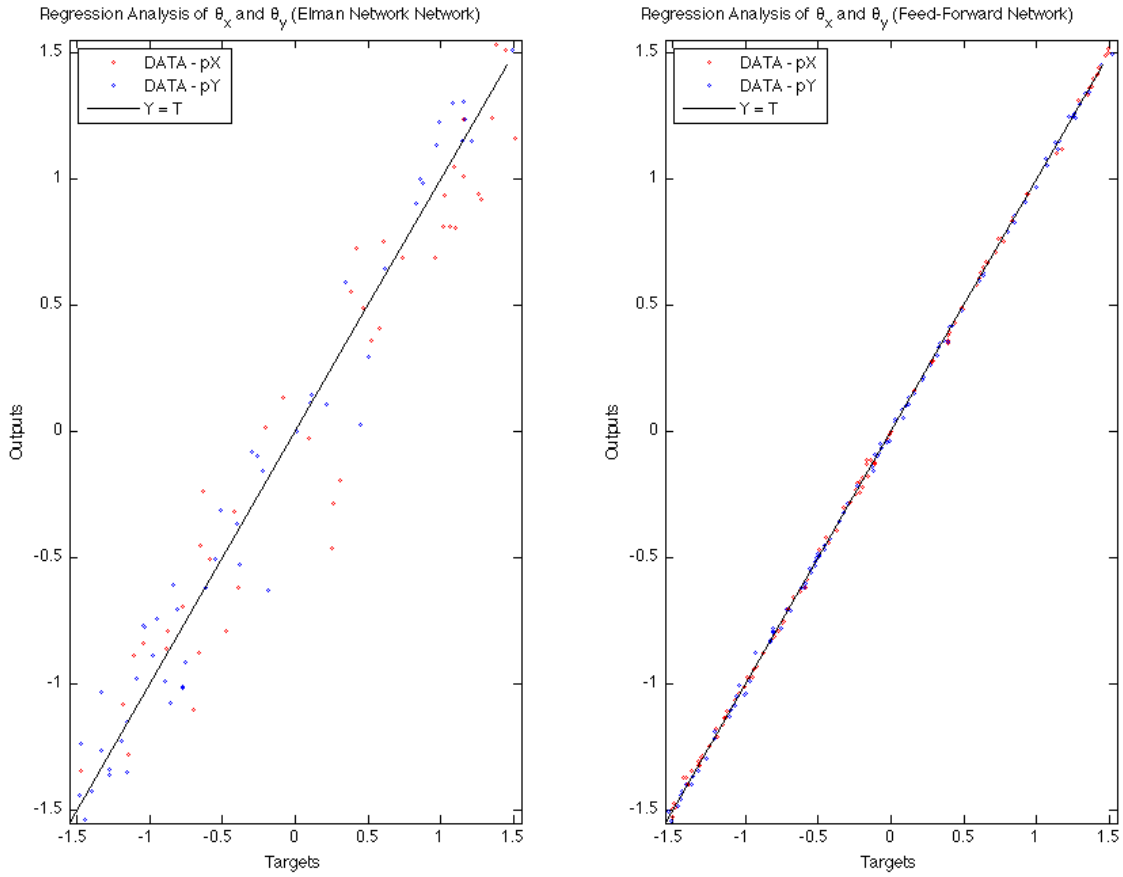


Figure 9: Regression analysis of the target and output temporal parameter values for Elman and Feed-Forward networks

The *reference* scenario, against which all scenarios are evaluated, assumes that 32 full-resolution noise-free projections are available. The neural network implementation is with the feed-forward network where the time delays are trained separately on two networks. The input to the neural network is the feature vector described previously. The rMSE values for θ_x and θ_y are 0.0062 and 0.0159 seconds, respectively.

The effect of the *number of projections* on accuracy is evaluated by changing the reference setting of $N = 32$ up to 64 and down to 16, 8, and 4. The results are shown in Figure 10. The root mean square error for 64 projections is 0.0063 and 0.0079 seconds for the *x* and *y* temporal delays. For 16 projections, they are 0.0164 and 0.0283 seconds. For 8 and 4 projections the root mean square error values increase at a gradual rate. As expected, as the number of projections decrease, the accuracy worsens. While the algorithm does not completely fail when fewer than 16 projections are used, such settings are probably of little practical use.

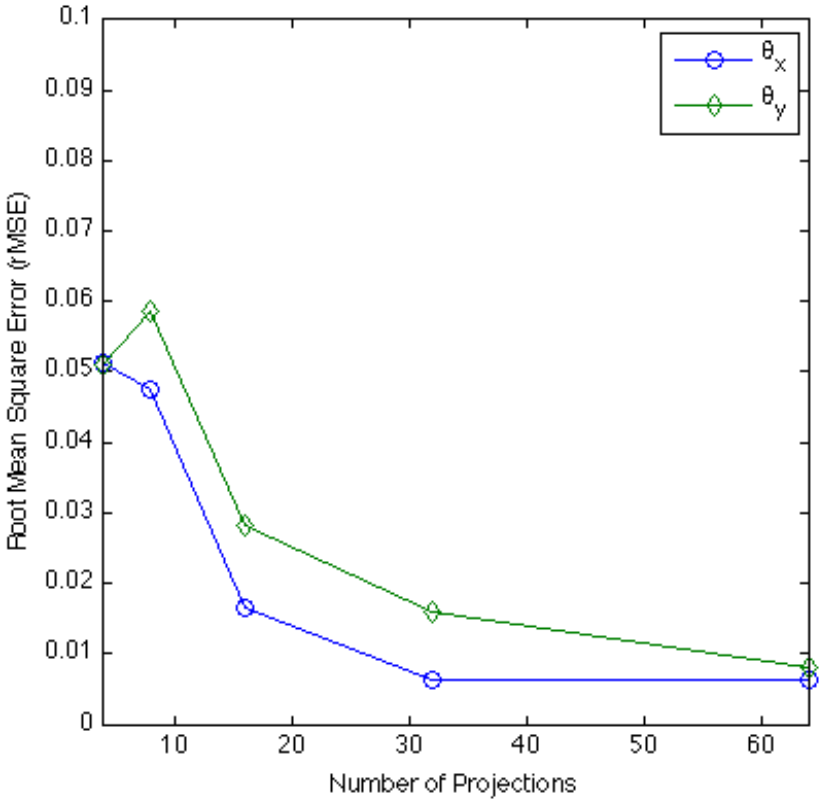


Figure 10: Estimation error as a function of the number of projections

Root Mean Square Error (seconds)	Number of Projections									
	64		32		16		8		4	
	θx	θy	θx	θy	θx	θy	θx	θy	θx	θy
1	0.0061	0.0068	0.0071	0.0157	0.0169	0.0278	0.0373	0.0589	0.0376	0.0427
2	0.0066	0.0081	0.0058	0.0150	0.0086	0.0252	0.0557	0.0550	0.0420	0.0500
3	0.0072	0.0064	0.0062	0.0178	0.0091	0.0250	0.0356	0.0598	0.0507	0.0491
4	0.0062	0.0077	0.0062	0.0161	0.0181	0.0258	0.0405	0.0549	0.0421	0.0541
5	0.0057	0.0081	0.0063	0.0160	0.0205	0.0287	0.0398	0.0643	0.0697	0.0596
6	0.0063	0.0089	0.0051	0.0139	0.0188	0.0309	0.0457	0.0613	0.0430	0.0565
7	0.0052	0.0074	0.0067	0.0164	0.0197	0.0294	0.0503	0.0586	0.0636	0.0454
8	0.0070	0.0084	0.0073	0.0163	0.0167	0.0272	0.0852	0.0558	0.0476	0.0416
9	0.0056	0.0082	0.0051	0.0150	0.0180	0.0358	0.0430	0.0633	0.0616	0.0578
10	0.0070	0.0086	0.0063	0.0173	0.0180	0.0276	0.0411	0.0555	0.0548	0.0539
Avg.	0.0063	0.0079	0.0062	0.0159	0.0164	0.0283	0.0474	0.0587	0.0513	0.0511

Table 1: 10-fold cross validation of the estimation error of the number of projections

Next, to assess the necessary *projection resolution*, the CT projections are downsampled by a factor of 2, then 4. The resolution of the parallel projections in the reference setting is equal to the spatial resolution of the reference CT image. When decreasing the projections resolution by a factor of 2, the average root mean square error values are 0.0082 seconds and 0.0168 seconds for θ_x and θ_y, respectively. When downsampled by 4, the average root means square error values are 0.0372 and 0.1048 seconds for the x and y delays. The results shown in Figure 11 and Table 2 indicate that downsampling by 2 is acceptable but the accuracy decreases significantly for higher ratios. The plot also suggests that higher resolution projections would have little benefit.

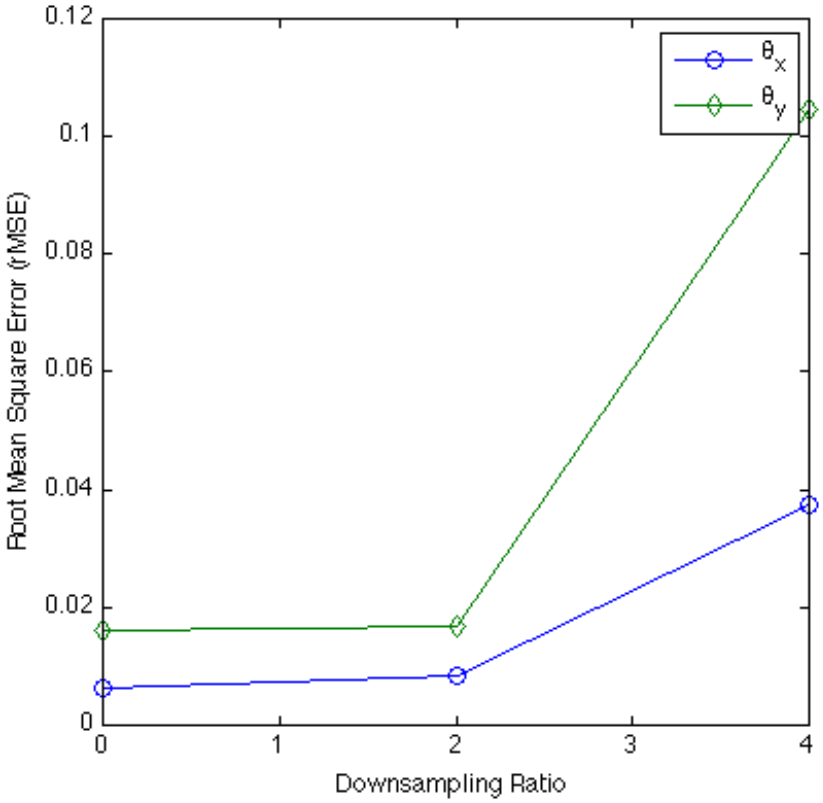


Figure 11: Estimation error as a function of the projection resolution

		Reference Scenario		Downsampled by 2		Downsampled by 4	
		θ_x	θ_y	θ_x	θ_y	θ_x	θ_y
Root Mean Square Error (seconds)	1	0.0071	0.0157	0.0070	0.0182	0.0372	0.1064
	2	0.0058	0.0150	0.0076	0.0149	0.0389	0.1020
	3	0.0062	0.0178	0.0082	0.0170	0.0372	0.1139
	4	0.0062	0.0161	0.0067	0.0163	0.0373	0.1089
	5	0.0063	0.0160	0.0088	0.0167	0.0387	0.1098
	6	0.0051	0.0139	0.0073	0.0159	0.0333	0.1005
	7	0.0067	0.0164	0.0104	0.0183	0.0344	0.1013
	8	0.0073	0.0163	0.0101	0.0164	0.0408	0.1014
	9	0.0051	0.0150	0.0076	0.0189	0.0374	0.1045
	10	0.0063	0.0173	0.0085	0.0156	0.0363	0.0991
	Avg.	0.0062	0.0159	0.0082	0.0168	0.0372	0.1048

Table 2: 10-fold cross validation of the estimation error of the projection resolution

The method's robustness in the presence of *projection noise* is also evaluated. Zero-mean white Gaussian noise is added to the projections in increasing amounts (5%, 10%, 20%, and 40%). These noise levels represent the ratio between the standard deviations of noise and projection. The results in Figure 12 show, unsurprisingly, higher estimation errors for increased noise. Noise levels higher than 10% are unacceptable.

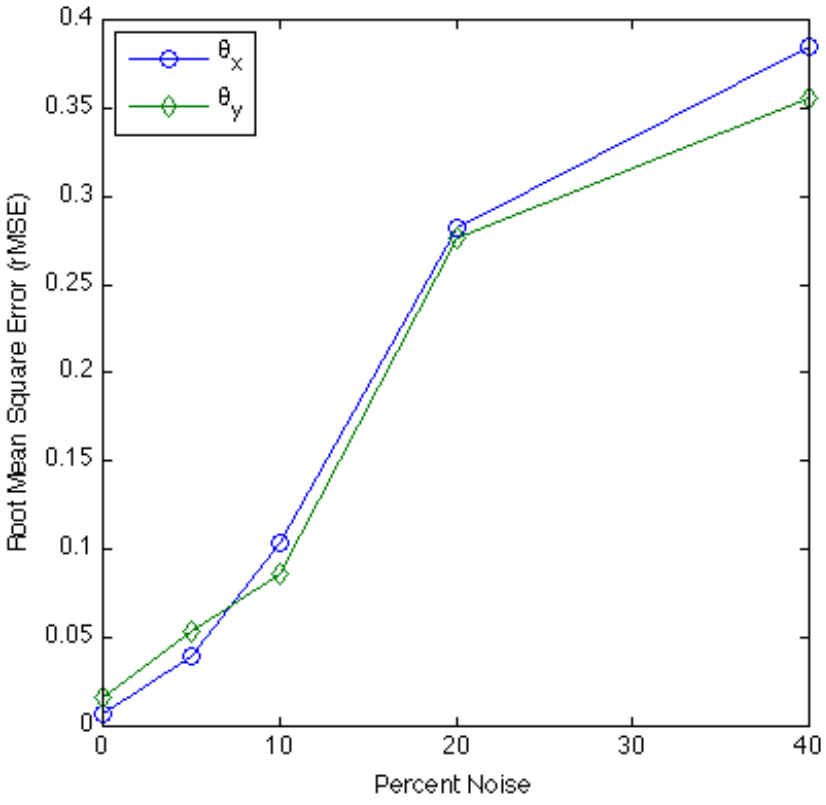


Figure 12: Estimation error as a function of the noise level

		Percent Projection Noise									
	Fold	0%		5%		10%		20%		40%	
		θ_x	θ_y	θ_x	θ_y	θ_x	θ_y	θ_x	θ_y	θ_x	θ_y
Root Mean Square Error (seconds)	1	0.0071	0.0157	0.0358	0.0527	0.1038	0.0815	0.2832	0.2584	0.3784	0.3247
	2	0.0058	0.0150	0.0407	0.0523	0.0936	0.1017	0.2878	0.2947	0.3911	0.3823
	3	0.0062	0.0178	0.0379	0.0584	0.1041	0.0875	0.2833	0.2605	0.3621	0.3378
	4	0.0062	0.0161	0.0415	0.0519	0.1055	0.0764	0.3035	0.2608	0.3903	0.3543
	5	0.0063	0.0160	0.0409	0.0611	0.1249	0.0912	0.2680	0.2792	0.4174	0.3673
	6	0.0051	0.0139	0.0379	0.0515	0.0840	0.0741	0.3018	0.2749	0.3916	0.3384
	7	0.0067	0.0164	0.0417	0.0600	0.1216	0.0938	0.3053	0.3014	0.4198	0.4122
	8	0.0073	0.0163	0.0422	0.0509	0.1046	0.0872	0.2735	0.2642	0.3501	0.3304
	9	0.0051	0.0150	0.0376	0.0466	0.1024	0.0811	0.2564	0.2938	0.3653	0.3744
	10	0.0063	0.0173	0.0376	0.0466	0.0874	0.0880	0.2564	0.2742	0.3913	0.3367
	Avg.	0.0062	0.0159	0.0394	0.0532	0.1032	0.0863	0.2820	0.2762	0.3857	0.3559

Table 3: 10-fold cross validation of the estimation error of the percent projection noise

In an attempt to assess the suitability of the features chosen for temporal parameter estimation, a final experiment was carried out, comparing feature-based estimation to estimation based on the entire projection sequence. The dimension of a full projection set is prohibitive (for both statistical and computational reasons), therefore we chose to downsample the reference image to 32×32 and to use at most 8 noise-free projections. This is clearly not a practical setting, however, the results

in Table 4 indicate that a larger feature set may achieve improved accuracy or allow the use of fewer projections.

	Full Projections		Features	
Number of Projections	rMSE (θ_x)	rMSE (θ_y)	rMSE (θ_x)	rMSE (θ_y)
8	0.0042	0.0050	0.0849	0.0603
4	0.0109	0.0108	0.1178	0.1136

Table 4: Suitability of the feature set: effect of full projections vs. set of statistical features

DISCUSSION AND CONCLUSIONS

A neural network model has been proposed to estimate the temporal delays in the 4D CT model described in Eq. 2. The model applies motion through a deformation vector field to a set of CT projections. Relevant features are extracted from these projections to estimate the temporal delays.

To evaluate the performance of the proposed method, two neural network implementations are compared, the Elman neural network and the Feed-Forward neural network. They are evaluated on ideal conditions with zero noise and features extracted from full resolution projections. The effect of a separate or singular neural network implementation is also considered. Imperfect conditions were introduced by reducing the number of projections, adding noise, and reducing

the projection resolution. Finally the statistical features extracted from the projection data set are compared to the full projection set.

The Elman and Feed-Forward networks were trained and tested on the same data set, and the mean square error values show that the Feed-Forward network fare outperforms the Elman network. The feed-forward network shows a decrease in error from the Elman network by 95%. From this we can infer, for the scope of this work, recurrent or feedback are not suitable.

Next the need for a singular neural network or separate networks is evaluated. With singular and separate neural networks, the root mean square error values are comparable,; however, the computation time is significantly less for the separate networks. When training and testing with separate networks, the computation time decreases by approximately 50%.

Several scenarios were considered for model where imperfect conditions were introduced. One such scenario is the where experiments were conducted to see the effect of the number of projections. Thirty-two projections were downsampled to 16, 8, and 4 projections. The projections were also increased to 64. The algorithm broke down for less than 16 projections. These results are acceptable because for a typical CT scans operate with a much larger set of projections. They usually range higher than 100 projections [6]. Thus, the feed-forward neural network model proves to be able to model the temporal delays accurately with a limited number of projections.

Later the effect of projection resolution was considered. It is seen downsampling by a factor of 2 and 4 did not show a significant increase in the

estimation error. This is significant because it suggests that increasing the resolution of the projections, which correlates to the spatial resolution of the CT image, will not have a significant improvement on the estimation error. As such, a higher resolution image and projections will not prove to have a significant benefit. As typical projections are of higher resolution, it is advantageous for computing time to a lower resolution projection data set [6].

The effect of noise was also considered. It can be seen that as the percent noise increases, the root mean square error values increase with respect to the noise. The model breaks down gradually, but noise values greater than 10% are not acceptable. This is expected as higher noise values give higher estimation errors. Literature shows that projection noise is not correlated in space or time, so increasing the spatial and temporal resolution by increasing the projection resolution may decrease effect of the amount of noise.

Finally the suitability of the feature set was considered. The feature set was compared to a full set of projections. While the projection set performs one order of magnitude better than the feature set, it is not realistic for fast computing. From the results for the previous experiments, it can be seen that the feature set is suitable for estimating the temporal delays accurately. There is still a difference between the accuracy with the full projections versus the features. This may be remedied by adding more features to the feature set. This is a possible future work that can be conducted.

To understand the significance of the results and put these rMSE numbers in context, recall that the delays take values from $-T/2$ to $+T/2$, where $T = 3.1$. Thus

the rMSE values for the reference setting are 0.6% of the breathing period, or 6% of the interval between the time bins of a typical conventional 4D-CT. An even more illuminating interpretation of these results is to consider the effect of the estimation error on the DVF accuracy. In a worst-case scenario, considering the maximum slope of the trace $b(t)$ in Eq. 2 and a maximum deformation of 3.5 cm, which is relative to normal deformation in the lungs which exhibits approximately 0.5 – 4.0 cm of deformation [4,5,6], the maximum DVF error is found to be 1.3 mm. This value is well within the accuracy of DVF estimation methods.

One limitation of the experimental setup is that the neural network is trained for a particular reference image and set of DVF basis vectors. When these change, it may therefore fail to accurately estimate θ_i values. Addressing this issue is beyond the scope of this feasibility study., and can be considered potential future work. However, the rather generic nature of the features and the substantial amount of a priori information that can be exploited (reference image, lengthy breathing trace, conventional 4D-CT) are grounds for optimism that the issue could be successfully addressed in future work. Another minor limitation is that the test DVF is perfectly modeled by Eq. 2. However, it has been shown previously that the equation accurately models realistic breathing motion [9]. Real anatomical motion will be used in future work.

It was hypothesized that temporal parameter estimation using neural networks will result in clinically accurate DVF estimation, as such a neural network-based method for estimating temporal parameters of a DVF model has been proposed. The method's average estimation error is less than 0.02 seconds, and its

worst-case DVF error is 1.3 mm. The method has been shown to accurately estimate the temporal parameters of the DVF model and has shown to have a worst-case DVF error of 1.3 mm. Considering the maximum deformation accounted for is 3.5 cm, an error of 1.3 mm is within the bounds of acceptable error. The accuracy of the proposed method has been evaluated in the presence of computational complexity constraints and CT projection noise. In the future, further experiments can be designed to evaluate the algorithm (1) with more features, (2) different reference images and DVF basis vectors, and (3) 4D moving anatomy.

CHAPTER 4

IMAGE REGISTRATION USING NON-UNIFORM SPLINES

USING NON-UNIFORM, NON-SEPARABLE SPLINES TO IMPROVE THE SPATIAL RESOLUTION OF DEFORMATION MODELS

IMAGE REGISTRATION

Image registration is a technique used to establish an exact point-to-point correspondence between two or more pictures. While the registration method proposed in this paper can be used in various imaging applications, it presented in the context of medical imaging where the images are representations of human anatomy. The images may be taken at different times, with different sensors, imaging modalities, or with different viewpoints [58,59]. Image registration can be also used to map an image of a subject to an image of a different subject or to a reference atlas image [60]. Generally, registration is accomplished by applying a spatial transformation to an image or image set such that it matches a reference image. The spatial transformation is iteratively adjusted to minimize the difference between the transformed image and the reference image through the use of an optimization algorithm [59].

Many image registration techniques assume a rigid-body, where the distance between any two points in the body is the same, regardless of the way the body is imaged. This has been shown to be a reasonable assumption for imaging the brain of a singular subject with the same or different modalities. For other parts of the body, or even imaging the brain of different subjects, this assumption does not hold, as internal organs change their position and shape during the imaging process and the size of the brain and other organs vary from individual to individual [59,60]. As such, techniques for non-rigid registration, also known as deformable image registration (DIR), have been developed to address these shortcomings [61].

Deformable image registration is used to map the movement of anatomy from one moment in time to another during imaging. This mapping is represented by a free-form deformation vector field (DVF), which relates the position of each anatomical element in the deformed image $S'(\boldsymbol{r})$ to the position in the reference image $S(\boldsymbol{r})$, as described in Eq. 1. Here, $\boldsymbol{r}$ is the 2D spatial variable and $\boldsymbol{d}(\boldsymbol{r})$ is the DVF. With this formulation, image registration becomes a problem of DVF estimation. In this paper we look at image registration for tumor tracking in time-dependent, or four-dimensional, CT studies (4D-CTs) in radiation therapy.

We know that radiation therapy is the preferred method of eradicating cancerous tumors located in the thorax and abdomen. To accurately target radiation to the tumor, its size, shape and location have to be known for both planning and delivery [6,4,8,62,48]. Image-guided radiation therapy (IGRT) techniques, such as conformal radiation therapy and intensity modulated radiation therapy, use CT imaging to optimize coverage of the target volume without affecting surrounding

healthy tissue [6,63]. A critical feature of IGRT is addressing tumor motion. The most prevalent cause for motion in the thorax and abdomen is breathing [6,11,64]. 4D-CT has been used for tumor tracking, where the delivery field follows the tumor during the entire breathing cycle, allowing the patient to breathe freely [65].

As stated previously, in the conventional 4D-CT method, a breathing trace is collected as CT projections are acquired, projections are partitioned into several bins, and a 3D-CT image is reconstructed for each time bin using filtered back-projection, resulting in multiple CTs covering a breathing period.

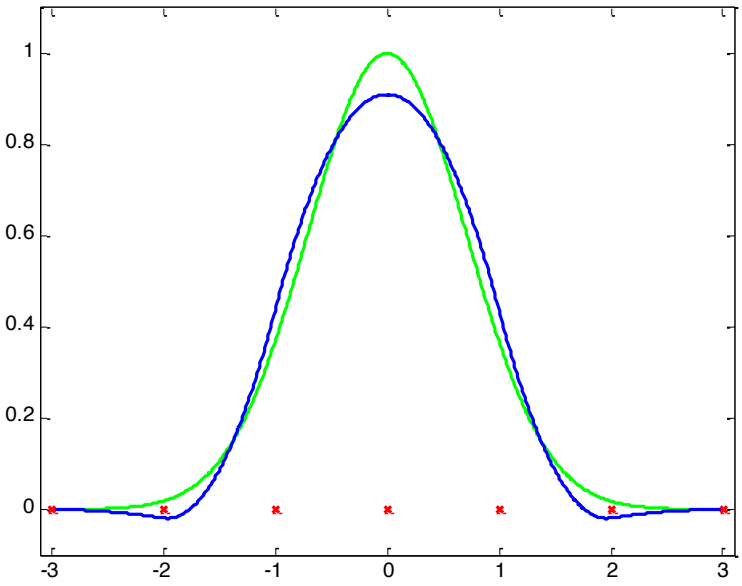
Image registration can be used to create an accurate and versatile high-resolution DVF for each pair of consecutive reconstructed CT images. A major concern for image registration is the appropriate modeling of DVFs in the presence of discontinuous motion between organs, for example the sliding motion that occurs between the pleural membranes and the lung wall due to contractions of the diaphragm from breathing inhalation and exhalation [66,67]. Estimating DVFs in image registration poses the significant challenge of modeling both the smooth movements inside organs, as well as the discontinuous motions of multiple organs moving freely [68,69].

MODELING WITH SPLINES

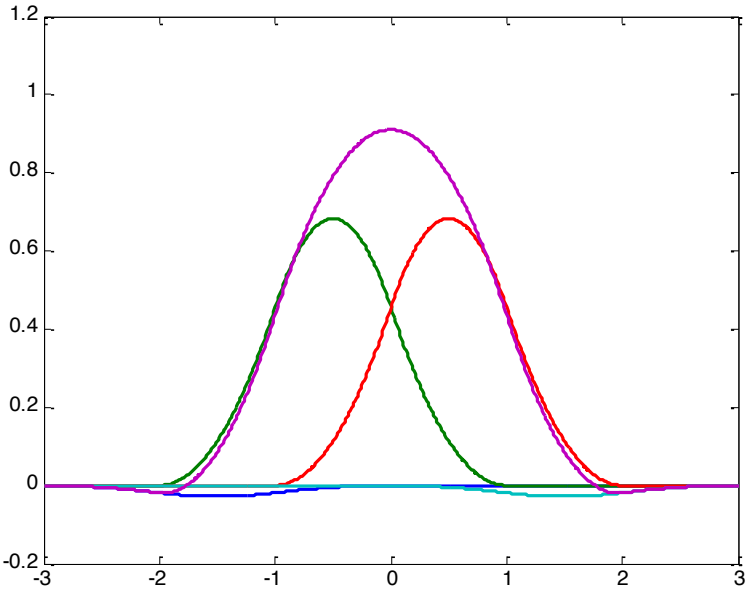
A majority of spline-based spatial modeling techniques in 4DCT imaging employ uniformly distributed splines. The basis functions are identical across the CT volume and they are centered at points on a uniform rectangular grid. It should be

pointed out that in image modeling, the splines are bivariate *surfaces* rather than the *curves* more commonly associated with the term *spline*. The disadvantage of such basis functions is that the spatial resolution is the same across the image. In most practical scenarios, anatomical details are not uniformly similar in space. One such example is the deformation, due to breathing, in the thorax. This is even more so when only part of the anatomy is in motion. As a result, models using uniform splines end up having a large number of parameters. A large number of degrees of freedom drastically diminishes the robustness of the algorithm, almost guarantees the solution will not be globally optimal, and increases computational complexity.

To illustrate the limitations of uniform splines, a few one-dimensional signals with both uniform and non-uniform splines are modeled, as shown in Figure 13, Figure 14, Figure 15, and Figure 16. Seven knots are used, depicted by x-marks on the horizontal axis, resulting in four basis functions. Even for a relatively smooth Gaussian-like function, the difference in modeling accuracy is visible. The relative RMS modeling error is 9.4% for uniform splines and 1.6% for non-uniform splines.

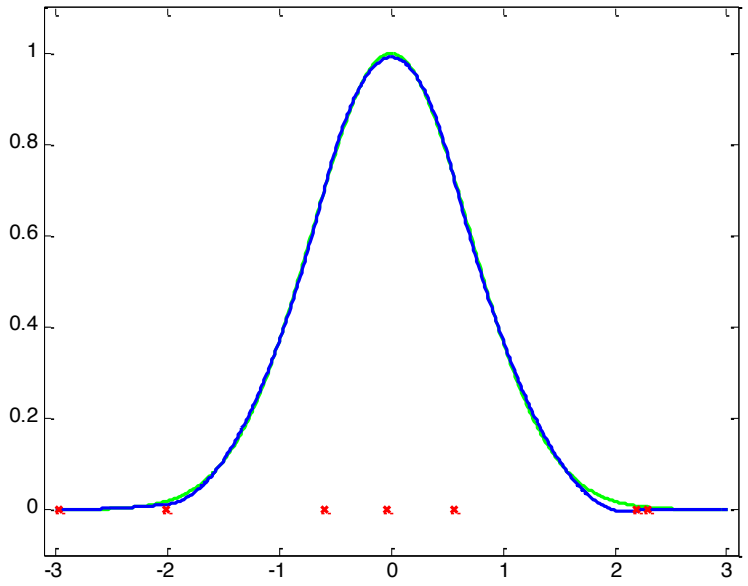


(a)

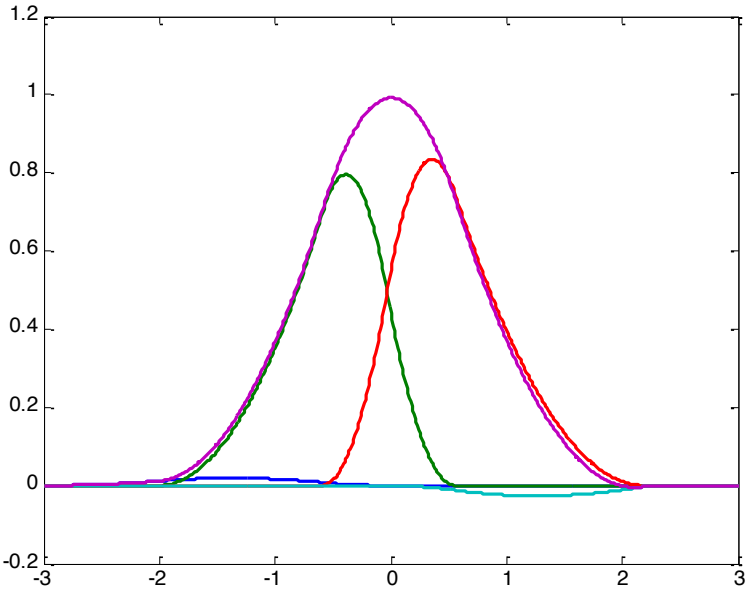


(b)

Figure 13: (a) Modeling a Gaussian signal with uniform splines; (b) The basis function decomposition

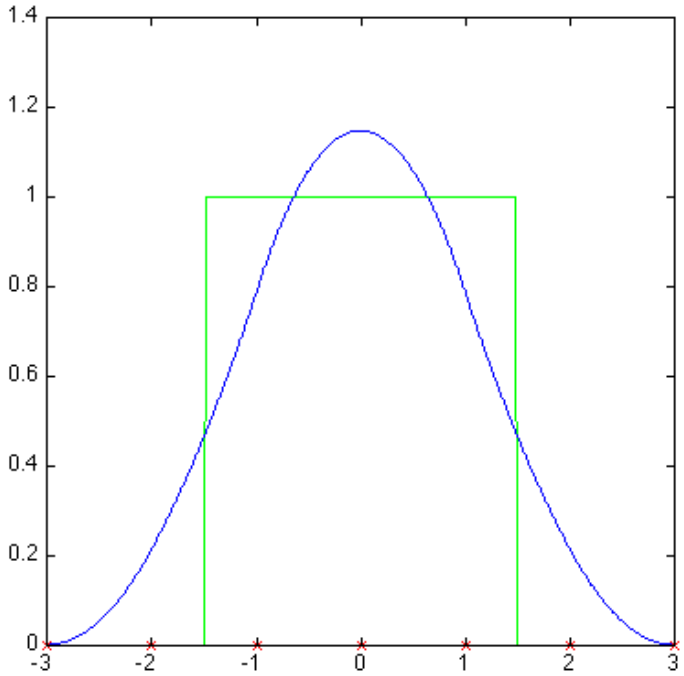


(a)

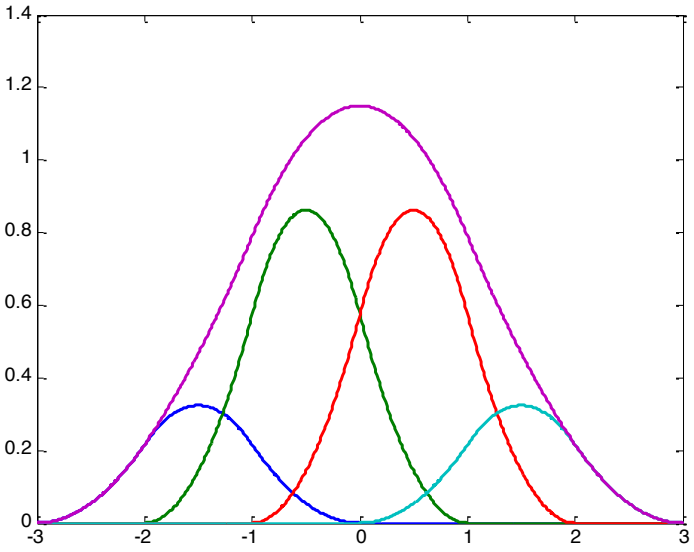


(b)

Figure 14: (a) Modeling a Gaussian signal with non-uniform splines; (b) The basis function decomposition

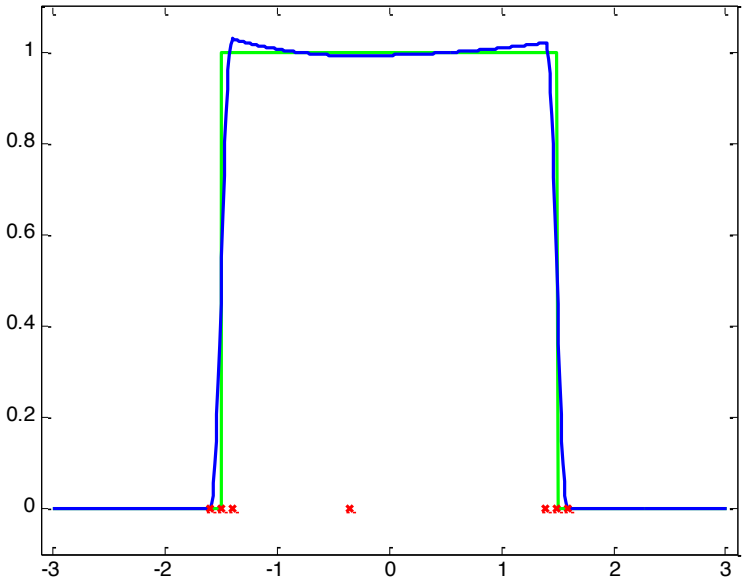


(a)

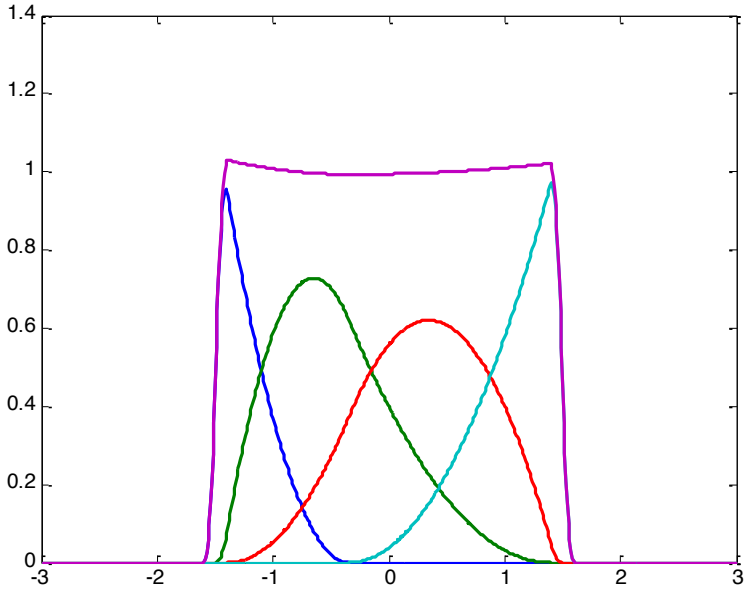


(b)

Figure 15: (a) Modeling a rectangular signal with uniform splines; (b) The basis function decomposition



(a)



(b)

Figure 16: (a) Modeling a rectangular signal with non-uniform splines; (b) The basis function decomposition

The advantage of non-uniform splines is even more visible for a signal with discontinuities, like the rectangular signal, shown in Figure 15 and Figure 16. He uniform splines blur the discontinuities present at the edges of the rectangular signal. This creates large errors. These errors can be reduced, but only with an increase in the number of knots. The advantage of non-uniform splines for a signal with discontinuities is clear when comparing the relative RMS modeling error for the same number of knots. The relative RMS modeling error is 32% for uniform splines and 8% for non-uniform splines.

In two or three dimensions, the limitations of uniform basis function placement are even more detrimental. Due to the exponential increase in degrees of freedom when going from 1D to 2D to 3D, many model parameters are wasted modeling parts of the anatomy with very little detail or very little motion. These parameters would be considerably more useful if they controlled the model in areas of high interest, such as the area around the moving tumor. In addition, while most of the anatomical motion is spatially smooth thanks to the inherent continuity of tissue, there are situations where the motion is discontinuous, for example for organs on either side of the diaphragm. Thus the DVF model needs to be able to handle discontinuous motion.

Attempts to extend DVF modeling to non-uniform splines have been hampered by the theoretical and computational complexity of multivariate splines. As a result, the only progress in this direction has been made by using separable splines [16]. These splines still use a rectangular grid for the basis functions but the

grid lines are non-uniform along each of the x, y, and z-axes. This approach works well if motion occurs in a large but spatially limited area of the anatomy, for example when the upper half of the volume exhibits significantly more motion than the lower half. But this method's performance is limited when motion occurs in a region that is not well aligned with the x, y, and z-axes. The aim of this paper is to develop a model that uses truly non-uniform splines, which can be located at arbitrary locations in the anatomy.

To justify the approach, a few test images that were modeled with uniform splines and non-uniform but separable splines were tested. Several results are presented below. In a first image featuring a vertical boundary, we first used a model using 4-by-4 basis functions positioned on a rectangular grid with x and y nodes depicted by the circle symbols. The non-separable splines produced a very rough approximation of the original. When the x and y grids are allowed to converge to optimum coordinates the approximation is significantly more accurate.

For an image with anatomical features that are not aligned with the x and y axes, such as the circular boundary shown below, the improvement obtained by allowing non-uniform x and y grids is negligible. These results indicate that separable splines fall short of the requirement for accurate modeling of realistic DVFs.

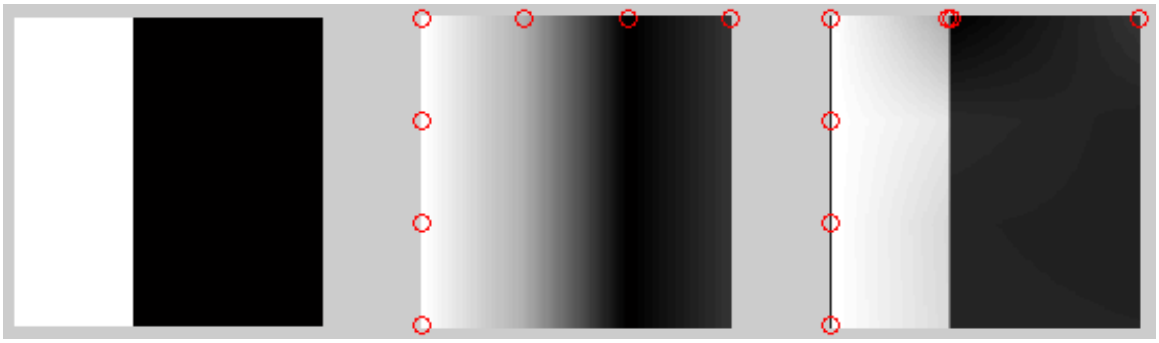


Figure 17: (a) Modeling a vertical boundary (b) using uniform separable splines (c) and non-uniform separable splines

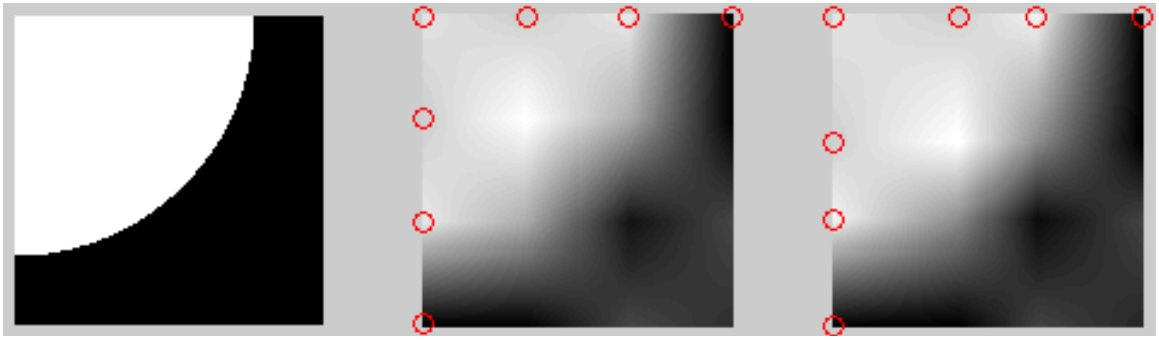


Figure 18: (a) Modeling a circular boundary (b) using uniform separable splines (c) and non-uniform separable splines

NON-UNIFORM, NON-SEPARABLE SPLINES IMPLEMENTATION

This paper seeks to develop a model that uses truly non-uniform splines, which can be located at arbitrary locations in the image. For simplicity, we chose thin-plate splines. The process by which the spline points are chosen is as follows: First, the gradients of DVF_x and DVF_y are estimated, resulting in four gradient images. A single gradient combining the variation in both the x and y directions is computed by summing the squares of these four images. The combined gradient is a local measure of the amount of discontinuity in the DVF. If N is the number of desired spline points, N/2 points are placed randomly where the combined gradient of the original DVFs is high. This is achieved by sampling a bivariate distribution equal to the combined gradient. As a result, more points are chosen in highly discontinuous areas. Finally, each of the N/2 points is split into 2 separate points a small distance apart in the direction of the gradient motion, i.e. perpendicular to the discontinuity, resulting in N spline points.

IMAGE REGISTRATION WITH NON-UNIFORM SPLINES

The image registration process that implements the proposed method is outlined in *Figure 19*. The reference image $S(\boldsymbol{r})$, and the deformed image $S'(\boldsymbol{r})$ are the inputs to the algorithm. The output is the optimal set of DVF model parameters, consisting of the amplitudes corresponding to the spline basis functions. Initially, all DVF model parameters are assumed to be zero. An initial DVF is computed using (1)

and applied to the reference image. The difference between the computed deformed image and the actual one is then used to drive a gradient-based optimization algorithm. The iterative process is continued until the mismatch converges to a minimum.

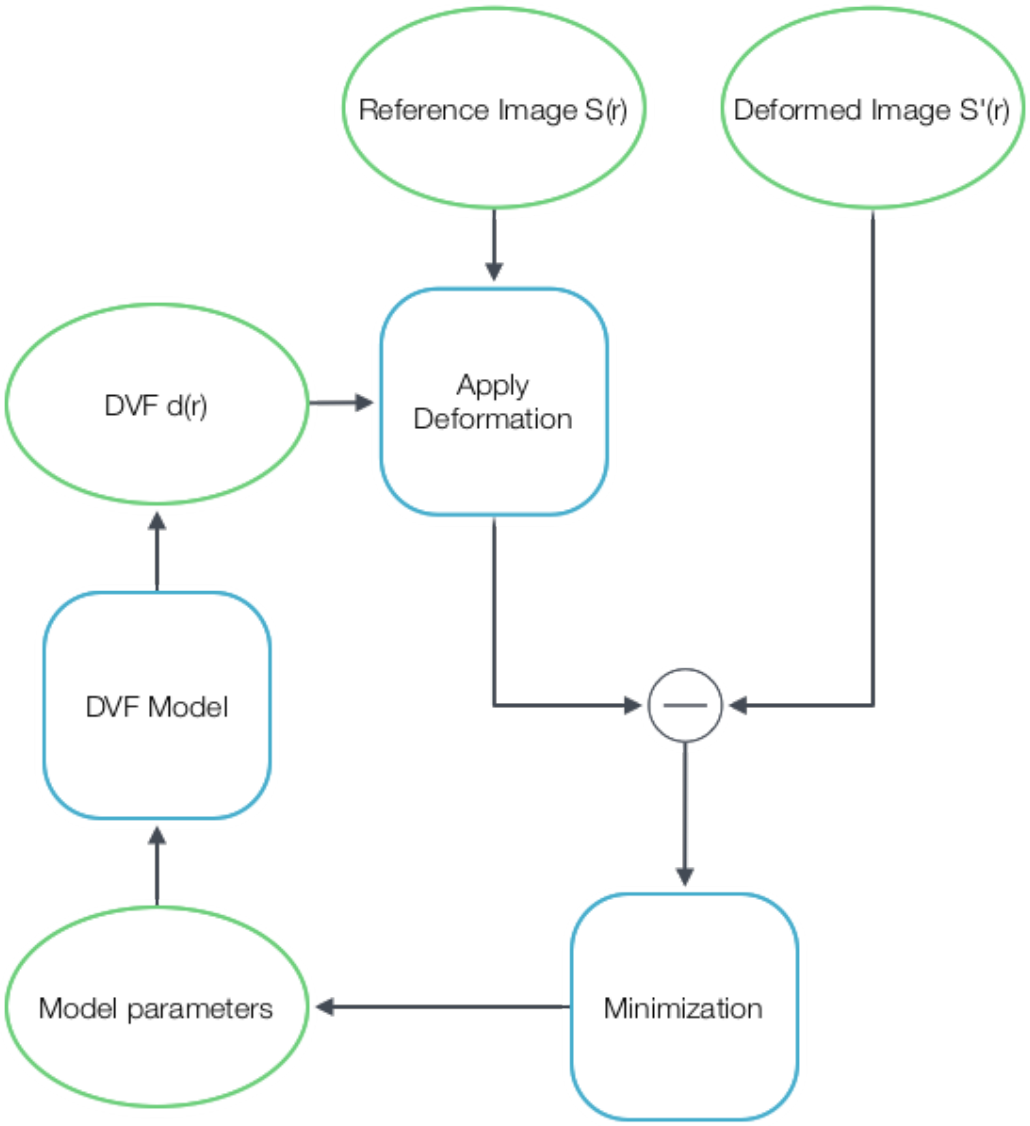


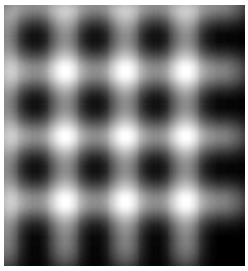
Figure 19: Schematic diagram of the image registration process.

EXPERIMENTAL SETUP

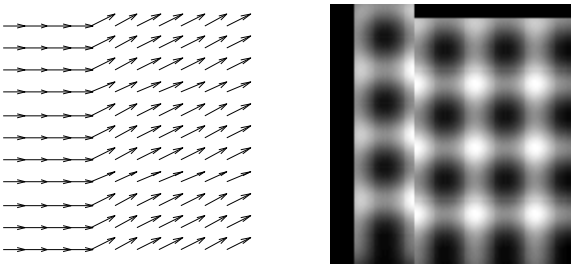
Two sets of experiments have been carried out to evaluate the proposed method's registration performance. First, the ability of non-uniform splines to model discontinuous motion is assessed. Second, the suitability of non-uniform spline-based for accurate image registration is investigated. The CT images in real clinical settings are 3D. In this paper, the feasibility of the proposed registration method is assessed in a simplified setting using 2D CT images.

IMAGE AND DISPLACEMENT DATA

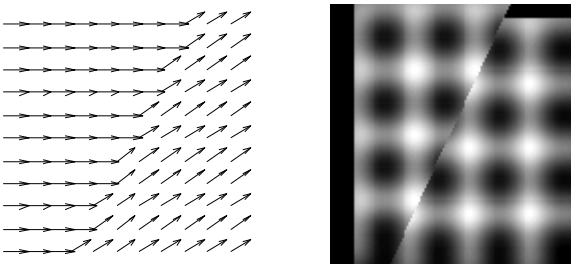
The proposed method is evaluated using a set of synthetic 2D images and DVFs. Three DVFs that represent sliding motion in the lung cavity along different types of boundary are considered: a vertical boundary, a linear diagonal boundary, and a circular boundary. The DVFs are illustrated by quiver plots in Figure 20 (b,c,d)..



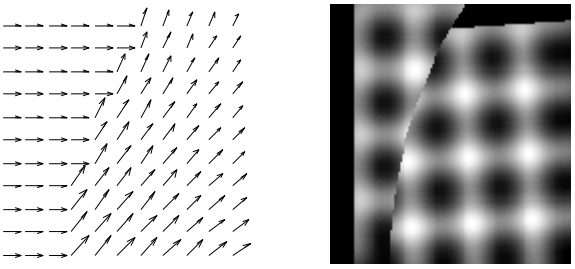
(a)



(b)



(c)



(d)

Figure 20: (a) Original reference image. (b, c, d) Test DVFs with vertical, diagonal, and circular boundaries and their corresponding deformed images.

The *vertical boundary* DVF corresponds to a global displacement to the right by 10% of the image size and a downward motion by 5% for the region to the right of the boundary. The *diagonal boundary* DVF corresponds to a global displacement to the right by 10% of the image size and a downward sliding along the boundary by about 6% for the region to the right of the boundary. The *circular boundary* DVF corresponds to a global displacement to the right by 10% of the image size and a downward rotation along the boundary by 3.6° for the region to the right of the boundary.

The test reference image is a 256x256-pixel smooth rectangular grid shown in Figure 20 (a). It is obtained by smoothing a perfect grid of pitch equal to 64 pixels in both directions with a Gaussian lowpass filter with standard deviation equal to 12.8. The three test DVFs are applied to the reference image; the three corresponding deformed images are also shown in Figure 20 (b,c,d).

MODELING EXPERIMENTS

For the first set of experiments, given the reference image, the deformed image, and the original deformation vector fields, we model a set of estimated deformation vector fields using splines and apply the resulting DVFs to the original deformed image which results in an estimated deformed image.

- Initially the DVFs are reconstructed using a set of 64 spline knots placed on a uniformly distributed grid.

- A combined gradient is computed for the original DVF and 30 spline points are placed randomly in areas where the gradient of the combined gradient is high.
- Each of the 30 knots is split into 2 separate knots placed one pixel apart in the direction of the gradient, resulting in a total of 60 knots.
- One knot is placed at each of the four corners of the image, bringing the total number of knots to 64, same as in the uniformly distributed grid.
- The estimated DVFs, both with uniformly distributed knots and randomly distributed knots, are optimized using the Levenberg-Marquardt iterative optimization method. The nonlinear least squares minimization method terminates when the mismatch in the current iteration is essentially the same (within 10-15) as in the previous iteration.

The accuracy of the method is evaluated by taking the RMS error of the estimated DVFs with the original DVFs.

REGISTRATION EXPERIMENTS

In real image registration applications, DVFs are evidently not available a priori; only the reference and deformed images are available. This makes the problem challenging for uniform spline modeling and even more so for non-uniform

splines, where an estimate of the gradient of the DVF is also necessary for spline point placement.

To concentrate on assessing the merits of non-uniform splines and avoid the difficulty of actual image registration, in this paper we focus on the challenge of choosing spline point locations from the images themselves when the DVF is actually available. The method we used is a two-step process. First, the DVF is modeled using uniform splines. The estimated DVF is applied to the reference image and the resulting image is subtracted from the original deformed image. The spline points are then randomly chosen in areas where this difference is high. The DVF is then modeled again with non-uniform splines and the modeling performance is compared to the ideal model obtained using the method described. We are confident that actual image registration, which will be addressed in future work, is achievable with a hierarchical, multi-resolution approach.

RESULTS

In the several experiments that were carried out, the image registration performance is quantified by the RMS modeling error. It is the RMS value of the difference between the actual DVF and the estimated DVF. Separate modeling errors are computed for the horizontal and vertical components of the DVF (DVF_x and DVF_y). An aggregate error is also computed as the RMS value of the *vector* modeling error. The reported error values are not normalized; they are absolute distances in an image with coordinates extending from -1 to $+1$ in both directions.

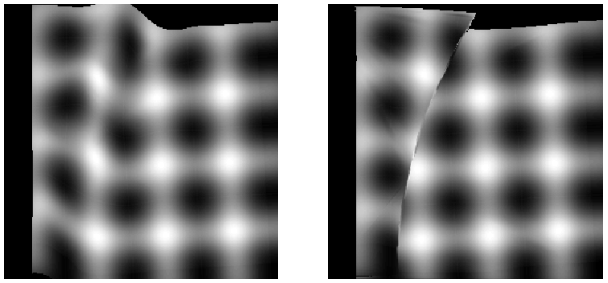
For instance, a horizontal error of 0.02 corresponds to 1% of the horizontal size of the image.

In a first experiment, the modeling ability of uniform and non-uniform splines for the three test DVFs were compared. The DVFs have been modeled using an 8x8 grid of uniform, separable splines and then using 64 non-uniform, splines with points placed intelligently using the procedure described previously in the experimental setup. For the vertical boundary, the aggregate modeling error for the DVF modeled using the uniform grid is 0.0232 and 0.0210 for the DVF modeled by the non-uniform, non-separable spline placement. For the diagonal boundary, the uniform grid had a larger aggregate error of 0.0278, and the non-uniform grid had the smaller error of 0.0177. A greater difference in error is seen with the circular boundary. The uniform grid gave an error of 0.0275 and the non-uniform grid, and error of 0.0169. The results are summarized in Table 5: Comparison of modeling performance for uniform and non-uniform splines.

	RMS Aggregate Error	
DVF	Uniform Grid	Non-Uniform Grid
Vertical	0.0232	0.0210
Diagonal	0.0278	0.0177
Circular	0.0275	0.0169

Table 5: Comparison of modeling performance for uniform and non-uniform splines.

To illustrate the modeling error, the estimated circular DVFs have been used to deform the original image and the resulting images are shown in Figure 21. Both the uniform and non-non uniform spline placements are able to accurately model the areas of little discontinuity. However, the uniform splines produced a rough estimation of the original DVF, unable to account for the sharp motion discontinuities. The non-uniform splines produce a visibly superior registered image accounting for the discontinuities.



(a) (b)

Figure 21: Modeled deformation of original image, from DVF model acquired using (a) uniform splines and (b) non-uniform splines.

A further optimization of the spline locations was attempted by allowing both spline amplitudes and locations to be free parameters and thus be simultaneously optimized. The amplitudes and locations previously obtained for the circular DVF have been used as the initial guess. The modeling error was reduced insignificantly (by about 10^{-4}); based on this result, no further efforts to optimize the locations have been made.

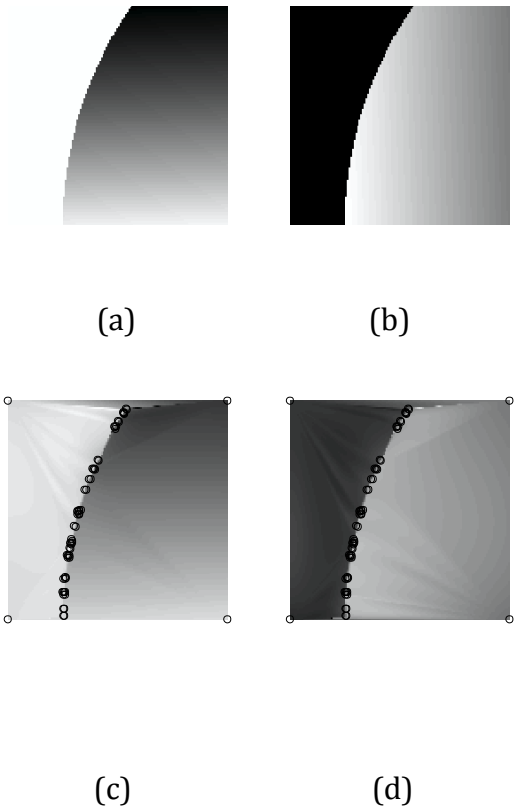


Figure 22: (a) Original DVF_x; (b) original DVF_y; (c) and (d) DVF_x and DVF_y modeled using gradient-based non-uniform, non-separable splines.

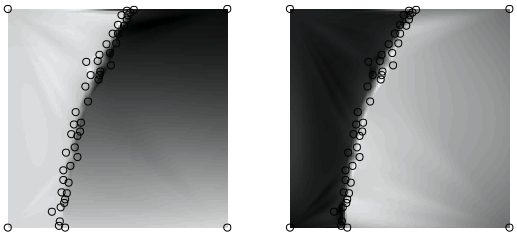
In a second experiment, we evaluated the effect of the number of free parameters on the modeling accuracy. Accuracy can always be improved by increasing the number of free parameters, but this approach has certain drawbacks.. A large number of parameters offer a larger number of degrees of freedom, which severely diminishes the robustness of the algorithm, almost guarantees that the solution will not be globally optimal, and increases computational complexity of the problem. Both uniform and non-uniform splines were used to model the circular

boundary DVF with a number of spline points varying from 36 points to 144 points. The results are shown in Table 6: Modeling error for different numbers of control points. It can be seen that as the number of free parameters increase, the aggregate error for the uniform grid reduces from 0.0310 to 0.0222. For the non-uniform spline placement, it can be seen that the aggregate error reduces from 0.0185 to 0.0112.

	RMS Aggregate Error	
Number of Spline Points	Uniform Grid	Non-Uniform Grid
36	0.0310	0.0185
49	0.0288	0.0171
64	0.0275	0.0169
144	0.0222	0.0112

Table 6: Modeling error for different numbers of control points

Next, the suitability of the proposed modeling method for image registration is evaluated as described previously in the experimental setup. The best trade-off between complexity (as given by the number of free parameters) and accuracy was obtained for 50 splines. The RMS aggregate modeling error is 0.0159. The estimated DVF and the image difference-based spline locations are shown in Figure 23.



(a) (b)

Figure 23: (a) DVF_x and (b) DVF_y modeled using image difference-based non-uniform, non-separable splines.

DISCUSSION AND CONCLUSIONS

A novel spline method has been proposed for image registration in the 4D CT model described in Eq. 2. The proposed method uses non-uniform, non-separable cubic splines to accurately model a DVF for image registration.

Multiple experiments were conducted to evaluate the image registration method, and the mean of comparison of the rMSE aggregate error between the actual and estimated DVFs. In the first experiment, three test DVFs were modeled using the conventional uniform grid method and the proposed intelligently selected non-uniform, non-separable spline method. Table 5 lists the summary of results. From the table, it can be seen the aggregate modeling error is reduced by 9%, from 0.0232 to 0.0210, for the uniform grid to the non-uniform, non-separable spline placement. This is for the vertical boundary. For the diagonal boundary, the aggregate modeling error is reduced by 36% from uniform to non-uniform. Finally

for the circular boundary the error is reduced by 39%. Not surprisingly, the least impressive improvement is achieved for the simplest, least realistic DVF that best aligns with the uniform rectangular grid. This is expected because, from the justification of the approach, uniform grids have been shown to model boundaries that are in the direction of the grid. Therefore, the uniform grid is able to fairly accurately model the vertical boundary. Even so, the non-uniform, non-separable spline based method is able to giver a smaller aggregate error, thereby modeling the vertical boundary more accurately than the grid.

The breakdown of the aggregate error in the horizontal and vertical directions depends on the nature of the DVF. For the three test DVFs, most of the gain is for the vertical deformation. For example, for the circular boundary DVF, the error is (0.0096, 0.0258) for uniform splines and (0.0070, 0.0154) for non-uniform splines.

The second method evaluated the effect of the number of free parameters on the modeling accuracy. For both uniform and non-uniform splines, it can be seen the modeling error decreased for the larger the number of free parameters. As stated previously, too many free parameters can reduce the robustness of the model. Table 6 shows the RMS aggregate error for uniform and non-uniform spline methods. The non-uniform, non-separable splines reduce the modeling error by 40% to 50% compared to the uniform splines, when going from 36 spline points to 144 spline points. A more illuminating interpretation of the results is that the accuracy achieved by 144 uniform splines can be achieved by at most 36 non-uniform splines. The number of free parameters is reduced by a factor of four. This

is significant because this shows the robustness of the non-uniform, non-separable spline model. From this we can show that the proposed model can accurately model the same DVF with a smaller number of points, thereby creating a more robust model .

The final experiment was conducted on image-to-image registration. The best result was obtained with 50 splines. The RMS aggregate modeling error is 0.0159. The RMS aggregate modeling error is 0.0159, which is slightly better than the error for 64 splines listed in Table 6. This somewhat unexpected result indicates that further research into optimal placement of spline points is warranted.

It was hypothesized that using non-uniform, non-separable splines in deformable image registration will result in greater registration accuracy than the conventional method of using uniform splines. Hence a novel method for image registration using non-uniform, non-separable cubic splines was proposed. The method has been evaluated in the context of 4D-CT imaging. Experimental results using a synthetic test image and three synthetic DVFs indicate that the proposed method reduces the modeling error by 40% on average compared to the standard methodology which uses a uniform grid of spline points. Alternately, it allows the number of free model parameters to be reduced by a factor of four while maintaining the same modeling accuracy. Therefore it can be seen that the non-uniform, non-separable spline method results in greater registration accuracy than the convention uniform grid of splines.

The promising results for the non-uniform, non-separable spline method are stimuli to further explore the possibility of new work in this field. In the future,

further experiments can be designed to determine the optimal spline placement for intelligently placing the spline knots. Further, projection-to-projection registration can be considered. The data evaluated in this work is on two-dimensional registration. Future work can be conducted on three-dimensional image registration to model the deformation in the *x*, *y*, and *z* directions. This can further expand to two-dimensional and three-dimensional projection-to-projection registration.

REFERENCES

[1] C H McCollough et al., "Strategies for reducing radiation dose in CT," *Radiol Clin N Am*, vol. 47, pp. 27-40, 2009.

[2] A Docef, M J Murphy, P J Keall, J V Siebers, and J F Williamson, "Deformed CT reconstruction from limited projection data," *International Congress Series*, no. 1281, pp. 104-108, 2005.

[3] R S Brock, A Docef, and M J Murphy, "Reconstruction of a cone-beam CT via forward iterative projection matching," *Medical Physics*, vol. 37, 2010.

[4] Laura A. Dawson and Michael B. Sharpe, "Image-guided radiotherapy: rationale, benefits, and limitations.," *Lancet Oncology*, no. 7, pp. 848-858, 2006.

[5] Y D Mutaf, J A Antolak, and D H Brinkmann, "The impact of temporal inaccuracies on 4DCT image quality," *Medical Physics*, no. 34, pp. 1615-1622, 2007.

[6] Lei Xing et al., "Overview of Image-Guided Radiation Therapy," *Medical Dosimetry*, vol. 31, no. 2, pp. 91-112, 2006.

[7] D Verellen, M De Ridder, and G Storme, "A (short) history of image-guided radiotherapy," *Radiotherapy and Oncology*, vol. 86, pp. 4-13, 2008.

[8] L A Dawson and D A Jaffray, "Advances in image-guided radiation therapy," *J Cin Oncol*, vol. 25, pp. 938-946, 2007.

[9] David Staub, Alen Docef, Robert S. Brock, and Constantin Vaman, "4D Cone-beam CT reconstruction using a motion model based on principal component analysis," *Medical Physics*, vol. 38, no. 12, pp. 6212-6220, December 2011.

[10] Elizabeth Zachariah. (2010) VCU Theses and Dissertations. Paper 2323. [Online]. http://scholarscompass.vcu.edu/etd/2323

[11] G S Mageras, A Pevsner, and E D Yorke, "Measurement of lung tumor motion using respiration-correlated CT," *Int J Radiat Oncol Biol Phys*, no. 60, pp. 933-941, 2004.

[12] V. Boldea, G. C. Sharp, S. B. Jiang, and D. Sarrut, "4D-CT lung motion estimation with deformable registration: Quantification of motion nonlinearity and hysteresis," *Medical Physics*, vol. 35, pp. 1008-1018, 2008.

[13] R Zeng, J A Fessler, and J M Balter, "Respiratory motion estimation from slowly rotating x-ray projections: theory and simulation," *Medical Physics*, vol. 32, no. 4, pp. 984-991, April 2005.

[14] D A Low, M Nystrom, and E Kalinin, "A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing," *Medical Physics*, no. 30, pp. 1264-1283, 2003.

[15] A Docef and M J Murphy, "Reconstruction of 4D deformed CT for moving anatomy," *International Journal of Computer Assisted Radiology and Surgery*, no. 3, pp. 591-598, 2008.

[16] Travis J Jacobson and Martin J Murphy, "Optimized know placement for B-splines in deformable image registration," *Medical Physics*, no. 38, p. 4579,

2011.

[17] M Baumann, T Holscher, and D Zips, "The future of IGRT - cost benefit analysis," *Acta Oncologica*, no. 47, pp. 1188-1192, 2008.

[18] M Lecchi, P Fossati, F Elisei, R Orecchia, and G Lucignani, "Current concepts on imaging in radiotherapy," *Eur J Nucl Med Mol Imaging*, no. 35, pp. 821-837, 2008.

[19] A B Wolbarst, *Physics of Radiology*. Madison, Wisconsin: Medical Physics Publishing, 2000.

[20] G Wang, H Yu, and B De Man, "An outlook on x-ray CT research and development," *Medical Physics*, no. 35, pp. 1051-1064, 2008.

[21] M J Budoff and J S Shinbane, "Overview. In: Cardiac CT Imaging: Diagnosis of Cardiovascular Disease," in *Computed tomography*. London: Springer, 2010, pp. 3-20.

[22] David P Gierga et al., "Quantification of respiration-induced abdominal tumor motion and its impact on IMRT dose distributions ," *International Journal of Radiation Oncology, Biology, Physics*, vol. 58, no. 5, pp. 1584-1595, 2004.

[23] S Shimizu et al., "Detection of lung tumor movement in real-time tumor-tracking radiotherapy.," *International Journal of Radiation Oncology, Biology, Physics*, vol. 51, no. 2, pp. 304-310, October 2001.

[24] G T Chen, J H Kung, and E Rietzel, "Four-dimensional imaging and treatment planning of moving targets," *Frontiers of radiaton therapy and oncology*, vol. 40, pp. 59-71, 2007.

[25] S S Korreman, T Juhler-Nøttrup, and A L Boyer, "Respiratory gated beam delivery cannot facilitate margin reduction, unless combined with respiratory correlated image guidance," *Radiotherapy & Oncology*, vol. 86, no. 1, pp. 61-68, January 2008.

[26] P Keall, "4-dimensional computed tomography imaging and treatment planning," *Seminars in Radiation Oncology*, no. 14, pp. 81-90, 2004.

[27] Cynthia L Eccles, Kristy K Brock, Jean-Pierre Bissonnette, Maria Andreia Hawkins, and Laura A Dawson, "Reproducibility of liver position using active breathing coordinator for liver cancer radiotherapy," *International Journal of Radiation Oncology Biology Physics*, vol. 64, no. 3, pp. 751-759, 2006.

[28] Laura A Dawson et al., "The reproducibility of organ position using active breathing control (ABC) during liver radiotherapy," *International Journal of Radiation Oncology, Biology, Physics*, vol. 51, no. 5, pp. 1410-1421, December 2001.

[29] B Schaly, J A Kempe, G S Bauman, J J Battista, and J Van Dyk, "Tracking the dose distribution in radiation therapy by accounting for variable anatomy," *Phys Med Biol*, no. 49, pp. 791-805, 2004.

[30] V Boldea, G C Sharp, S B Jiang, and D Sarrut, "4D-CT lung motion estimation with deformable registration: Quantification of motion nonlinearity and hysteresis," *Medical Physics*, no. 35, pp. 1008-1018, 2008.

[31] Jan-Jakob Sonke, Lambert Zijp, Peter Remeijer, and Marcel van Herk, "Respiratory correlated cone beam CT," *Medical Physics*, vol. 32, no. 4, pp. 1176-

86, 2005.

[32] S S Vedam et al., "Acquiring a four-dimensional computed tomography dataset using an external respiratory signal.," *Physics in Medicine and Biology*, vol. 48, no. 1, pp. 45-62, January 2003.

[33] P J Keall, G Starkschall, and H Shukla, "Acquiring 4D thoracic CT scans using a multislice helical method," *Phys Med Biol*, no. 49, pp. 2053-2067, 2004.

[34] E Rietzel, T Pan, and G TY Chen, "Four-dimensional computed tomography: Image formation and clinical protocol.," *Medical Physics*, no. 32, pp. 874-889, 2005.

[35] T Pan, T Y Lee, E Rietzel, and G T Chen , "4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT," *Medical Physics*, vol. 31, no. 2, pp. 333-340, 2004.

[36] P J Keall et al., "Four-dimensional radiotherapy planning for DMLC-based respiratory motion tracking," *Medical Physics*, no. 32, pp. 942-951, 2005.

[37] E C Ford, G S Mageras, E Yorke, and C C Ling, "Respiration-correlated spiral CT: A method of measuring respiratory-induced anatomic motion for radiation treatment planning," *Medical Physics*, no. 30, pp. 88-97, 2003.

[38] J H Killoran, A M Allen, B H Kann, and Y Lyatskaya, "Inter fractional variability of breathing phase definition as determined by fiducial location," *Medical Physic*, vol. 35, no. 2, pp. 753-763, 2008.

[39] Yuwei Chi, Jian Liang, Xu Qin, and Di Yan, "Respiratory motion sampling in 4DCT reconstruction for radiotherapy," *Medical Physics*, vol. 39, no. 4, pp. 1696-

1703, 2012.

[40] A F Abdelnour et al., "Phase and amplitude binning for 4D-CT imaging," *Physics in Medicine and Biology*, vol. 52, no. 12, pp. 3515-3529, May 2007.

[41] J Wong, M B Sharpe, and D A Jaffray, "The use of active breathing control (ABC) to reduce margin for breathing motion," *Int J Radiat Oncol Biol Phys*, no. 44, pp. 911-919, 1999.

[42] A Trofimov, E Rietzel, and H Lu, "Temporo-spatial IMRT optimization: Concepts, implementation and initial results," *Phys Med Biol*, no. 50, pp. 2779-2798, 2005.

[43] V R Kini et al., "Patient training in respiratory-gated radiotherapy," *Medical Dosimetry*, vol. 28, no. 1, pp. 7-11, 2003.

[44] C J Haasbeek et al., "Impact of audio-coaching on the position of lung tumors," *International Journal of Radiation Oncology, Biology, Physics*, vol. 71, no. 4, pp. 1118-1123, July 2008.

[45] E M Leter, F Cademartiri, P C Levendag, H Stam, and P J Nowak, "Four-dimensional multislice computed tomography for determination of respiratory lung tumor motion in conformal radiotherapy," *International Journal of Radiation Oncology, Biology, Physics*, vol. 62, no. 3, pp. 888-892, July 2005.

[46] Ju Young Song et al., "Respiratory motional effect on cone-beam CT in lung radiation surgery," *Medical Dosimetry*, vol. 32, no. 2, pp. 117-125, 2009.

[47] G Chen, J Tang, and S Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly

undersampled projection data sets," *Medical Physics*, vol. 35, pp. 660-663, 2008.

[48] M. Lecchi, P. Fossati, F. Elisei, R. Orecchia, and G. Lucignani, "Current concepts on imaging in radiotherapy," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 35, pp. 821-837, 2008.

[49] E. Yorke, K. E. Rosenzweig, R. Wagman, and G. S. Mageras, "Interfractional anatomic variation in patients treated with respiration-gated radiotherapy," *Journal of Applied Clinical Medical Physics*, vol. 6, pp. 19-32, 2005.

[50] M I Jordan and C M Bishop, "Neural Networks," *ACM Computing Surveys*, no. 28, pp. 73-75, 1996.

[51] R E Uhrig, "Introduction to articial neural networks," in *Industrial Electronics, Control, and Instrumentation, 1995., Proceeding of 1995 IEEE IECOn 21st International Conferance*, vol. 1, 1995, pp. 33-37.

[52] Simon S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 1998.

[53] J L Elman, "Finding Structure in Time," *Cognitive Science*, vol. 14, pp. 179-211, 1990.

[54] H B Demuth and M H Beale, "Elman Recurrent Neural Network," in *Neural Network Toolbox*, 4th ed. Natick, MA: MathWorks, 1996.

[55] M J Murphy and S Dieterich, "Comparative performance of linear and nonlinear neural networks to predict irregular breathing," *Physics in Medicine and Biology*, vol. 51, no. 22, p. 5903, Nov 2006.

[56] M J Murphy and D Pokhrel, "Optimization of an adaptive neural network to predict breathing," *Medical Physics*, vol. 36, no. 1, pp. 40-47, January 2009.

[57] A, Greville, T Ben-Israel, *Generalized Inverses: Theory And Applications*.: Springer, 2003.

[58] Lisa Gottesfeld Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 326-376, December 2002.

[59] Derek L. G. Hill, Philipp G. Batchelor, Mark Holden, and David J. Hawkes, "Medical image registration," *Physics in Medicine and Biology*, vol. 46, no. 3, pp. 1-45, June 2000.

[60] Brian F. Hutton, Michael Braun, Lennart Thurfjell, and Dennys Y. H. Lau, "Image registration: an essential tool for nuclear medicine," *European Journal of Nuclear Medicine*, vol. 29, pp. 559-577, 2002.

[61] D. Sarrut, "Deformable registration for image-guided radiation therapy," *Medical Physics*, vol. 16, no. 4, pp. 285-297, 2006.

[62] M Baumann, T Holscher, and D Zips, "The future of IGRT - cost benefit analysis," *Acta Oncologica*, vol. 47, no. 7, pp. 1188-1192, 2008.

[63] D. Verellen, M. De Ridder, and G. Storme, "A (short) history of image-guided radiotherapy," *Radiotherapy and Oncology*, vol. 86, pp. 4-13, 2008.

[64] P. Keall, "4-dimensional computed tomography imaging and treatment planning," *Seminars in Radiation Oncology*, vol. 14, pp. 81-90, 2004.

[65] F. Khan et al., "The use of 4DCT to reduce lung dose: a dosimetric analysis,"

Medical Dosimetry, vol. 34, no. 4, pp. 273-278, 2009.

[66] Z. Wu, E. Rietzel, V. Boldea, D. Sarrut, and G. Sharp, "Evaluation of deformable registration of patient lung 4DCT with sub-anatomical region segmentations," *Medical Physics*, vol. 35, no. 2, pp. 775-781, 2008.

[67] D. Yang et al., "4D- CT motion estimation using deformable image registration and 5D respiratory motion modeling ," *Medical Physics*, vol. 35, no. 10, pp. 4577-4590, 2008.

[68] L. Risser, F. X. Vialard, H. Y. Baluwal, and J. A. Schnabel, "Piecewise-diffeomorphic image registration: application to the motion estimation between 3D CT lung images with sliding conditions," *Medical Image Analysis*, vol. 17, pp. 182-193, 2013.

[69] D. F. Pace, S. R. Alward, and M. Niethammer, "A Locally Adaptive Regularization Based on Anisotropic Diffusion for Deformable Image Registration of Sliding Organs," *IEEE Transactions in Medical Imaging*, vol. 32, no. 11, pp. 2114 - 2126, July 2013.

[70] David Kincaid and Ward Cheney, *Numerical Analysis: Mathematics of Scientific Computing*. Singapore.

[71] Pascal Getreuer, "Linear Methods for Image Interpolation," *Image Processing On Line*, 2011.

[72] H Zhong, E Weiss, and J V Siebers, "Assessment of dose reconstruction errors in image-guided radiation therapy," *Physics in Medicine and Biology*, vol. 53, pp. 719-736, 2008.

VITA

Elizabeth Shobha Zachariah was born on August 30, 1986 in Kottayam, Kerala, India and is an American citizen. She was raised in Richmond, Virginia and graduated from Douglas S. Freeman High School in June 2004. She obtained her Bachelor of Science in Biomedical Engineering from Virginia Commonwealth University, Richmond, Virginia in May 2008 and her Master of Science in Biomedical Engineering from Virginia Commonwealth University, Richmond, Virginia in December 2010.