

Towards Physiological-based Respiratory Motion Modeling: Improvement on 4D Image  
Quality and Feature-based Deformable Image Registration

by

Xiao Liang

Graduate Program in Medical Physics  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Jing Cai, Co-Supervisor

\_\_\_\_\_  
Fang-Fang Yin, Co-Supervisor

\_\_\_\_\_  
Anuj Kapadia, Chair

\_\_\_\_\_  
Mustafa R Bashir

\_\_\_\_\_  
Lei Ren

\_\_\_\_\_  
W Paul Segars

Thesis submitted in partial fulfillment of  
the requirements for the degree of Doctor of Philosophy in the  
Graduate Program in Medical Physics  
in the Graduate School of  
Duke University

2017

ABSTRACT

Towards Physiological-based Respiratory Motion Modeling: Improvement on 4D Image  
Quality and Feature-based Deformable Image Registration

by

Xiao Liang

Graduate Program in Medical Physics  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Jing Cai, Co-Supervisor

\_\_\_\_\_  
Fang-Fang Yin, Co-Supervisor

\_\_\_\_\_  
Anuj Kapadia, Chair

\_\_\_\_\_  
Mustafa R Bashir

\_\_\_\_\_  
Lei Ren

\_\_\_\_\_  
W Paul Segars

An abstract of a thesis submitted in partial fulfillment of  
the requirements for the degree of Doctor of Philosophy in the  
Graduate Program in Medical Physics  
in the Graduate School of  
Duke University

2017

Copyright by  
Xiao Liang  
2017

## Abstract

**Purpose:** To develop methods towards physiological-based respiratory motion management, through the enhancement of image quality of four dimensional magnetic resonance imaging (4D-MRI), as well as the deformable image registration (DIR) accuracy utilizing the data commonly available in radiotherapy treatment.

**Methods:** First part of the work is on the development of a probability density function (PDF) based method to improve the image quality of 4D-MRI. The overall idea is to identify a few main breathing cycles (and their corresponding weightings) that can best represent the main breathing patterns of the patient and then reconstruct a 4D image set for each of the identified main breathing cycles. This method is implemented in three steps: (1) The breathing signal is decomposed into individual breathing cycles, characterized by amplitude,  $A$ , and period,  $T$ ; (2) Individual breathing cycles are grouped based on values of  $A$  and  $T$  to determine the main breathing cycles. If a group contains more than 10% of all breathing cycles in a breathing signal, it is determined as a main breathing pattern group and is represented by the average of individual breathing cycles in the group; (3) For each main breathing cycle, a set of 4D images is reconstructed using a result-driven sorting method adapted from our previous study. The probability-based sorting method was first tested on 26 patients' breathing signals to evaluate its feasibility of improving tumor motion displacement PDF. Breathing

motion PDFs were calculated for the original breathing signal (reference PDF), the average single-cycle breathing cycle ( $PDF_{\text{single}}$ ), and the main cycles determined by the probability-based sorting method ( $PDF_{\text{prob}}$ ). Similarity of  $PDF_{\text{single}}/PDF_{\text{prob}}$  to the reference PDF was evaluated using the Dice Similarity Coefficient (DSC) between the areas under the  $PDF_{\text{single}}/PDF_{\text{prob}}$  and under the reference PDF, noted as  $DSC_{\text{single}}/DSC_{\text{prob}}$ . The new method was subsequently tested for a sequential image acquisition scheme using the 4D digital eXtended Cardiac Torso (XCAT) phantom. Performance of the sorting methods was evaluated in terms of “tumor” volume precision and accuracy as measured by the 4D images, and also the accuracy of AIP of the 4D images.

The second part of the work is on the development of a radial basis function (RBF)-based hybrid DIR framework utilizing sparsely available measured motion information. The framework is carried out in three steps: (1) A base, intensity-based displacement vector field (DVF) is converted to an intensity-based coefficient matrix comprising expansion coefficients for the Wendland’s RBF; (2) The intensity-based coefficient matrix is modified under the guidance of sparsely distributed measured motion information to generate the hybrid coefficient matrix; (3) The hybrid coefficient matrix is converted to the hybrid DVF. The hybrid DIR framework was tested on an in-house built lung motion phantom and compared with five different DIR methods including Velocity, MIM, ILK, OHS, and Elastix. Intensity-based and hybrid DVFs were compared to the known DVFs of the phantom. Synthesized EOI (primary) images were

generated by deforming the phantom EOE (secondary) images with the intensity-based/hybrid DVFs. The intensity-based/hybrid synthesized EOI images were compared with the phantom EOI images. The effects of the number of landmarks were also studied by implementing the hybrid DIR process with 60-600 landmarks in the lungs.

The third part of the work is on the development of a method to automatically extract physiological landmarks (lung vessel bifurcations) using a novel random walking strategy. The lung images were first pre-processed with segmentation, binarizaion, and skeletonization. After the pre-processing, a vessel tree skeleton was acquired. For bifurcation extraction, first, end voxels were detected. Then a tracer would start from an end voxel and randomly move to one of its neighbors until reaching another end voxel. Numbers of being passed by the tracer were recorded for each voxel and a voxel was determined as a bifurcating voxel if it was visited more than all of its neighbors. The method was evaluated on a 7-layer simulated bifurcating tree, generated based on a modified Holton's method for generation of botanical trees.

**Results:** For the first part, improved similarity of breathing motion PDF to reference PDF was observed as compared to PDF<sub>single</sub>, indicated by the significant increase in DSC (DSC<sub>prob</sub> = 0.89±0.03, and DSC<sub>single</sub> = 0.83±0.05, p-value <0.001). Based on the simulation study on XCAT, the probability-based method outperforms the conventional phase-based methods in qualitative evaluation on motion artifacts and quantitative evaluation on tumor volume precision and accuracy and accuracy of AIP of the 4D images.

In the second part, DIR errors were reduced in all tested cases after applying the hybrid DIR framework. Intensity-based only vs hybrid 3D DIR errors per voxel are 1.27 mm vs 1.07 mm, 1.56 mm vs 1.55 mm, 6.33 mm vs 4.81 mm, 3.53 mm vs 2.94 mm, and 2.00 mm vs 1.67 mm for for Velocity, MIM, ILK, OHS, and Elastix respectively. Reduction in intensity difference between the phantom EOI images and the synthesized EOI images are also observed in the hybrid synthesized EOI images. Study on the impacts of the number of landmarks confirms that a larger amount of landmarks can improve the accuracy of the hybrid DVF, and also indicates that the hybrid DIR framework may result in a larger improvement to intensity-based DVFs of poorer accuracy.

In the third part, an algorithm based on modified Holton's model has been implemented to generate bifurcating vessel trees for testing the automatic bifurcation extraction method. Based on the results on a 7-layer bifurcating tree, the extraction method is able to detect more than 85% of the ground-truth bifurcations with 95% of the detected bifurcations located within 2 voxels from the nearest ground-truth bifurcation. The detection rate rises to 90% of the maximum detection rate in 4 iterations. The percentage of correctly detected bifurcations remains stable for different number of iterations ranging from 1 to 100. However, the variation of the percentage increases as the number of iterations lowers.

**Conclusion:** In this thesis, we developed a number of novel methods on the improvement in 4D image quality and hybrid deformable image registration, creating a more solid ground toward physiological plausible respiratory motion modeling. In particular, we have developed a probability-based multi-cycle sorting method for improved 4D image reconstruction and demonstrated the feasibility of a novel probability-based multi-cycle 4D image reconstruction method, presenting potential advantages over the conventional phase-based reconstruction method for radiation therapy motion management. In addition, we developed a hybrid DIR framework which is capable of potentially reducing registration errors by utilizing sparsely distributed measured motion information. The flexibility of the framework is expected to enhance the DIR accuracy in radiotherapy research. Furthermore, we have developed an automatic method to extract vessel bifurcations based on the random walking strategy and demonstrated its high accuracy and efficiency in a preliminary evaluation.

# Contents

Abstract .....	iv
List of Tables .....	xii
List of Figures .....	xiii
Acknowledgements .....	xvi
1. Introduction .....	1
1.1 Respiratory Motion and Motion Modeling in Radiotherapy .....	1
1.2 4D Imaging in Radiotherapy .....	2
1.2.1 Overview .....	2
1.2.2 Impacts of Breathing Motion Irregularity .....	5
1.3 DIR in Radiotherapy .....	6
1.3.1 The Displacement Vector Field (DVF).....	7
1.3.2 Current Approaches to DVF Generation .....	8
1.3.3 DIR Errors and Evaluation.....	9
1.4 Overview of the Topics.....	12
3. Development of a Probability-based Multi-cycle Sorting Method for Improved 4D Image Reconstruction.....	14
3.1 Motivation - From Single-cycle to Multi-cycle Sorting .....	14
3.2 The Probability-based Multi-cycle Sorting Method .....	15
3.3 Feasibility Study of the Probability-based Multi-cycle Sorting Method .....	19
3.3.1 Methods and Materials.....	19
3.3.2 Results .....	21

3.4 Evaluation of the Potential Clinical Efficacy of the Probability-based Multi-cycle Sorting Method Using 4D Digital Human Phantoms .....	24
3.4.1 Methods and Materials .....	24
3.4.1.1 Applicability of the probability-based multi-cycle sorting method .....	24
3.4.1.2 Simulation for 2D acquisition.....	24
3.4.1.3 Evaluation of the efficacy of the probability-based sorting method .....	27
3.4.2 Results .....	29
3.5 Discussion.....	38
4. Development of a Hybrid DIR Framework Using Sparsely-distributed Measured Motion Information .....	44
4.1 Motivation - Utilizing Sparsely-distributed Measured Motion Information.....	44
4.2 An RBF-based Hybrid DIR Framework .....	45
4.3 Investigation of the Hybrid DIR Framework on Digital Lung Phantom .....	48
4.3.1 Methods and Materials.....	48
4.3.1.1 The Tagging-MR based Thorax Digital Motion Phantom.....	48
4.3.1.2 Testing on an in-house developed lung motion phantom.....	49
4.3.2 Results .....	51
4.4 Discussion.....	58
5. Exploration of New Methods for Physiological Feature Extraction.....	62
5.1. Motivation – Automation of Bifurcation Extraction.....	62
5.2. Automatic Extraction of Vessel Bifurcations Using a Random Walking Method	62
5.2.1. Methods and Materials .....	62
5.2.1.1 Generation of a Bifurcating Tree for Vessel Tree Simulation .....	62

5.2.1.2 Vessel Tree Extraction on Patient Images.....	68
5.2.1.3 Extraction of Bifurcations Using Random Walking.....	70
5.2.2 Results .....	72
5.2.2.1 Simulated Vessel Trees.....	72
5.2.2.2 Bifurcations Extracted from the Simulated Vessel Tree .....	73
6. Conclusion .....	77
References .....	79
Biography .....	83

## List of Tables

Table 1. DSC for 26 patient breathing signals, quantifying similarity between $PDF_{\text{single}}/PDF_{\text{prob}}$ and $PDF_{\text{ref}}$ , and the difference in DSC for $PDF_{\text{single}}$ and $PDF_{\text{prob}}$ .....	23
Table 2. Quantitative results of tumor volume variation and accuracy for the two conventional sorting methods and the probability-based sorting method. ....	33
Table 3. Quantitative results of AIP difference from the reference AIP for the two conventional sorting methods and the probability-based sorting method. ....	36
Table 4. CoM shifts of the target from the reference AIP to the 4D AIPs.....	37

## List of Figures

Figure 1: Coronal images of 4D-CT (left) and 4D-MRI (right) at the same location.....	5
Figure 2: Workflow of the probability-based multi-cycle 4D reconstruction method.....	16
Figure 3: Probability-driven sorting for an example case. (a) The original breathing curve. (b) Grouping result of the individual breathing cycles. Each dot represents an individual cycle characterized by its amplitude and period. (c) Multiple trajectories resulting from the probability-driven sorting method (solid) overlaid with average single-cycle breathing cycle (dashed, repeated three times). (d) A comparison among the reference breathing motion PDF (dashed), breathing motion PDF for the single-cycle trajectory (dash-dot), breathing motion PDF for the multiple trajectories (solid).....	21
Figure 4: The study design for the comparison of the probability-based sorting method with the conventional phase sorting methods using a digital phantom.....	26
Figure 5: (a) The artificial two-pattern breathing signal. (b) The two main breathing patterns in the artificial breathing signal. (c) 10-phase 4D images generated using the phase sorting method for cine acquisition. (d) 10-phase 4D images generated using the phase sorting method for sequential acquisition. (e) 10-phase 4D images, cycle 1 (top) and cycle 2 (bottom), generated using the probability-driven sorting method. ....	31
Figure 6: (a) An example patient breathing signal acquired during CT simulation. (b) The three main breathing patterns identified with the probability-driven sorting method. (c) 10-phase 4D images generated using the phase sorting method for cine acquisition. (d) 10-phase 4D images generated using the phase sorting method for sequential acquisition. (e) 10-phase 4D images, cycle 1 (top), cycle 2 (middle), and cycle 3 (bottom), generated using the probability-driven sorting method.....	32
Figure 7: The top row shows the AIPs of the liver region generated using (from left to right) the phase sorting for cine acquisition, the phase sorting for sequential acquisition, the probability-driven sorting, and the reference, based on (a) the artificial breathing signal and (d) a patient breathing signal. The middle row shows the intensity difference maps of AIPs for (from left to right) the phase sorting for cine acquisition, the phase sorting for sequential acquisition, and the probability-driven sorting based on (b) the artificial breathing signal and (e) a patient breathing signal. The bottom row shows the profile of intensity difference along the blue lines for the phase sorting for cine acquisition (dash-dot), the phase sorting for sequential acquisition (dashed), and the probability-driven sorting (solid). ....	35

Figure 8: The framework for DIR improvement, supplementing the current framework of DIR generation in clinic. ....	46
Figure 9: Pre- and post- improvement DVFs in the superior inferior (SI) direction and 3D registration error maps are shown for the five DVFs generated with (a) Velocity, (b) MIM, (c) ILK in DIRART toolbox, (d) OHS in DIRART toolbox, and (e) Elastix. In each sub-figure, the ground-truth DVF in the SI direction is shown at the top left, pre-improvement DVF in the SI direction shown at the top middle, post-improvement DVF in the SI direction shown at the top right, pre-improvement 3D registration error maps shown at the bottom middle, post-improvement 3D registration error map shown at the bottom right. ....	52
Figure 10: The top row (a) – (e) shows the difference maps between the primary image and the image synthesized by deforming the secondary image with the pre-improvement DVFs. The bottom row (f) – (j) shows the difference maps between the primary image and the image synthesized by deforming the secondary image with the post-improvement DVFs. DIR 1 to DIR 5 correspond to Velocity, MIM, ILK, OHS, and Elastix, respectively. ....	53
Figure 11: The left figure (a) illustrates the impact of the number of ground truth displacements on the post-improvement 3D registration errors. The right figure (b) illustrates the impact of the number of ground truth displacements on the percentage post-improvement 3D registration errors normalized to the pre-improvement 3D registration errors. DIR 1 to DIR 5 correspond to Velocity, MIM, ILK, OHS, and Elastix, respectively. ....	56
Figure 12: Impact of distribution of ground-truth points on the performance of the DIR improvement framework for the five DVFs generated with (a) Velocity, (b) MIM, (c) ILK in DIRART toolbox, (d) OHS in DIRART toolbox, and (e) Elastix. In each row, the figures from left to right show the pre-improvement 3D registration error map, the post-improvement 3D registration error map for a minimum on-grid ratio of 100%, the post-improvement 3D registration error map for a minimum on-grid ratio of 50%, the post-improvement 3D registration error maps.....	57
Figure 13: Work flow for generation of a bifurcating tree. node: current node in the tree; root: the root node, the first node of the tree; depth: depth of the current node, depth of the root node set as 0; max depth: the number of bifurcating levels of the tree, or the depth of the tree; left child: left child node of the current node; right child: right child node of the current node; parent: parent node of the current node. ....	65

Figure 14: This flow chart shows how a bifurcation is generated, i.e., how two child branches are attached to the current node of interest (red dot) in one recursive step. The existing bifurcating tree is in blue with the parent branch of the current node of interest in orange.  $(x_0, y_0, z_0)$ : the coordinates of the current node of interest.  $(x_P, y_P, z_P)$ : the coordinates of the other end of the parent branch.  $\lambda$ : length of the parent branch.  $\lambda_1$ : length of the left child branch (green).  $\lambda_2$ : length of the right child branch (green).  $\alpha_1$ : magnitude of the angle at which the left child branch rotates about the y axis.  $\alpha_2$ : magnitude of the angle at which the right child branch rotates about the y axis.  $\beta$ : magnitude of the angle at which the entire bifurcation (green) rotates about the z axis. 67

Figure 15: (a) The workflow of deriving vessel bifurcations from clinical images. Operations of interest are highlighted. (b) Example of extracted lung vessels trees (yellow) with bifurcation landmarks (red) based on a high quality MR image of the lungs (isovoxel 1.75 mm using radio TrueFISP sequence)..... 70

Figure 16: Snapshots of the vessel trees generated with different sets of parameters, (a)  $\alpha=60^\circ$ ,  $\beta=360^\circ/\psi*\text{level}$ , (b)  $\alpha=60^\circ+(\text{level}-1)*10^\circ$ ,  $\beta=360^\circ/\psi*\text{level}$ , (c)  $\alpha=30^\circ$ ,  $\beta=360^\circ/\psi*\text{level}$ , (d)  $\alpha=60^\circ$ ,  $\beta=180^\circ/\psi*\text{level}$ ,  $\psi=(1+\text{sqrt}(5))/2$ . In each row, snapshots of a simulated vessel tree taken at multiple angles are shown. .... 72

Figure 17: Detected (red dots) and ground-truth bifurcations (green dots) superimposed on the simulated vessel tree..... 74

Figure 18: Number of bifurcations detected vs number of iterations..... 74

Figure 19: Percentage of detected bifurcations within (a) 1 voxel (red), (b) 1.5 voxels (green), (c) 2 voxels (blue), and (d) 3 voxels (black) of the nearest ground-truth bifurcation. .... 75

## **Acknowledgements**

I would like to express my gratitude to Dr. Jing Cai and Dr. Fang-Fang Yin for their insightful comments and patient guidance for this study. I would also like to thank the other members of my thesis committee, Dr. Anuj Kapadia, Dr. Lei Ren, Dr. W Paul Segars and Dr. Mustafa Bashir, for their time and effort. I would like to give a special thanks to Juan Yang for her help throughout my work and specifically in DVF generation. Last but not least, I would like to thank all the students in our laboratory for their constructive suggestions and perspectives.

# 1. Introduction

## 1.1 *Respiratory Motion and Motion Modeling in Radiotherapy*

Cancers in the thorax and abdomen may move during respiration. Respiratory variations in patient anatomy are potential sources of error in radiotherapy. If not properly managed, they can lead to deleterious effects at every stage of image-guided radiation therapy (IGRT), such as artifacts during image acquisition [1-4], inaccurate target delineation during treatment planning [5, 6], enlarged target volume during radiation delivery. These uncertainties may contribute to suboptimal treatment outcome by underdosing the target and/or overdosing neighboring healthy tissues [7-9]. This is especially of concern in stereotactic body radiation therapy (SBRT), which produces a much higher fractional dose (10-30 Gy/fraction) than conventional radiation therapy (1.5-3.0 Gy/fraction) delivered to very small target volumes [10, 11].

Image artifacts caused by respiratory motion are characterized by blurring or discontinuity of moving tissues in the images [4, 12-18]. For fast imaging techniques such as computed tomography (CT), moving tumor/tissues may be extended or compressed in the directions of breathing motion, causing potential error in target volume delineation and dose calculation in radiation therapy. For slow imaging techniques such as cone-beam CT (CBCT)[6], single-photon emission computerized tomography (SPECT)[19] and positron emission tomography (PET)[20], the resulting images reflect the averaged intensity of the anatomy throughout the breathing cycle.

To estimate and correct for the effects of respiratory motion in radiation therapy, various respiratory motion modeling methods have been developed [21, 22]. As direct measurement is difficult, respiratory motion is usually estimated as the displacements between corresponding spatial points by performing deformable image registration (DIR) on multi-phase images. Although DIR has been widely utilized in radiotherapy-related research, many studies have shown that the performance of DIR is highly variable between different applications, depending on both image quality (artifacts, low intensity, etc.) and registration process. DIR errors are commonly seen and well expected in current applications, which may affect the accuracy of motion modeling and subsequently the management of respiratory motion in radiation therapy[23]. Therefore, it is warranted that the two fundamental tools in radiotherapy, four-dimensional (4D) imaging and DIR, be revisited and further improved, which, in turn, may enhance the quality of motion modeling.

## ***1.2 4D Imaging in Radiotherapy***

### **1.2.1 Overview**

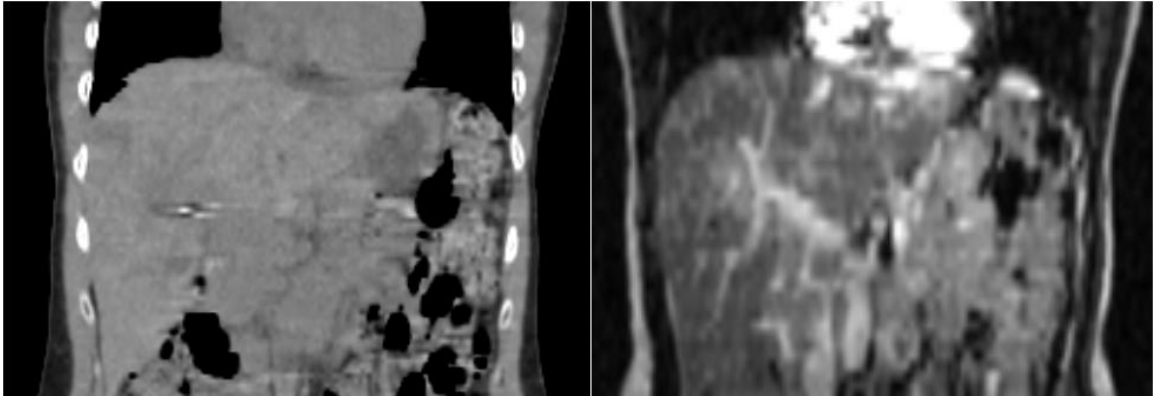
4D imaging techniques have been developed to account for breathing motion for various imaging modalities[1, 3, 12, 18, 22, 24], aiming at reducing or eliminating the breathing-related motion artifacts by sorting image data into separate respiratory bins, providing a clear shot of anatomy at each phase of the breathing cycle. In IGRT, 4D imaging plays an essential role, often performed as the first step of motion management

to obtain comprehensive respiratory motion information. 4D images reveal detailed, dynamic information of anatomy during breathing, and are used to determine treatment target volumes for mobile tumors, to estimate the 4D dose summation, as well as to study organ biomechanics (such as lung ventilation) for treatment response applications. The current standard of 4D imaging technique is 4D-CT, widely used for evaluation of respiratory motion in clinical radiation therapy [4, 6, 16-18]. 4D CT works in cine mode, meaning that, instead of one image being acquired at each slice, multiple images will be acquired continuously at each slice location for a minimal duration of a respiratory cycle. Each image will be subsequently assigned to a respiratory phase of the breathing cycle. Images from all slices in the same frame form a volume and anatomic changes can be seen between different frames. A unique advantage of 4D-CT is that it provides information of electron density, which is necessary for dose calculation.

While having proved effective for cancers in the thorax, 4D-CT lacks sufficient soft-tissue contrast for cancers in the abdomen and pelvis. In contrast, magnetic resonance imaging (MRI) provides improved tumor and soft-tissue contrast compared to CT, and involves no radiation hazard. 4D-MRI is an emerging technique and not yet a routine procedure in radiotherapy. However, its role in radiotherapy is promising, especially for applications in the abdomen and pelvis because of the high soft-tissue contrast in the abdominal region where 4D-CT images tend to exhibit a poor contrast. Absence of radiation hazard also allows prolonged acquisition, which may lead to better

image quality. Recent advances in MRI, especially in fast imaging and ultra-short echo-time (UTE) imaging, permit high quality MR imaging of the thorax [25-27]. Therefore, 4D-MRI has the potential to be the optimal technique for imaging respiratory motion management for the entire torso.

Currently, 4D-MRI can be implemented in either cine mode or sequential mode. Cine mode acquisition is similar to the acquisition method used in 4D-CT. Images at a slice location are acquired repeatedly for no shorter than a respiratory cycle and are subsequently sorted into different frames based on their phases. Sequential mode is also commonly used, especially for those sequences that do not support cine mode acquisition. In sequential acquisition, images are acquired from the first slice location to the last and the acquisition will be repeated from the first slice location again until sufficient images have been collected for every slice location. The images will be sorted into corresponding frames by matching the time stamps of the images with the associated respiratory curve. Sequential mode typically takes longer than cine mode, implying that redundant information is acquired. This redundancy makes sequential mode more robust to irregular breathing. Removal of images acquired during irregular breathing will render the 4D-MRI reconstruction impossible in the cine acquisition mode but is not likely to cause the same consequence in the sequential acquisition mode.



**Figure 1: Coronal images of 4D-CT (left) and 4D-MRI (right) at the same location.**

### **1.2.2 Impacts of Breathing Motion Irregularity**

It has been shown that the quality of 4D imaging is affected by breathing irregularity [2, 4, 6]. Clements, et al [6] studied the effect of irregular breathing patterns in 4D-CT on a Perspex thorax phantom imaged with the same acquisition settings used for stereotactic lung patients. It has been found that when irregular motion was introduced 4D-CT volumes were reduced by up to 20% and internal target volume (ITV) discontinuity was present in 4D-CT maximum intensity projection (MIP) images. In a study conducted by Sheng, et al [28], real time sagittal dynamic magnetic resonance imaging was resorted using a 4D-CT acquisition method based on phantom or patient images to simulate the imaging rebinning processes. The accuracy of average intensity projection (AIP) of 4D images was quantitatively evaluated by comparing AIPs generated from resorted dynamic MRI images and the full set of dynamic MRI images. It has been found that the AIP similarity shows a linear negative correlation with the

respiratory variability normalized by tumor sizes. The impact of irregular motions on 4D-PET image reconstruction has also been investigated by Tian, et al.[2] by evaluating the shape of reconstructed point source activity distributions. It has been found that irregular motions with amplitude fluctuation and baseline drift partly influence the target shape, causing elongation of targets in the motion direction in 4D-PET images. To cope with breathing irregularity, breathing guidance techniques such as audio instructions, video guiding interface, have been developed to facilitate regular and predictable breathing motion. Although breathing guidance techniques have been reported to effectively improve breathing regularity, applicability of these techniques is restricted to patients capable of following guidance instructions [29-31].

### ***1.3 DIR in Radiotherapy***

DIR, owing to its ability to geometrically align two images, is becoming increasingly important for managing these variations in modern radiotherapy[23, 32]. Studies have shown that distinct DIR algorithms perform differently at various clinical sites and thus need to be applied judiciously[23, 33, 34]. In the thorax, applications of DIR in radiotherapy range from four-dimensional radiotherapy[35, 36], adaptive radiotherapy[8, 37], to treatment assessment[38-43]. The accuracy of DIR algorithm directly impacts the success of its clinical application. Careful assessment of DIR is a critical necessity before they may be used to inform clinical decision making. In contrast

to the maturity of DIR in the research community, clinical implementation of DIR is much further behind. One of the significant limitations is insufficient understanding of the uncertainties associated with the use of DIR. Current methods of DIR focus on *morphological* similarity but not on the *physiological* plausibility of the deformation. That said, increased morphological similarity of the aligned data does not always imply increased registration accuracy[44]. There is a clear need to fundamentally improve the accuracy of DIR based on physiological plausibility.

### **1.3.1 The Displacement Vector Field (DVF)**

The DVF, representing correspondence between voxels in the primary and secondary volumes or frames, is introduced to describe deformation between the two volumes. DVF can be generated via a variety of commercially available software and open-source DIR programs. Typically, displacement vectors are defined at each voxel in the primary volume and end at the most probable positions that voxels in the primary volume correspond to in the secondary volume. Since deformation in anatomy observed in 4D images essentially results from the respiratory motion, in theory, the DVF represents how each voxel in the primary volume move from the phase of the primary volume to the phase of the secondary volume. Therefore, the DVF can be used as quantitative motion information for accumulated dose calculation.

Since images are composed of voxels with fixed physical size whereas tissues may change in shape, the anatomy represented in a single voxel in a phase can expand

to several voxels in another phase and tissue occupying multiple voxels in a phase can shrink to only one voxel in another phase. Due to this characteristic, it is noteworthy that the correlation described by the DVF is not voxel-to-voxel or physically invertible.

### **1.3.2 Current Approaches to DVF Generation**

DIR algorithms can be generally divided into two categories: intensity-based and feature-based. Intensity-based DIR algorithms are focused on minimizing the intensity difference between the deformed source image and the target image, with little to none consideration on actual physiological motion. Feature-based DIR algorithms basically first identify pairs of matching features in the source and target images. Displacement vectors at non-feature locations are interpolated from the displacement vectors of the features. In this approach, density of features will affect the accuracy of interpolation and for most clinical images, it may take enormous time to manually identify sufficient amount of landmarks for feature-based DIR. A combination of the two approaches may perform better than using either one alone. We aim to develop a technique capable of reducing errors in DIR by integrating sparsely-distributed physiological ground truth of motion information into DVFs generated by DIR algorithms.

Hybrid DIR methods, i.e., DIR based on both image intensity and image features, have been recently developed for image registration [38, 45-47]. It is sound to expect that hybrid DIR can improve the registration accuracy (or, the physiological plausibility of the DIR) because physiologically-measured ground truth deformation are used to guide

the image registration process. In the 2010 MICCAI (The Medical Image Computing and Computer Assisted Intervention Society) meeting, hybrid DIR methods were ranked high among various DIR algorithms for lung image registration[48], demonstrating their great promises for motion modeling applications.

### **1.3.3 DIR Errors and Evaluation**

It has been widely recognized that DIR is subject to registration errors. Although certain restraints are used in DIR algorithms to ensure smoothness and avoid unrealistic changes of DVF, no consideration of physiological motion of real anatomy is integrated in the DVF optimizing process. Consequently, a DVF may be a mathematical solution to the optimization problem but makes no sense physiologically. Therefore, DIR must be carefully examined and valuated against known truth, which is not always readily available. In recognition of the importance of DIR accuracy, performance of DIR on 4D-CT of the lung has been extensively examined using phantoms and patients' data. In a multi-institutional study[23], post-registration mean absolute errors in lung 4D-CT were 0.5-1.2 mm (left-right (LR)), 0.4-1.9 mm (anterior-posterior (AP)), and 0.7-1.9 mm (superior-inferior (SI)), all of which are less than the slice thickness, 2.5 mm. As respiration also induces considerable motion in the liver, post-registration mean absolute errors in liver were found in the same study to be 0.9-1.9 mm (LR), 1.2-3.3 mm (AP), and 1.2-7.4 mm (SI). This increased registration error may be attributed to poor soft tissue contrast in liver 4D-CT images.

The sources of DIR errors vary for different DIR algorithms as they employ different transformation models, matching criteria, and optimization methods. Based on matching criterion, DIR is often categorized into two groups: feature-based and intensity-based. Feature-based DIR methods focus on geometrically matching the landmarks, and use some form of interpolation to derive the displacements at non-landmark locations. As a result, the accuracy of feature-based DIR is usually high near the landmarks, but low far from the landmarks. On the other hand, intensity-based DIR methods aim to match two images based on image intensity, treating all image voxels equally and therefore the entire image is taken into account during registration. Intensity-based DIR methods typically perform better at regions with good contrast than regions lacking intensity contrast [49, 50]. Current intensity-based DIR methods focus on morphological similarity but not on the physiological plausibility of the deformation. Studies have shown that an increased morphological similarity of the aligned data does not always imply increased registration accuracy[44].

It is common to use digital and physical phantoms for DIR evaluation. The phantoms can tolerate virtually any amount of radiation dose without raising concerns about overdose, enabling high quality images to be acquired for DVF generation. Limitations of phantoms are as obvious. Most phantoms, especially physical phantoms, are not likely to contain as detailed anatomy as in human body. Tissues are often assumed to be homogeneous, lacking internal variation in density, composition,

elasticity, etc. This homogeneity naturally leads to motion homogeneity, meaning that an organ tends to move as a whole as a rigid body, which is not the case in the human body.

Patient or volunteer study is preferred for comprehensive evaluation of DIR. Implanted markers, anatomical landmarks, and contrast agents such as hyperpolarized gas have been used to provide the ground truth. Displacements derived from DIR algorithms are compared with displacements obtained from tracking or measurement in images. Limitations still present in these approaches. Implanting markers is invasive and may not be justifiable for health volunteers and many patients. Land markers may not be suitable for images with low quality such as low contrast, high noise. For instance, the liver in 4D-CT images tends to be homogeneous without many landmarks. The major limitation of the hyperpolarized method is its limited application to the lungs.

It is worth noting that since DIR is usually evaluated via matching landmarks which in fact enhance local contrast, results of these studies are indicating DIR accuracy in high contrast region. For applications where accurate displacement vectors throughout the region of interest (including low contrast region) are required, DIR accuracy remains in question. Recent studies that specifically investigated DIR accuracy in low contrast regions concluded that DIR algorithms can yield satisfactory matching of the contour but the internal displacement vectors may be subject to large errors due to the lack of contrast [49, 50].

## **1.4 Overview of the Topics**

Motivated by the aforementioned needs, this dissertation work aims to develop methods towards physiological-based respiratory motion management. This will be achieved through the enhancement of image quality of 4D-MRI, as well as the development of physiological-driven DIR based on high quality 4D-MRI. The proposed framework, once fully developed, will provide the foundation for 4D-MRI based motion management for the entire torso for radiotherapy. It will not only allows us to more accurately evaluate tumor respiratory motion, more properly design treatment plans, and more precisely delineate target volumes, but also more reliably assess treatment response and based on which to more soundly adapt treatment strategies.

The long-term goal of this project is to develop physiological plausible respiratory motion modeling methods for precision radiotherapy applications. The specific objective of this dissertation work is to develop improved methods for 4D-MRI image reconstruction and feature-based DIR, which are two fundamentally important aspects of physiological plausible respiratory motion modeling. The overall rationale of this work is that by more effectively utilizing more physiologically-measured features in the DIR process; the physiological plausibility of the DIR will be improved, so is the accuracy of DIR. To achieve this goal, the specific aims of the dissertation are designed such that each aim will focus one aspect of the problem, from 4D image reconstruction to 4D

image registration. Each following chapter has a different emphasis but towards a common goal which is to improve DIR by employing more physiologically-based ground truth information.

## **3. Development of a Probability-based Multi-cycle Sorting Method for Improved 4D Image Reconstruction**

### ***3.1 Motivation - From Single-cycle to Multi-cycle Sorting***

Instead of eliminating breathing irregularity, previous studies suggest possibility of reducing deleterious effects of breathing irregularity by incorporating breathing variation into 4D reconstruction process. Previous studies on breathing motion variation suggest that when breathing signals become longer, the probability density function (PDF) of breathing motion displacement will approach stable[51, 52]. Based on this finding, although breathing motion varies in both amplitude and period from cycle to cycle, each patient may have a relatively stable PDF. In addition, a previous study has shown that the breathing motion PDF of a single breathing cycle is dramatically different from that true stabilized PDF resulted from many breathing cycles [28]. However, conventional 4D sorting methods generate only one cycle, incapable of revealing breathing variation information. The lack of breathing variation information in 4D images may result in inaccurate target delineation and erroneous image intensity distribution, potentially causing inaccuracy in patient setup and dose calculation. For some imaging modalities such as CT, MRI using cine acquisition, the incompleteness of breathing variation information for each slice is inherent because image data at each slice location are acquired continuously in a short period of time, typically slightly over a breathing cycle, thus experiencing little to no breathing variation. For some other modalities such as MRI using sequential acquisition, it is possible to extract breathing

variation information from raw data because image data at each slice location are acquired over relatively long intervals, typically several breathing cycles.

Based on this observation, we aim to develop a novel, probability-based sorting method capable of generating multiple breathing cycles of 4D images to overcome the aforementioned limitations of single-cycle sorting methods. First, the feasibility of improving PDF accuracy using the new sorting method will be tested on patients' breathing signals. Second, the clinical efficacy of the new method will be evaluated on digital human phantoms, and compared with conventional phase-based sorting methods in terms of motion artifacts and accuracy of AIP images.

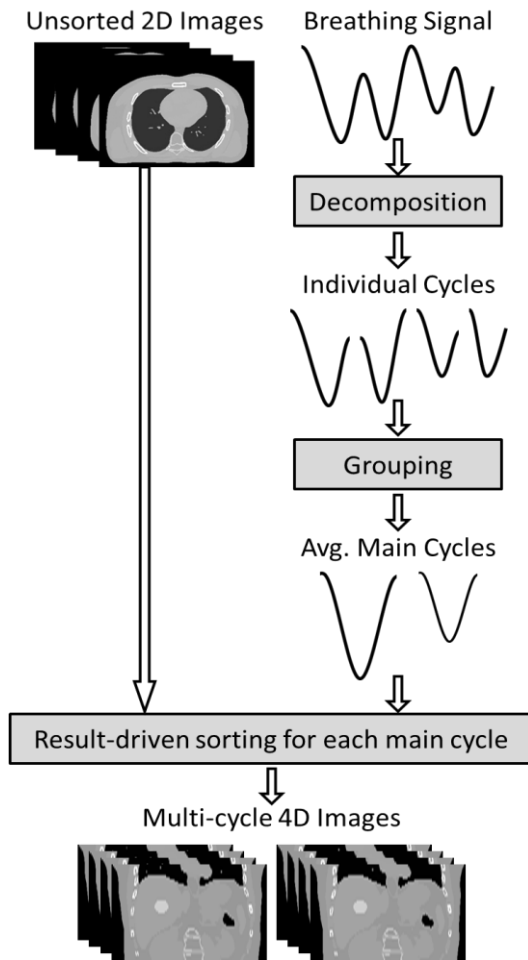
### **3.2 The Probability-based Multi-cycle Sorting Method**

Figure 2 shows the general workflow of the probability-based multi-cycle 4D sorting method. The overall idea is to identify a few main breathing cycles (and their corresponding weightings) that can best represent the main breathing patterns of the patient and then reconstruct a 4D image set for each of the identified main breathing cycles. The method is carried out in the following three steps:

(1) Decomposition of breathing signals: The breathing signal is decomposed into individual breathing cycles, which are defined as segments of the breathing signal between two consecutive end-of-exhale (EOE) peaks. Each individual cycle is characterized by amplitude,  $A$ , and period,  $T$ , calculated as follows,

$$A_i = \frac{d_{peak,i} + d_{peak,i+1}}{2} - d_{valley,i} \quad [1]$$

$$T_i = t_{peak,i+1} - t_{peak,i} \quad [2]$$



**Figure 2: Workflow of the probability-based multi-cycle 4D reconstruction method**

where  $d$  is the displacement and  $t$  is the time. Considering the complexity of patient breathing pattern, for example, fluctuation, peak detection must be manually checked before further analysis to ensure the segmented individual cycles are reasonable;

(2) Determination of main breathing cycles: Individual breathing cycles are grouped based on A and T to determine the main breathing cycles. In practice, the bin size of T is set to 0.5 seconds if the average breathing period is less than 4 seconds, and 1.0 second otherwise. To accommodate potentially large variation of breathing amplitude among different patients, the amplitude range that a group covers for each patient is chosen adaptively, being 20% of the displacement range that 95% of the data points fall in. The amplitude and period coverage is so chosen for the purpose of differentiating individual cycles with considerably different amplitudes and period while keeping the total number of identified main breathing patterns low. If a group contains more than 10% of all breathing cycles in a breathing signal, it is determined as a main breathing pattern group and is represented by the average of individual breathing cycles in the group, including an averaged motion trajectory, an average period  $T_{\text{main}}$ , and frequency of occurrence  $f_{\text{main}}$ . In practice, there are two important considerations in applying this method. First, abnormal breathing variations are undesired and should be excluded. Second, the number of main breathing cycles should be small (preferably  $< 4$ ) to achieve reasonable image acquisition time.

(3) Generation of 4D datasets for main breathing cycles: For each main breathing cycle, a set of 4D images is reconstructed using a result-driven sorting method adapted from our previous study [53]. The result-driven sorting method: As a brief introduction, the result-driven sorting method has been developed for the purpose of effectively

utilizing redundant images when multiple images are assigned to a bin. Firstly, the average respiratory trajectory is calculated from the entire breathing signal as the targeted motion result. With the average respiratory trajectory, it is possible to identify the amplitude of a bin, defined as the amplitude of the average trajectory at the center of a bin. In each bin, the image with the amplitude closest to the amplitude of this bin is selected to generate 4D-MRI dataset. In this study, an average respiratory cycle is calculated for each main breathing cycle. The full EOE to end-of-inhalation (EOI) range of the averaged motion trajectory is equally divided into 10 amplitude bins with 5 on the descending slope and 5 on the ascending slope. The total number of bins to be filled is the number of amplitude bins, 10, multiplied by the number of main breathing cycles. Each frame of acquired images is assigned to one or more of these bins based on the breathing amplitude when the frame was acquired. A frame is allowed to be assigned to bins in different main breathing cycles but will not be assigned to bins on the opposite slope at this stage. Operation for empty bins: After assigning all the frames, there may be bins with either more than one frame or no frame. For empty bins, attempts to fill the bin will be made in the following order, copying from the bin in the opposite slope, interpolating from adjacent slices at the same phase, interpolating from the same slice at adjacent phases. Operation for over-filled bins: For bins filled with multiple frames, the frame with the closest amplitude to the center of the bin is selected. The reconstruction is completed when every bin has been filled with one and only one frame.

### 3.3 Feasibility Study of the Probability-based Multi-cycle Sorting Method

#### 3.3.1 Methods and Materials

The probability-based sorting method was first tested on 26 patients' breathing signals acquired with the RPM system (Varian Medical Systems, Palo Alto, CA) to evaluate its feasibility of improving tumor motion displacement PDF. The probability-based sorting method was evaluated by comparing its PDF to that of the conventional phase-based single-cycle 4D sorting method. Each RPM breathing signal was processed so that multiple main breathing cycles with high frequency of occurrence were identified. Subsequently, a breathing motion PDF for the probability-based sorting method,  $PDF_{prob}(x)$ , was calculated as follows,

$$PDF_{prob}(x) = \frac{\sum_j p_j(x) T_{main,j} f_{main,j}}{\sum_j T_{main,j} f_{main,j}} \quad [3]$$

where  $p_j(x)$  is the breathing motion PDF for main breathing cycle  $j$ . A single-cycle motion trajectory with 20 phase bins was also generated for each RPM curve by employing the phase sorting method being widely used for 4D image sorting and reconstruction. Each bin covered 5% of the cycle. Each data point was assigned to one of the 20 bins based on its phase  $\varphi$ , determined as follows,

$$\varphi = \frac{t - t_{peak,i}}{t_{peak,i+1} - t_{peak,i}} \times 100\% \quad [4]$$

where  $t$  is the time of a data point between the  $i$ th and the  $i + 1$ th peaks. The single-cycle motion trajectory was converted to a breathing motion PDF for the phase

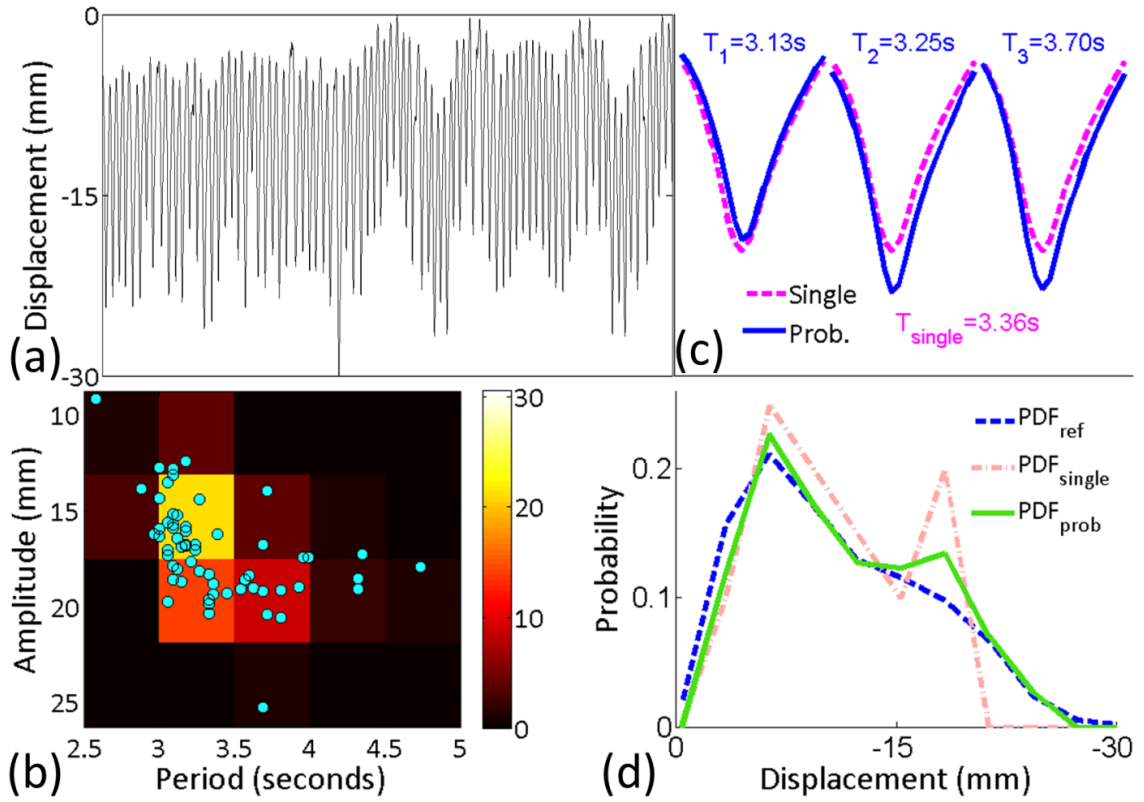
sorting method,  $PDF_{single}(x)$ . The original RPM breathing signal was also converted to a breathing motion PDF as reference,  $PDF_{ref}(x)$ . To quantitatively evaluate the similarity of PDFs from 4D sorting methods with the reference PDF, Dice Similarity Coefficient (DSC) was calculated for each of the 26 RPM breathing signals between  $PDF_{single}(x)$  and  $PDF_{ref}(x)$ , and  $PDF_{prob}(x)$  and  $PDF_{ref}(x)$ , respectively, as follows,

$$DSC_{single} = \frac{2|S_{single} \cap S_{ref}|}{|S_{single}| + |S_{ref}|} \quad [5]$$

$$DSC_{prob} = \frac{2|S_{prob} \cap S_{ref}|}{|S_{prob}| + |S_{ref}|} \quad [6]$$

where  $S_{single}$ ,  $S_{prob}$ , and  $S_{ref}$  are the area under  $PDF_{single}(x)$ ,  $PDF_{prob}(x)$ , and  $PDF_{ref}(x)$ , respectively. A paired t-test was conducted to test against the null hypothesis that the probability-based sorting method has no advantage over the phase sorting method on similarity measured by DSCs between 4D and reference motion PDF.

### 3.3.2 Results



**Figure 3: Probability-driven sorting for an example case. (a) The original breathing curve. (b) Grouping result of the individual breathing cycles. Each dot represents an individual cycle characterized by its amplitude and period. (c) Multiple trajectories resulting from the probability-driven sorting method (solid) overlaid with average single-cycle breathing cycle (dashed, repeated three times). (d) A comparison among the reference breathing motion PDF (dashed), breathing motion PDF for the single-cycle trajectory (dash-dot), breathing motion PDF for the multiple trajectories (solid).**

Figure 3 shows the results of an example case. The original breathing signal (Figure 3a) demonstrated the variations in breathing amplitude and period between individual breathing cycles. Three main breathing patterns (Figure 3b) were clustered and identified from the original breathing signal through the breathing signal analysis,

and their corresponding average breathing cycles were shown in Figure 3c. The three main breathing cycles derived from the probability-based method have shown amplitude, period and frequency of occurrence (15.3 mm, 3.13 s, 48%; 18.3 mm, 3.25 s, 32%; 18.6 mm, 3.70 s, 20%) different from those attributes of the single cycle (15.8 mm, 3.36 s, 100%). Figure 3d shows the comparison between reference motion PDF ( $PDF_{ref}$ ), PDF from single cycle sorting ( $PDF_{single}$ ), and PDF from probability-based sorting ( $PDF_{prob}$ ). It is apparent that  $PDF_{prob}$  more closely represents  $PDF_{ref}$  than  $PDF_{single}$ . There are two clear improvements in  $PDF_{prob}$  as compared to  $PDF_{single}$ . One is that the enhanced representation towards the inferior end of the respiratory motion, where  $PDF_{single}$  has a probability of zero. The other is that the smoother descending transition in  $PDF_{prob}$ , where  $PDF_{single}$  has a sharp peak against a monotonically decreasing descending slope in  $PDF_{ref}$ .

Table 1 summarizes the results on the 26 patients' breathing signals. On average, improved similarity of  $PDF_{prob}$  to  $PDF_{ref}$  was observed as compared to  $PDF_{single}$ , indicated by the increase in DSC (probability-based sorting,  $0.89 \pm 0.03$ , and single-cycle sorting,  $0.83 \pm 0.05$ ). A paired t-test showed the increase of DSC was significant with a p-value  $< 0.001$ .

**Table 1.** DSC for 26 patient breathing signals, quantifying similarity between PDF<sub>single</sub>/PDF<sub>prob</sub> and PDF<sub>ref</sub>, and the difference in DSC for PDF<sub>single</sub> and PDF<sub>prob</sub>

Case #	DSC		$\Delta$ DS C
	Singl e	Prob .	
1	0.82	0.84	0.02
2	0.83	0.88	0.05
3	0.90	0.91	0.01
4	0.83	0.90	0.07
5	0.88	0.92	0.04
6	0.93	0.90	-0.03
7	0.80	0.88	0.08
8	0.80	0.87	0.07
9	0.83	0.89	0.07
10	0.81	0.89	0.08
11	0.78	0.88	0.10
12	0.87	0.92	0.06
13	0.77	0.87	0.10
14	0.77	0.91	0.14
15	0.87	0.83	-0.04
16	0.82	0.86	0.04
17	0.88	0.96	0.08
18	0.90	0.93	0.04
19	0.86	0.92	0.06
20	0.81	0.93	0.12
21	0.82	0.90	0.08
22	0.78	0.91	0.12
23	0.79	0.90	0.11
24	0.80	0.85	0.05
25	0.85	0.89	0.04
26	0.76	0.86	0.10
<b>Mean <math>\pm</math> S.D.</b>	<b>0.83<math>\pm</math> 0.05</b>	<b>0.89<math>\pm</math> 0.03</b>	<b>P&lt;0. 001</b>

### ***3.4 Evaluation of the Potential Clinical Efficacy of the Probability-based Multi-cycle Sorting Method Using 4D Digital Human Phantoms***

#### **3.4.1 Methods and Materials**

##### **3.4.1.1 Applicability of the probability-based multi-cycle sorting method**

This method requires image data (images, or image raw data such as projections in CT and k-space data in MRI) to be randomly acquired in relative to breathing variations, i.e., different parts of the anatomy can experience the different breathing variations throughout the entire image acquisition process. As a result, this method cannot be applied to some 4D imaging techniques whose image acquisition scheme does not meet this criterion. For example, it is inapplicable to current cine/helical 4D-CT techniques, in which each slice is imaged for only a short period of time (slightly greater than one breathing cycle) and therefore does not experience the breathing variations throughout the image acquisition. As for the breathing signal, this method can be applied with various types of breathing surrogates, such as Real-time Position Management (RPM), diaphragm, tumor, internal markers, as long as it provides consistent breathing signal from the same source.

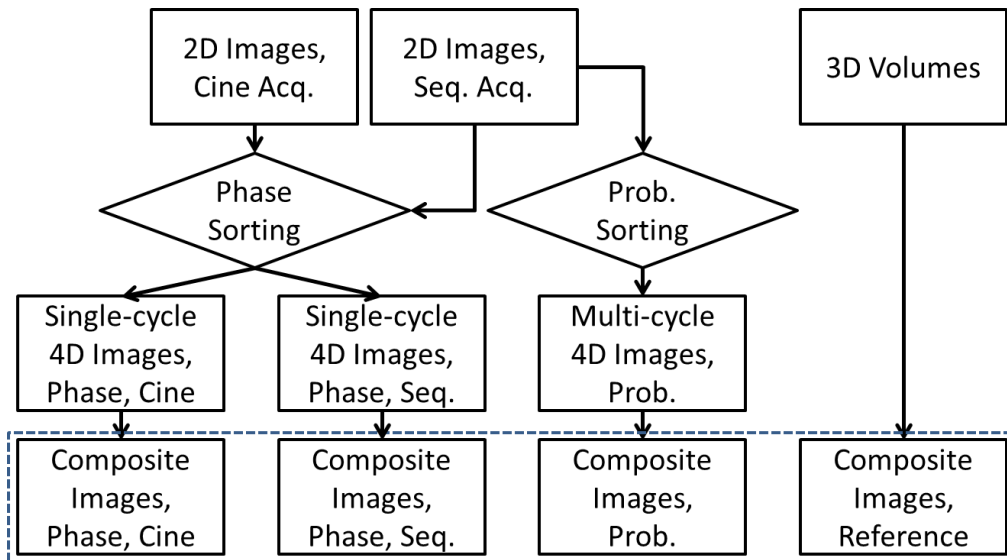
##### **3.4.1.2 Simulation for 2D acquisition**

Feasibility of probability-based multi-cycle sorting method was tested for a sequential image acquisition scheme using the 4D digital eXtended Cardiac Torso

(XCAT) phantom [54]. The sequential image acquisition scheme, which was previously demonstrated for T2-weighted 4D-MRI reconstruction [53], meets the requirement of probability-based multi-cycle 4D reconstruction since the image data is sampled randomly in relative to breathing variation.

To evaluate the potential clinical efficacy of the probability-based multi-cycle sorting method, the conventional phase-based single-cycle sorting method was also simulated using the same corresponding XCAT phantom for comparison. As outlined in Figure 4, the phase-based sorting method was simulated for two different image acquisition schemes: cine acquisition and sequential acquisition, respectively. The former is the current clinical standard method used in 4D-CT, and the latter is the same acquisition scheme as used in the probability-based method. Both image acquisition methods have been described in details in the literature [1, 53, 55, 56], and will only be briefed here. In cine acquisition, image data are continuously acquired at a slice location for at least a breathing cycle before data at the next location are acquired. The data needed for 4D reconstruction are ready when all slice locations are imaged. In sequential acquisition, a slice location is imaged only once before the next slice location is imaged. The acquisition process will repeat multiple times for the entire volume until a certain percentage of all required data are obtained. 3D volumes were also acquired at different displacements along the breathing curves to serve as reference.

The simulation of the probability-based sorting method and the conventional phase-based sorting methods can be summarized into four steps: (1) generating a transverse 2D XCAT image at the first slice location with the amplitude of the breathing signal at the first acquisition; (2) repeating Step (1) for the next acquisitions with corresponding amplitudes. For cine acquisition scheme, repeat 20 times before moving to the next slice; for sequential acquisition, move to the next slice and repeat from the first slice until 95% of the bins are filled with at least one image; (3) assign each 2D XCAT image to a bin based on the sorting method being simulated; and (4) generating simulated 4D-MRI dataset with the 2D XCAT image closest to the center of each bin.



**Figure 4: The study design for the comparison of the probability-based sorting method with the conventional phase sorting methods using a digital phantom.**

### 3.4.1.3 Evaluation of the efficacy of the probability-based sorting method

Depending on the source of breathing signal for XCAT, the simulation was performed in two types of scenarios: 1) the 4D XCAT phantom was set to move with an artificially-generated breathing signal consisting of two breathing patterns that are repeated alternatively; motion amplitudes and periods are 14.5 mm and 2.77 s, and 30.0 mm and 2.25 s, respectively. Simulation using the artificial breathing signal was to demonstrate and validate the proposed methodology in the simplest scenario; and 2) the 4D XCAT phantom was set to move with patient's RPM breathing signal (n=6). For all cases, the peak-to-peak motion amplitude for the diaphragm was set to 30 mm. The XCAT images were generated in the transverse plane using the following parameters: in-plane resolution: 1 mm; matrix size: 512x512; slice thickness: 5 mm; number of slices: 50. The simulated volume covers the upper abdominal region from the bottom of the liver at the EOI phase, to the highest point of the diaphragm at the EOE phase. To simulate the clinical images for patients with liver cancer, a spherical target of 3 cm in diameter was added in the liver and set to move with the liver during respiration. The XCAT images were generated with similar appearance to T2-weighted MRI images by using the activity mode of the XCAT phantom, in which intensities of the organs and tissues were assigned using values extracted from FRFSE T2-weighted MRI images.

Performance of the sorting methods was first evaluated in terms of target volume precision and accuracy as measured by the 4D images. The target volume precision and accuracy, labeled as  $CV_{pre}$ , and  $CV_{acc}$ , respectively, were calculated as,

$$CV_{pre} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (TV_i - \overline{TV})^2}}{\overline{TV}} \quad [7]$$

$$CV_{acc} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (TV_i - TV_{ref})^2}}{TV_{ref}} \quad [8]$$

where  $n$  is number of phases,  $TV_i$  is target volume at the phase  $i$ ,  $\overline{TV}$  is average target volume,  $TV_{ref}$  is reference target volume, CV stands for coefficient of variation.

Furthermore, the performance of the sorting methods was evaluated in terms of the accuracy of the AIP of the 4D images. This is because AIP reflects to some extent the motion probability of the tissues, and is often used for treatment planning and image guidance. For the conventional phase-based sorting methods, the AIP,  $AIP_{phase}$ , was calculated as the simple average of all phases of the 4D images. For the probability-based method, the AIP was calculated as a weighted average of the phases of the multi-cycles. Mathematically, it was calculated as:

$$AIP_{prob} = \frac{\sum_j \sum_k I_{j,k} \mu_{j,k} T_{main,j} f_{main,j}}{\sum_j T_{main,j} f_{main,j}} \quad [9]$$

where  $I_{j,k}$  is the 3D volume at the  $k$ th amplitude bin of the  $j$ th main breathing pattern,  $\mu_{j,k}$  is the ratio of the duration of the  $k$ th amplitude bin of the  $j$ th main breathing pattern to the period of the  $j$ th main breathing pattern,  $T_{main,j}$ . The reference

AIP,  $AIP_{ref}$ , used as ground truth for evaluation, was determined as the average of all image data obtained throughout the entire image acquisition period.

Two approaches were adopted to evaluation of accuracy of the AIPs. First, the evaluation was focused on image intensity using two metrics:  $dAIP_V$ , which measures intensity difference per voxel averaged over the entire volume, and  $dAIP_P$ , which measures intensity difference per voxel averaged over a profile passing through the center of target along the SI direction. They were calculated as:

$$dAIP_V = \frac{1}{n_V} \sum_{(i,j,k) \in V} |AIP(i,j,k) - AIP_{ref}(i,j,k)| \quad [10]$$

$$dAIP_P = \sqrt{\frac{\sum_{(i,j,k) \in P} (AIP(i,j,k) - AIP_{ref}(i,j,k))^2}{n_P}} \quad [11]$$

where  $n_V$  is the number of voxels in the volume,  $n_P$  is the number of voxels on the profile,  $AIP(i,j,k)$  is the image intensity of AIP ( $AIP_{phase}$ , or  $AIP_{prob}$ ) at  $(i,j,k)$  in reconstructed images, and  $AIP_{ref}(i,j,k)$  is the image intensity at  $(i,j,k)$  in reference images. Second, a more clinically relevant evaluation on accuracy of the AIPs was performed by comparing the shifts of center of mass (CoM) of the target in the 4D AIPs from CoM of the target in the reference AIP.

### 3.4.2 Results

Figure 5 shows the sorting and reconstruction results using the phase-based and the probability-based methods for the two-pattern artificial breathing signal. The artificial breathing signal (Figure 5a) represents a simplified scenario of breathing variation. The two main breathing patterns comprising the artificial signal are shown in

Figure 5b. In 4D images reconstructed using the conventional phase-based methods (Figure 5c-d), major motion artifacts can be observed in and around the target and the diaphragm regions, especially in phase 4, 5, and 6, where tissues are duplicate and/or missing. In contrast, in the two cycles of 4D images reconstructed with the probability-based sorting method (Figure. 5e), visible motion artifacts are minimal.

Figure 6 shows the sorting and reconstruction results using the phase-based and the probability-based methods for an example patient's breathing signal. Breathing variation in both amplitude and period (Figure 6a) can be appreciated. Three main breathing cycles (Figure 6b) were identified using the probability-based reconstruction method with attributes described in Sec. 3A. Single-cycle 4D images reconstructed using the conventional phase-based methods (Figure 6c, row 1, and 2) suffer from motion artifacts to a larger extent as compared to the results for the artificial breathing profile as duplicate and missing tissues can be observed in phase 4-9 for both cine and sequential acquisition. In all three cycles of the 4D images reconstructed using the probability-based method (Figure 6c, row 3-5), motion artifacts remain minimal.

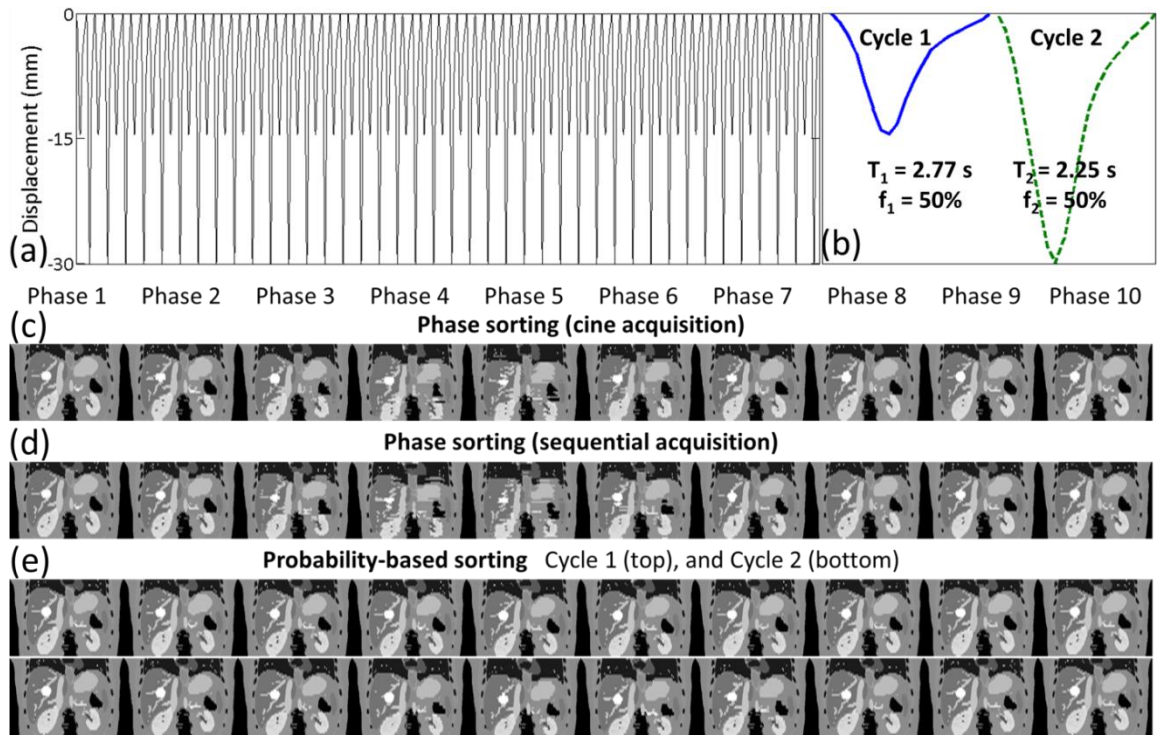


Figure 5: (a) The artificial two-pattern breathing signal. (b) The two main breathing patterns in the artificial breathing signal. (c) 10-phase 4D images generated using the phase sorting method for cine acquisition. (d) 10-phase 4D images generated using the phase sorting method for sequential acquisition. (e) 10-phase 4D images, cycle 1 (top) and cycle 2 (bottom), generated using the probability-driven sorting method.

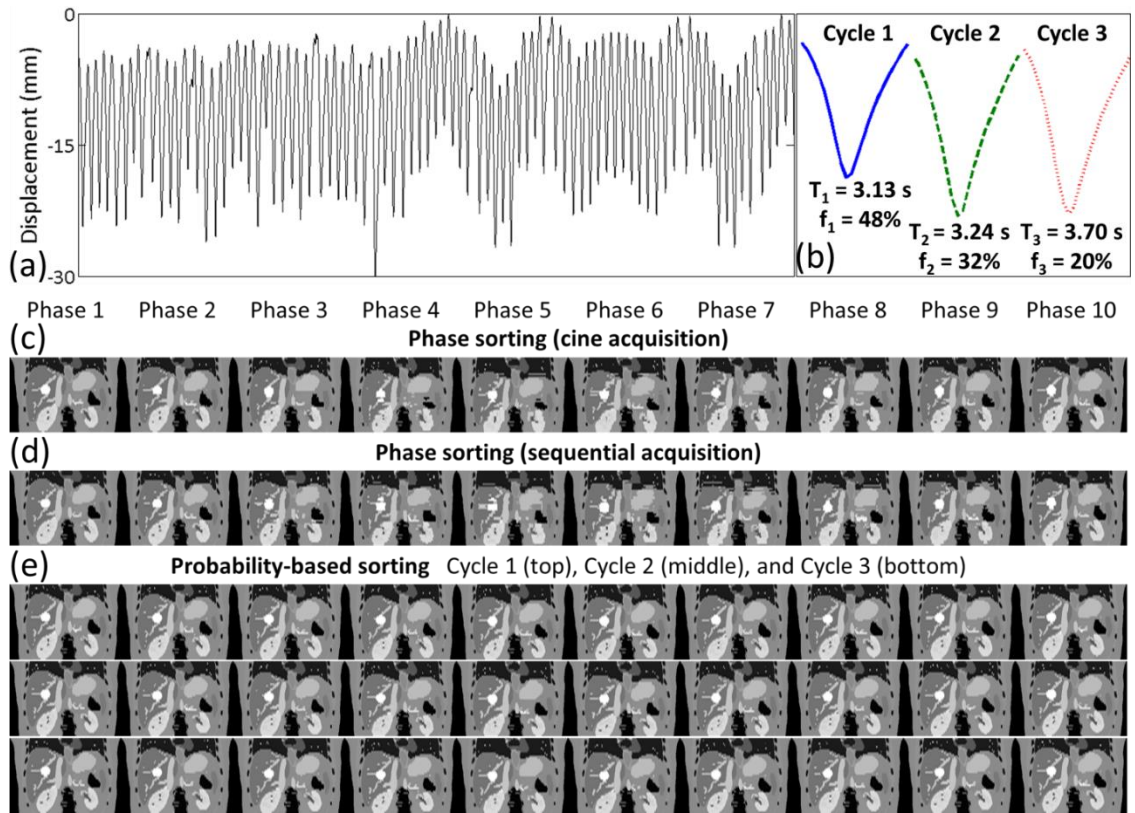


Figure 6: (a) An example patient breathing signal acquired during CT simulation. (b) The three main breathing patterns identified with the probability-driven sorting method. (c) 10-phase 4D images generated using the phase sorting method for cine acquisition. (d) 10-phase 4D images generated using the phase sorting method for sequential acquisition. (e) 10-phase 4D images, cycle 1 (top), cycle 2 (middle), and cycle 3 (bottom), generated using the probability-driven sorting method.

Table 2. Quantitative results of tumor volume variation and accuracy for the two conventional sorting methods and the probability-based sorting method.

	$CV_{pre}$			$CV_{acc}$		
	<b>Cine</b>	<b>Seq.</b>	<b>Prob.</b>	<b>Cine</b>	<b>Seq.</b>	<b>Prob.</b>
artificial	0.17	0.22	0.04	0.16	0.22	0.04
pt #1	0.12	0.14	0.02	0.13	0.13	0.02
pt #2	0.15	0.14	0.03	0.15	0.14	0.03
pt #3	0.08	0.28	0.04	0.13	0.25	0.04
pt #4	0.13	0.10	0.03	0.17	0.13	0.04
pt #5	0.10	0.17	0.04	0.11	0.17	0.04
pt #6	0.06	0.11	0.02	0.07	0.17	0.02
<b>mean±SD</b>	0.12±0.04	0.17±0.06	0.03±0.01	0.13±0.03	0.17±0.05	0.03±0.01
<b>p-value</b>	<0.001	<0.001	/	<0.001	<0.001	/

Table 2 shows the quantitative results on motion artifacts, evaluated by comparing target volumes measured at each phase against target volumes measured in 3D volumes as reference. For both the artificial and the patient cases, the probability-based sorting is able to generate targets with closest volume to reference targets on average. Best target volume consistency is also achieved by the probability-based sorting method with the lowest standard deviations of target volumes among the three reconstruction methods. In addition to target volumes,  $CV_{pre}$  and  $CV_{acc}$  were calculated to evaluate the precision and the accuracy of target volume. For both metrics, the

probability-based sorting method has shown significant lower values, indicating improved target volume consistency and accuracy.

Figure 7 demonstrated AIPs of the 4D images at a region including the liver and the diaphragm and corresponding results on the accuracy of AIP. Results for the artificial breathing profile are shown in the left (Figure 7a-c) and results for an example patient breathing signal in the right (Figure 7d-e). In the AIPs from the conventional phase-sorting methods, motion artifacts are visible in the diaphragm region, represented by discontinuity and deformation of image intensity distribution. Also, the smeared targets in AIPs from the conventional methods exhibit noticeable difference from the targets in reference AIPs in terms of intensity distribution. In contrast, the AIPs from the probability-based method show minimal motion artifacts and similar intensity distribution in the target region as compared to the reference AIPs (Figure. 7a, d). Intensity difference maps of AIPs are shown in Figure 7b, e for better visualizing the AIP accuracy for the three methods compared against the reference AIP. Since the liver in the XCAT phantom is assigned with a uniform intensity, major differences of AIPs are found at the diaphragm and the target regions for all three methods. However, the intensity differences are smaller for the probability-based method. Intensity differences along the blue line were plotted, showing that the intensity difference profile for the probability-based method has smaller maximum difference than profiles for the conventional methods for both cases (Figure 7c, f).

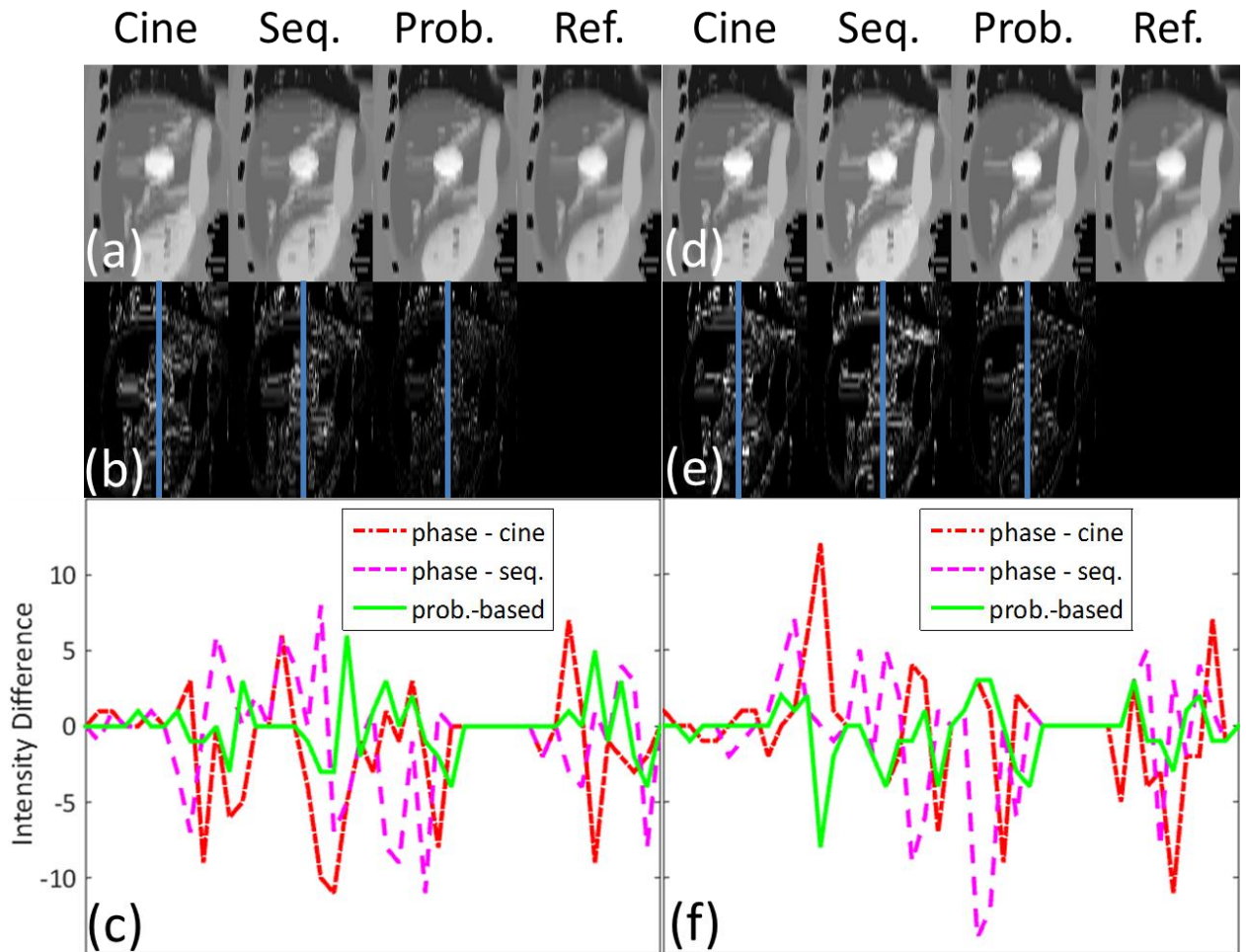


Figure 7: The top row shows the AIPs of the liver region generated using (from left to right) the phase sorting for cine acquisition, the phase sorting for sequential acquisition, the probability-driven sorting, and the reference, based on (a) the artificial breathing signal and (d) a patient breathing signal. The middle row shows the intensity difference maps of AIPs for (from left to right) the phase sorting for cine acquisition, the phase sorting for sequential acquisition, and the probability-driven sorting based on (b) the artificial breathing signal and (e) a patient breathing signal. The bottom row shows the profile of intensity difference along the blue lines for the phase sorting for cine acquisition (dash-dot), the phase sorting for sequential acquisition (dashed), and the probability-driven sorting (solid).

AIP accuracy was also quantitatively measured by  $dAIP_v$  and  $dAIP_p$ , which reflect AIP accuracy in the entire volume and along the profile, respectively, as summarized in Table 3. Based on significant decrease of both metrics, the best AIP accuracy was achieved with the probability-based sorting method.

The CoM shifts of the targets from the reference AIP to the 4D AIPs are shown in Table 4. Rank sum tests were conducted to check if there were significant differences between the target CoM shifts for the conventional phase-based sorting methods and the probability-based sorting method. Based on the results of the statistical tests, the probability-based sorting method can generate AIPs with target CoMs significantly closer to actual CoMs than the sequential acquisition phase-based sorting. Although not statistically significant based on current results, a trend was observed that the probability-based sorting method might perform better than the cine acquisition phase-based sorting in terms of proximity of target CoMs to reference target CoMs, possibly as a result of large variation between the target CoM shifts for the cine acquisition phase-based sorting.

**Table 3.** Quantitative results of AIP difference from the reference AIP for the two conventional sorting methods and the probability-based sorting method.

	$dAIP_v$			$dAIP_p$		
	Cine	Seq.	Prob.	Cine	Seq.	Prob.
artificial	1.46	1.27	0.65	3.89	3.93	1.88
pt #1	1.52	1.30	0.92	6.85	4.36	2.16

pt #2	1.36	1.42	0.81	3.69	4.05	2.01
pt #3	2.23	1.62	1.36	6.62	4.51	3.81
pt #4	1.66	1.25	0.95	5.03	3.42	2.98
pt #5	1.29	1.33	0.93	3.20	3.14	2.53
pt #6	1.60	1.34	0.89	5.87	5.23	2.65
<b>mean±SD</b>	1.59±0.31	1.36±0.13	0.93±0.22	5.02±1.47	4.09±0.70	2.57±0.67
<b>p-value</b>	0.002	0.011	/	0.002	0.002	/

**Table 4.** CoM shifts of the target from the reference AIP to the 4D AIPs.

	CoM shifts (mm)		
	Cine	Seq.	Prob.
artificial	7.1	3.3	0.1
pt #1	0.6	3.1	1.1
pt #2	3.0	3.7	2.9
pt #3	1.8	3.7	0.4
pt #4	0.2	1.5	0.5
pt #5	0.1	1.0	0.5
pt #6	1.4	1.6	0.2
<b>mean±SD</b>	2.04±2.46	2.53±1.15	0.80±0.97
<b>p-value</b>	0.30	<0.01	/

### 3.5 Discussion

In this study, a novel probability-based multi-cycle 4D reconstruction method was described and evaluated for its feasibility and potential clinical efficacy. The probability-based method was first applied on 26 patients' RPM breathing curves to extract main breathing patterns. The PDF from the probability-based method has been shown to be more accurate than the PDF from conventional single-cycle methods in terms of similarity to the reference PDF from original breathing signals. This result also indicates that the new method can be potentially used for generating more accurate AIPs. This potential was confirmed in the digital phantom simulation study, in which the XCAT 4D images were generated using an artificially generated breathing curve and six patients' RPM breathing curves. The ability of the probability-based method to generate 4D datasets with better motion PDF and thus more accurate AIPs can be theoretically deducted. Suppose a patient breathing signal  $x(t)$  continuously recorded over  $t = 0 \rightarrow T$  can be decomposed into monotonic segments,  $x_k(t)$ . For each segment, one and only one inverse function,  $t_k(x)$ , exists. Therefore,  $PDF_{ref}(x)dx$ , i.e., the probability of motion displacement staying between  $x$  and  $x + dx$ , can be expressed as,

$$PDF_{ref}(x)dx = \frac{\sum_k |dt_k|}{T} = \frac{\sum_k |t_k'(x)|}{T} dx \quad [12]$$

$$PDF_{prob}(x) = \frac{\sum_j p_j(x) T_{main,jf_{main,j}}}{\sum_j T_{main,jf_{main,j}}} \quad \text{same as [3] above}$$

The equation for calculating  $PDF_{prob}$ , introduced in Sec. 2.B, is shown above for reference. In the extreme case when each individual breathing cycle is grouped as a

main breathing cycle,  $f_{main,j} = 1, \forall j$ , as  $p_j(x)$  becomes PDF for cycle  $j$ ,  $p_j(x)T_{main,j}dx$  represents the total time of displacement staying between  $x$  and  $x + dx$  during cycle  $j$ .

Thus we have

$$\sum_j p_j(x)T_{main,j}f_{main,j}dx = \sum_j \sum_i |t_{j,i}'(x)|dx = \sum_k |t_k'(x)|dx \quad [13]$$

where  $t_{j,i}'(x)$  is the derivative of the inverse function of the  $i$ th monotonic segment in cycle  $j$ .  $\sum_j T_{main,j}f_{main,j}$  simply becomes  $T$ . Therefore, in the limit of large number of main breathing cycles,

$$PDF_{prob}(x) \rightarrow PDF_{ref}(x)$$

As for conventional single-cycle sorting methods,  $PDF_{single}(x)$  may improve as the number of phases increases but is not guaranteed to converge to  $PDF_{ref}(x)$ . The above demonstration is intended to facilitate understanding of the results in Figure 4 rather than to serve as a rigid proof of the properties of  $PDF_{prob}(x)$  as number of main breathing cycle increases.

Currently, in free-breathing treatment, on-board CBCT, effectively an AIP over multiple breathing cycles, is registered with the AIP of planning single-cycle 4D images for patient alignment. However, as introduced in Sec. 1, it has been shown that PDFs of single breathing cycles are considerably different from PDFs of multiple breathing cycles[28], which may affect appearance of AIPs of targets and, in turn, patient setup accuracy. For the purpose of more accurate target alignment, it is reasonable to register on-board CBCT with AIPs from the probability-based sorting instead of AIPs from

conventional single-cycle sorting methods in free-breathing treatment, as this study shows that the probability-based sorting is able to provide more accurate PDFs than single-cycle sorting methods. Results from evaluation of accuracy of the target CoMs in the AIPs suggest that the probability-based sorting may provide AIPs with more accurate and stable target CoMs than the sequential acquisition phase-based sorting method. As there is large variation of the target CoM shifts between studied cases for the cine acquisition phase-based sorting, more inclusive studies are warranted to confirm the advantage of the probability-based sorting over the cine acquisition phase-based sorting. More accurate PDFs derived from the probability-based sorting method may also benefit PDF-based treatment planning and dose calculation and extend the applicability of PDF-based treatment planning for treating moving targets using the free-breathing technique [28, 51, 52]. Currently, dose calculation for free-breathing treatment is performed on AIP images. For a moving target, the dose to AIP is calculated, which can be different from the actual dose to target. Sheng *et. al.* calculated actual doses to target by convoluting the dose distribution with the target PDF and also proposed PDF-based treatment planning based on PDF-based dose calculation for the purpose of reducing dose to tissues surrounding the target. It was found in their study that if a highly reproducible PDF is used in optimization, the gross tumor volume (GTV) dose coverage can be maintained and the mean lung dose reduced by 9.6%. However, they showed that the dosimetric error increased nearly exponentially as the PDF error

increased [28]. Based on their study, as a rough estimate, a 6% increase (from 83% to 89% as demonstrated in this study) in tumor motion PDF reproducibility could lead to a 33% decrease in dosimetric error in GTV coverage.

In this study, the proposed probability-based sorting method was only tested in a 4D-MRI sequential acquisition scenario. It is expected, however, that the new method can be potentially applied for other 4D reconstructions. A direct extension of our current work is to apply the method for 4D-MRI reconstruction based on retrospective k-space sorting. In k-space sorting, the k-space can be divided into many small segments, each of which is associated with a time point and assigned to a respiratory phase. Regardless of the MRI data acquisition scheme (radial, spiral, etc.), the assignment of respiratory phases to the k-space segments is a random process, which meets the criteria for probability-based sorting. In addition, we expect the probability-based sorting method can also be applied to other imaging modalities. One possible application is for 4D-PET reconstruction because of the randomness of decay events relative to breathing variation. PET events detected within a short period of time can be used to reconstruct a volume with low Signal-to-Noise Ratio (SNR), which is associated with a time point on the breathing signal. These volumes will be subsequently resorted into bins for multiple main breathing cycles. Due to the low SNR of each volume, each bin must have a sufficient number of volumes for generating multi-cycle 4D images with acceptable image quality. Despite the potentially broad applicability, it is notable that this method

may not work for 4D-CT due to the fact that 4D-CT is limited to cine acquisition mode. As a result, raw data for each slice only cover a small portion of breathing curves, which often shows similar breathing pattern without much variation. Extending CT acquisition time may not provide sufficient coverage for breathing variation and increased exposure to ionizing radiation is not desirable.

It has been shown in this study that the probability-based reconstruction method, capable of generating 4D images with multiple main breathing patterns, can improve 4D image quality. The improvement can be further enhanced by optimizing the algorithm. For example, in the grouping process, the groups were delimited by amplitude and period of individual breathing cycles as shown in Figure 4(b). This approach, however, may separate individual cycles with close amplitudes and/or periods into two groups. This result is because of grouping individual cycles based on pre-set grids. One possible remedy to this problem is replacing the current grouping method with data-based grouping, or clustering method such as k-means clustering method with optional optimization of the number of groups. The number of groups may be reduced while keeping the PDF similarity as high as current results. Another modification to the current method that can potentially improve the outcome is taking the location of individual breathing cycles into consideration. Baseline shift has been observed in some patient's breathing curves. In the current method, if two individual breathing cycles share the same amplitude and period but have different locations, they will end up in

the same group. For patient with moderate-to-none baseline shift, the reconstructed images will not be affected. In some patients, however, considerable baseline shift has been observed. In this situation, the final motion range may be under-estimated if positions of individual breathing cycles are not taken into consideration. To extend our method to patients with large baseline shifts, each individual breathing cycle will be characterized with three attributes, amplitude, period, and position. The probability-based sorting method identifies main breathing patterns based on analysis of breathing curves captured by external surrogates. In our digital phantom study, the breathing curves were directly used to guide the motion of the phantom so the internal-external breathing motion is perfectly correlated. For patient imaging, although previous studies has pointed out the internal-external breathing motion correlation is usually high, it is still desired to test our method on real patient imaging to evaluate how the method perform with the presence of uncertainty in internal-external breathing motion correlation.

## **4. Development of a Hybrid DIR Framework Using Sparsely-distributed Measured Motion Information**

### ***4.1 Motivation - Utilizing Sparsely-distributed Measured Motion Information***

Current commercial DIR software and open-source DIR algorithms are mostly intensity-based methods, equipped with no or limited functions of feature (landmark) - based improvement. The algorithms allowing for feature-based improvement often rely on a limited number of manually selected image landmarks, which is time-consuming and inconvenient for clinical applications. On the other hand, sophisticated hybrid DIR algorithms capable of handling a large amount of features are currently under development and are not fully ready and commercially available yet for wide clinical applications. In addition, as these hybrid algorithms either first perform feature-based DIR whose results are used as a start point or constraints for the subsequent intensity-based DIR or adopt a coupled approach by performing feature-based and intensity-based DIR simultaneously, the final registration performance may rely on the quality of the feature-based registration. However, this methodology may not be robust and convenient for clinical applications because the amount of image features is variable, which is an uncertain factor to the success of the feature-based registration and, in turn, the hybrid process.

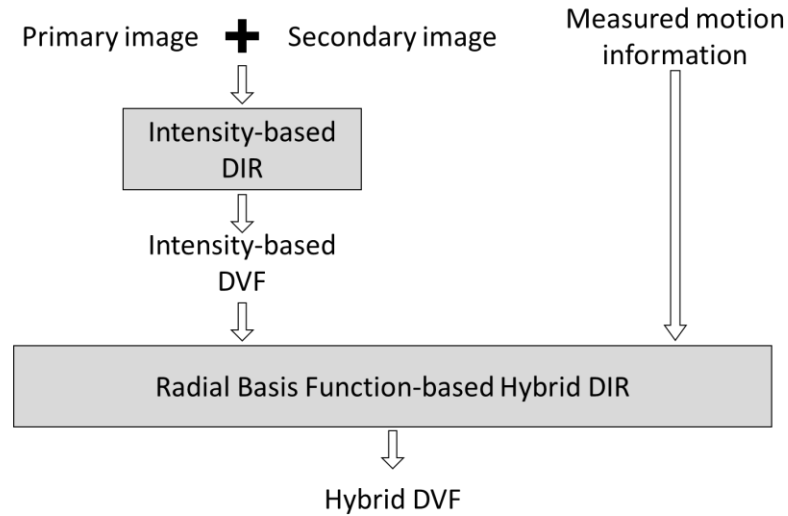
The aim of this work is to develop a robust hybrid DIR framework by incorporating physiologically measured landmark deformation into intensity-based

DVF. The combination of the two DIR methods is effectively achieved by transforming the motion pattern derived by the landmarks to the surrounding intensity-based DVFs via radial basis function (RBF)-based expansion. The feasibility of the framework was tested with multiple current intensity-based DIR algorithms on an in-house developed digital lung phantom with a ground-truth dense DVF.

## ***4.2 An RBF-based Hybrid DIR Framework***

The overall idea of the framework is to adjust the displacement vectors at non-measured points with guidance of both known displacements at measured points and relationship in displacement between the non- and measured points extracted using radial basis function. The framework is carried out in three steps:

(1) A base, intensity-based DVF is converted to an intensity-based coefficient matrix comprising expansion coefficients for the Wendland's RBF; (2) The intensity-based coefficient matrix is modified under the guidance of sparsely distributed measured motion information to generate the hybrid coefficient matrix; (3) The hybrid coefficient matrix is converted to the hybrid DVF.



**Figure 8: The workflow of the hybrid DIR framework**

Figure 8 shows the design of the robust hybrid DIR framework. The overall idea is to take available measured displacements of image features/landmarks to enhance physiological plausibility of DIR and reduce DIR errors in a time-efficient manner for clinical applications. The framework is considered robust because the input data are flexible. The base DVF matrix can be from any DIR algorithm, and the measured displacements can be of any size with any distributions and obtained by any means (manually, automatically, or combined).

The hybrid DIR framework is carried out in three steps.

(1) The intensity-based DVF is reshaped to an  $n$ -by-1 vector, where  $n$  is the number of voxels in the region of interest, labeled as  $f_{pre}$ . It is then converted to an intensity-based expansion coefficient vector ( $n$ -by-1),  $\lambda_{pre}$ , using the 3D Wendland's

compactly supported RBF with second derivative continuity.  $\lambda_{pre}$  is calculated by solving the following equation,

$$A\lambda_{pre} = f_{pre} \quad [1]$$

where  $A = (\phi_{i,j})_{n \times n}$ ,  $n$  is the number of the voxels in the region of interest, and  $\phi_{i,j}$  is defined by the following equation:

$$\phi_{i,j} = \begin{cases} (1 - r_{ij}/r_0)^4(4r_{ij}/r_0 + 1), & 0 \leq r_{ij} \leq r_0 \\ 0 & , r_{ij} > r_0 \end{cases} \quad [2]$$

where  $r_{ij}$  is the Euclidean distance between voxels  $i$  and  $j$ ,  $r_0$  is the range of support for the Wendland's RBF.

(2) The measured displacements of image landmarks are utilized to modify  $\lambda_{pre}$ , to generate a hybrid expansion coefficient map,  $\lambda_{post}$ , such that the following condition is satisfied,

$$A_{truth}\lambda_{post} = f_{truth} \quad [3]$$

where  $A_{truth} = (\phi_{i,j})_{m \times n}$  is a submatrix of  $A$ , consisting of the rows for the voxels where measured displacements are located,  $m$  is the number of the voxels where the measured displacements are available, and  $f_{truth}$  is an  $m$ -by-1 vector consisting of the measured displacements. As Eq. 3 represents an underdetermined linear system, a constraint is introduced to determine the solution  $\lambda_{post}^*$ ,

$$\lambda_{post}^* = arg \min_{\lambda_{post}} \|\lambda_{post} - \lambda_{pre}\|_2^2 \quad [4]$$

(3) The hybrid expansion coefficient map,  $\lambda_{post}$ , is converted back to a hybrid DVF,  $f_{post}$  using the following equation.

$$A\lambda_{post} = f_{post} \quad [5]$$

### ***4.3 Investigation of the Hybrid DIR Framework on Digital Lung Phantom***

#### **4.3.1 Methods and Materials**

##### **4.3.1.1 The Tagging-MR based Thorax Digital Motion Phantom**

The hybrid DIR framework was evaluated using an in-house developed thorax digital phantom which consists of two volumetric images of the lungs at end-of-inhalation (EOI) and end-of-exhalation (EOE) phases, and the ground-truth DVF between the two phases. The phantom DVF is determined using a hyperbolized gas tagging MRI technique which provides an in vivo, direct measurement of regional lung deformations during respiration [57, 58]. The generation of the thorax digital phantom has been described in details previously [59] and will only be briefed here. Subjects were imaged at the EOI and the EOE phases using a high-resolution (2.5mm isovoxel) 3D proton MRI, as well as a hybrid MRI which combines hyperpolarized gas tagging MRI and a low-resolution (4.5 mm isovoxel) proton MRI. A sparse tagging displacement vector field (tDVF) was derived from the hyperpolarized gas tagging MRI by tracking the displacement of tagging grids between EOI and EOE. A motion model of the entire thorax was determined based on the tDVF and the high-resolution MR images, and then

applied to the high-resolution EOI image (primary image in the phantom) to create a deformed EOE image (secondary image in the phantom), forming the virtual phantom where the motion model provides the ground truth of deformation.

#### 4.3.1.2 Testing on an in-house developed lung motion phantom

The hybrid DIR framework was tested with DVFs generated with five DIR algorithms: Velocity, MIM, ILK and OHS algorithms in DIRART toolbox, and Elastix. The registration was performed between the EOI and EOE images of the phantom to generate the corresponding DVFs. The lungs of the phantom were segmented and the DVFs within the lungs were used as the intensity-based DVFs for the study. To expedite the calculation, the intensity-based DVFs were down-sampled by a factor of 6 in each direction. The intensity-based DVFs is one of the two inputs for the hybrid DIR framework,  $f_{pre}$ . The matrix  $A$  was generated for all the voxels retained in the intensity-based DVFs. The measured deformations were selected from the true DVFs of the phantom in a random fashion, i.e., a set of randomly selected voxels (the number of voxels is user determined) within the lungs were designated as landmark voxels with measured deformations. The list of measured deformations of landmark voxels is the second input for the hybrid DIR framework,  $f_{truth}$ . The matrix  $A_{truth}$  consists of  $\phi_{i,j}$  between the landmark voxels and all the voxels retained in the DVFs. Each row of  $A_{truth}$  contains  $\phi$  between the  $i^{th}$  landmark voxel and all the voxels retained in the intensity-based DVFs. Since the intensity-based DVFs and thus the hybrid DVFs were down-

sampled, the hybrid DVFs were restored to full size via linear interpolation for the purpose of comparison with the ground-truth DVF of the digital phantom.

To evaluate the performance of the hybrid DIR framework, intensity-based and hybrid registration errors were calculated and by comparing the intensity-based and hybrid DVFs against the ground-truth DVF of the phantom. The 3D registration error was quantitatively measured at each voxel,  $e_{3D,i}$ , and for the entire volume,  $e_{3D}$ , which were calculated as follows respectively,

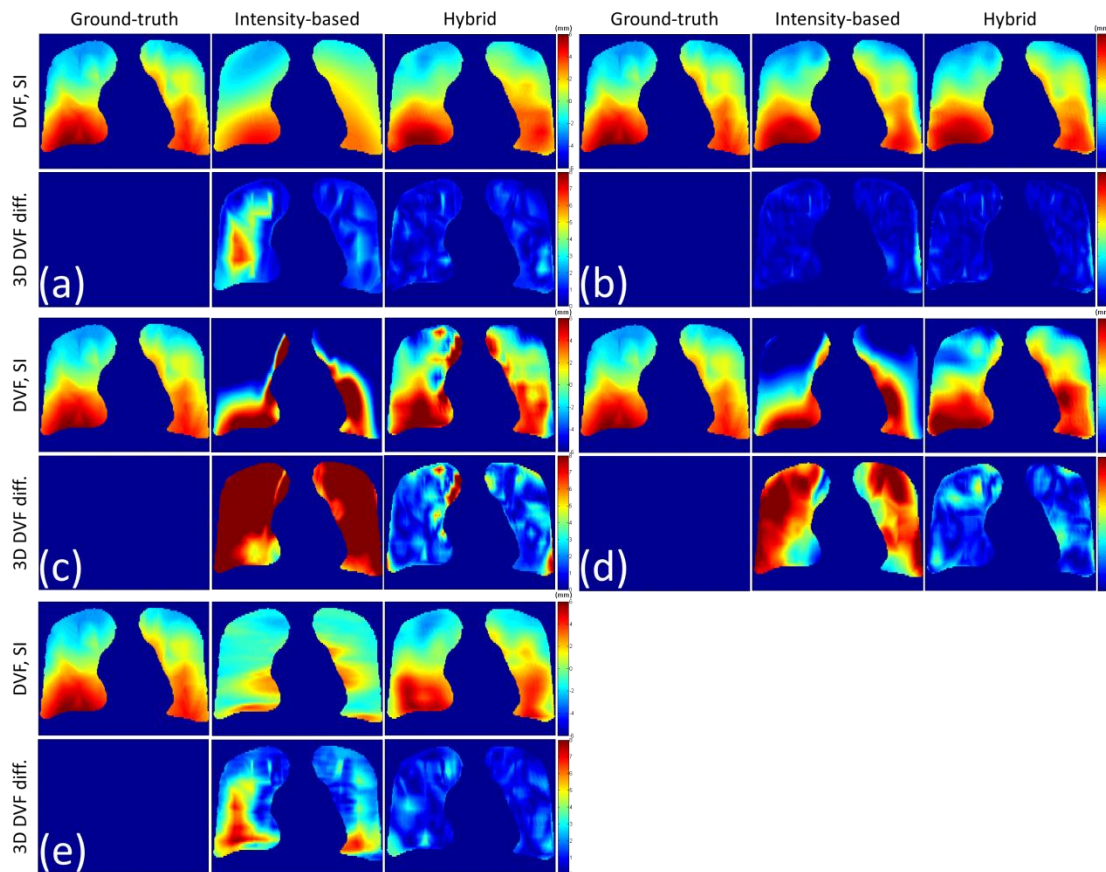
$$e_{3D,i} = \|\vec{f}_i - \overrightarrow{f_{gt\ map,i}}\|_2 \quad [6]$$

$$e_{3D} = \frac{\sum_{i=1}^n e_{3D,i}}{n} \quad [7]$$

where  $\vec{f}_i$  represents either the intensity-based or hybrid displacement vector at the  $i$ th voxel,  $\overrightarrow{f_{gt\ map,i}}$  is the measured displacement at the  $i^{th}$  voxel, and  $n$  is the number of voxels. In addition to the 3D registration errors, the performance of the hybrid DIR framework was also evaluated using the intensity difference map between the original phantom EOI image and EOI images synthesized by applying hybrid DVFs to the phantom EOE image. Furthermore, considering in practice the number and the distribution of available landmarks vary greatly depending upon the application site, we investigated the impact of these two factors on the effectiveness of the DIR improvement method (indicated by the reduction of 3D registration error). Particularly for landmark distribution, we evaluated several scenarios when different percentage of landmarks is located on the grids of the down-sampled intensity-based DVFs.

### 4.3.2 Results

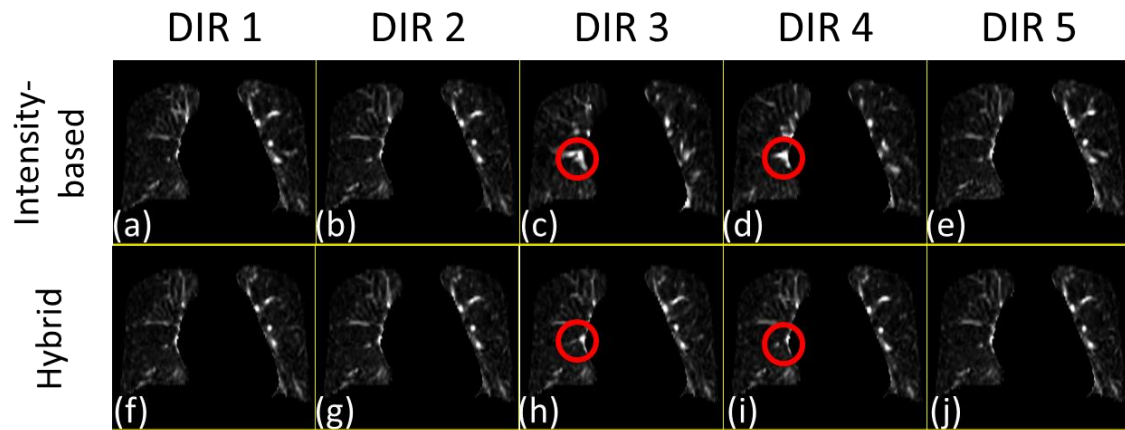
The 3D registration errors, defined as the vector difference between the intensity-based or hybrid displacement vector and the ground truth displacement vector, have been found to decrease after correction. Intensity-based and hybrid 3D registration errors per voxel for Velocity, MIM, ILK, OHS, and Elastix are 2.43 mm vs 1.24 mm, 0.81 mm vs 0.87 mm, 8.20 mm vs 2.17 mm, 4.74 mm vs 1.53 mm, 2.87 mm vs 1.30 mm. Figure 9 shows the reduced displacement vector errors in the superior-inferior (SI) direction after hybrid DIR for an example case. The intensity errors, defined as the intensity difference between the primary volume synthesized with intensity-based or hybrid DVF and the original primary volume, were also reduced by 0.37 in unit of image intensity on average after correction. Intensity-based and hybrid intensity errors per voxel for Velocity, MIM, ILK, OHS, and Elastix are 2.43 vs 2.12, 1.42 vs 1.78, 3.65 vs 2.88, 3.22 vs 2.46, 2.87 vs 2.52 in unit of image intensity. Figure 10 shows the reduced intensity errors for primary volumes synthesized with hybrid DVFs.



**Figure 9: Intensity-based and hybrid DVFs in the SI direction and 3D registration error maps are shown for the five DVFs generated with (a) Velocity, (b) MIM, (c) ILK in DIRART toolbox, (d) OHS in DIRART toolbox, and (e) Elastix. In each sub-figure, the ground-truth DVF in the SI direction is shown at the top left, intensity-based DVF in the SI direction shown at the top middle, hybrid DVF in the SI direction shown at the top right, intensity-based 3D registration error maps shown at the bottom middle, hybrid 3D registration error map shown at the bottom right.**

Figure 9 shows the results for applying the hybrid DIR framework to the intensity-based DVFs generated with the five intensity-based DIR algorithms. The most prominent reduction in the registration errors is observed in the right lung. In Figure 11(a), (d) and (e), the large area within which intensity-based errors are greater than 5 mm disappears in the hybrid error map. Since the intensity-based errors in the left lung

and are small compared with the errors in the right lung, the reduction of error is not obvious in the registration error maps. However, when comparing the error distribution in the intensity-based and hybrid DVF with that in the ground-truth DVF, it is clear that the error distribution in the hybrid DVFs, e.g., the yellow region with a sinuous shape, is more similar to the ground-truth error distribution than the over-smoothed error distribution in the intensity-based DVFs. The quantitative results shows that reduced registration errors are found in all tested DVFs. Intensity-based and hybrid 3D errors per voxel for Velocity, MIM, ILK, OHS, and Elastix are 1.27mm vs 1.07mm, 1.56mm vs 1.55mm, 6.33mm vs 4.81mm, 3.53mm vs 2.94mm, 2mm vs 1.67mm.

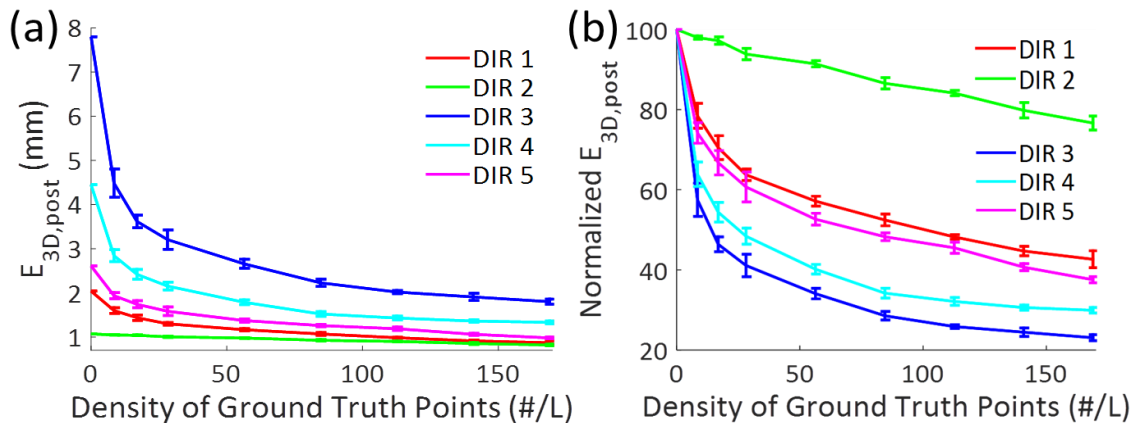


**Figure 10:** The top row (a) – (e) shows the difference maps between the primary image and the image synthesized by deforming the secondary image with the intensity-based DVFs. The bottom row (f) – (j) shows the difference maps between the primary image and the image synthesized by deforming the secondary image with the hybrid DVFs. DIR 1 to DIR 5 correspond to Velocity, MIM, ILK, OHS, and Elastix, respectively.

Figure 10 shows the results for the differences, or the intensity error maps, between the primary image and the image synthesized by deforming the secondary image with the intensity-based or hybrid DVFs. The improvement in the accuracy of intensity is not as visually obvious as that in the accuracy of displacement vectors. However, in certain locations with a large area of intensity errors, such as the red circles in the intensity-based images (Figure 10(c) and (d)) and the hybrid images (Figure 10(h) and (i)), the reduction of intensity errors can be clearly seen. Quantitatively, intensity-based and hybrid intensity errors are reduced by 0.37 in unit of image intensity on average.

Figure 11 shows the impacts of the number of measured displacements and the intensity-based registration errors on the hybrid registration errors. Overall, the hybrid registration errors decrease as the number of measured displacements increases for all the DVFs tested. As can be seen in Figure 11(a), for DVFs with average intensity-based 3D registration errors up to 8mm, the hybrid 3D registration errors can reach smaller than 2mm, which is comparable with the typical slice thickness of CT images. Meanwhile, the marginal reductions of the hybrid registration errors diminish. It is also observed that the order of magnitude of the hybrid errors for the five DVFs remains the same as that of the intensity-based errors, revealing a correlation between intensity-based and hybrid registration accuracy. In contrast, it is observed in Figure 11(b), the

ratio of hybrid to intensity-based errors is inversely correlated with the intensity-based error, indicating larger improvement for DVFs with poorer accuracy.



**Figure 11: The left figure (a) illustrates the impact of the number of measured displacements on the hybrid 3D registration errors. The right figure (b) illustrates the impact of the number of measured displacements on the percentage hybrid 3D registration errors normalized to the intensity-based 3D registration errors. DIR 1 to DIR 5 correspond to Velocity, MIM, ILK, OHS, and Elastix, respectively.**

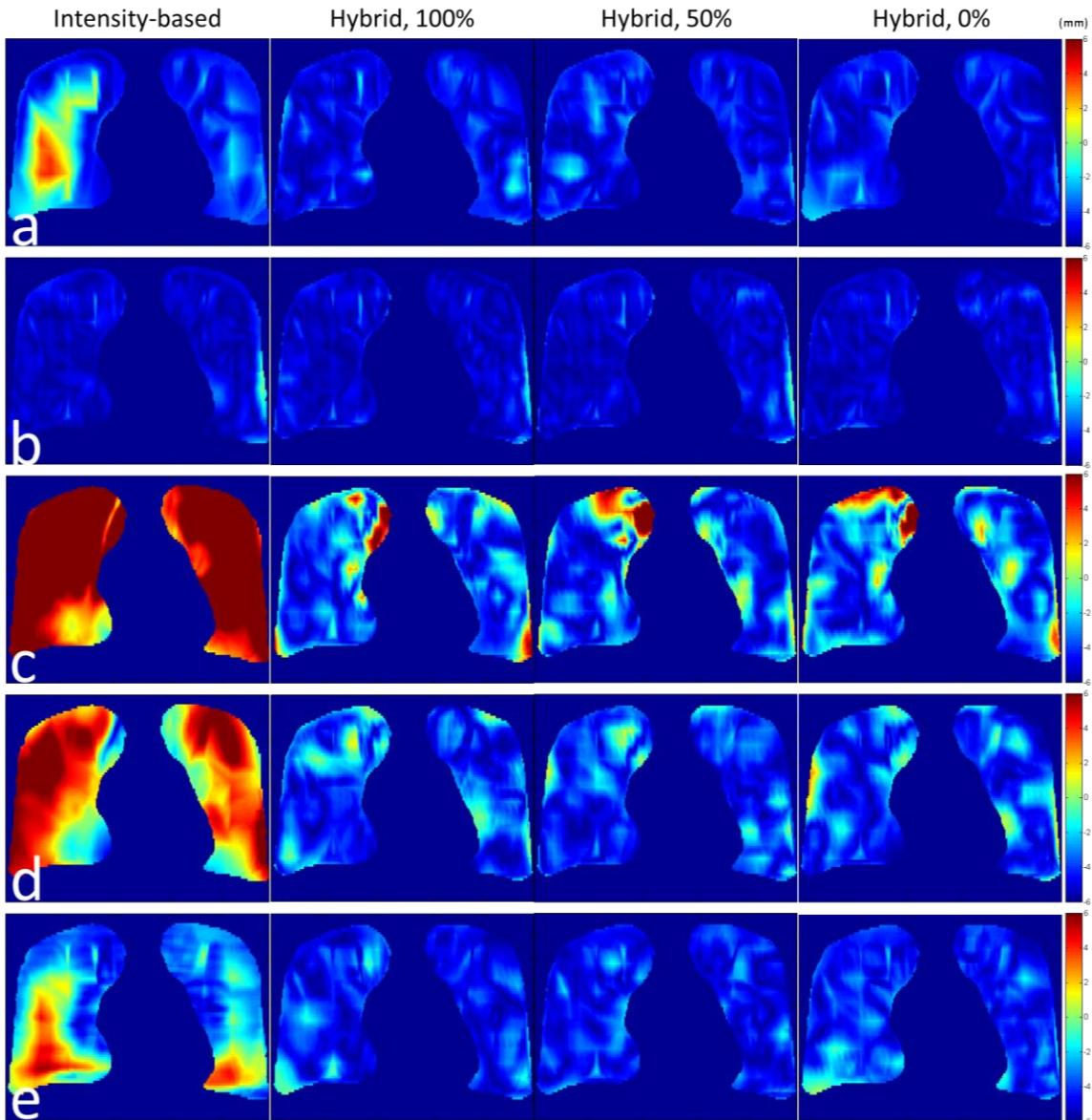


Figure 12: Impact of distribution of measured points on the performance of the hybrid DIR framework for the five DVFs generated with (a) Velocity, (b) MIM, (c) ILK in DIRART toolbox, (d) OHS in DIRART toolbox, and (e) Elastix. In each row, the figures from left to right show the intensity-based 3D registration error map, the hybrid 3D registration error map for a minimum on-grid ratio of 100%, the hybrid 3D registration error map for a minimum on-grid ratio of 50%, and the hybrid 3D registration error map for a minimum on-grid ratio of 0%.

Figure 12 shows the performance of the hybrid DIR framework for the five DVFs using different minimum percentage of the measured voxels belonging to the down-sampled voxels (on-grid ratio). For minimum on-grid ratios ranging from 100% (Figure 12(b)) to 0% (Figure 12(d)), the overall reduction of registration errors is comparable. Hybrid 3D registration errors for DVF generated with Velocity: 100%,  $1.07 \pm 0.04$  mm, 50%,  $1.06 \pm 0.05$  mm, 0%,  $1.06 \pm 0.03$  mm; MIM: 100%,  $0.90 \pm 0.01$  mm, 50%,  $0.90 \pm 0.02$  mm, 0%,  $0.90 \pm 0.03$  mm; ILK: 100%,  $2.14 \pm 0.15$  mm, 50%,  $2.14 \pm 0.09$  mm, 0%,  $2.05 \pm 0.05$  mm; OHS: 100%,  $1.45 \pm 0.03$  mm, 50%,  $1.55 \pm 0.06$  mm, 0%,  $1.56 \pm 0.10$  mm; Elastix: 100%,  $1.16 \pm 0.02$  mm, 50%,  $1.21 \pm 0.05$  mm, 0%  $1.25 \pm 0.03$  mm.

#### **4.4 Discussion**

It is noteworthy that the results shown in Figure 11 suggest that it is feasible to reduce DIR errors by utilizing a relatively small number of measured displacement vectors. Regarding application of the proposed method in clinical practice, although it is tempting to use more measured points for better DIR accuracy, it is not always efficient to do so because the extra effort to identify more measured points may not be compensated as the marginal reduction of registration error decreases. Also, it may not be necessary to use as many measured points as possible because when the registration error is reduced to below image resolution (typically around 1 mm), the registration error will no longer be a limiting factor of registration accuracy. Based on these

observations, an effective way to obtain a DVF with good quality may be applying the proposed hybrid DIR framework on a DVF with small registration errors.

The results demonstrate the efficacy of this hybrid DIR framework in terms of DVF accuracy and image accuracy. However, some residual errors can be observed in the peripheral regions close to the boundary of the lungs, such as a region in the upper right lungs facing the left lung shown in Figure 9(c). This kind of uncertainty in the hybrid DVFs may be caused by instability of the RBF interpolation near the boundaries. As most boundary voxels are located outside the grid points, the voxels near the boundaries are extrapolated. Also, the distribution of the measured displacements may have an impact on the DIR errors in the hybrid DVFs. For each algorithm, measured points were randomly sampled 5 times to evaluate the impact of distribution of measured points. The maximum coefficient of variation of hybrid mean 3D registration error is 7.0% for ILK DVFs. CV for other DVFs are 3.3% for Velocity, 1.2% for MIM, 2.1% for OHS, and 1.7% for Elastix. Although the measured points are randomly selected, it cannot be guaranteed that the distribution is uniform. Since a higher on-grid ratio tends to distribute the measured displacements more uniformly, smaller DIR errors in the hybrid DVFs may be correlated with a more uniform distribution of ground-truth displacements. If confirmed by further investigation, extraction of measured displacements may be more efficient by using a uniform distribution rather than a distribution with varied densities of measured displacements. Despite these limitations,

it has been shown that the proposed framework is capable of considerably reducing DIR errors for DVFs with larger-than-voxel size DIR errors.

The hybrid DIR framework has shown potential to reduce DIR error with flexibility. However, the limitation of this study is that the framework has only been tested on one digital phantom due to the limited availability to ground-truth DVFs, a difficulty that often baffles DIR studies. One possible way to further validate our method on patient images can be carried out in three steps: first, identify certain amount of landmarks in both the primary and the secondary images; second, using part of the ground truth points for correction; third, the remaining ground truth points are used for evaluation of the post-correction errors.

Wendland's compactly supported function is chosen for the current framework to enable correlation of motion at one point with points nearby while limiting the maximum correlation range. Rationale of this model is that in a relatively homogenous tissue, motion is expected to be continuous. It is also reasonably expected that motion at one point should have minimal impact on the motion at points far away. This selection of RBF also provides a potential to further improving the calculation speed. As the RBF values between points further than a certain distance are all zeros, matrix  $A$  are sparse, leaving potential for improvement on storage and calculation efficiency.

In practical application, consideration should be given to proximity of the measured points. It is not uncommon to have two or more available measured points

close to each other. In theory, as the motion are smooth, having more than one measured points are not likely to provide much additional information for DIR improvement.

From a practical point of view, two measured points in proximity will have similar rows or columns in matrix  $A$ , leading to numerically unstable matrix when inverted.

Therefore, imposing a limit on minimum spacing between the may be beneficial. Further studies on the impact of distribution of measured points on DIR improvement are warranted.

In conclusion, a hybrid DIR framework capable of reducing registration errors by utilizing sparsely available measured motion information has been developed and evaluated. It has been shown that DIR errors could be considerably reduced with less than 0.5% of total number of voxels.

## **5. Exploration of New Methods for Physiological Feature Extraction**

### ***5.1. Motivation – Automation of Bifurcation Extraction***

In chapter 4, a flexible hybrid DIR framework utilizing sparsely available measured motion information has been presented. Based on the evaluation results, the framework is expected to significantly reduce DIR errors for most part of tissue. However, the uncertainty from DIR may still exist in the hybrid DVFs. It is possible to eliminate this uncertainty from DIR algorithms by solely using physiological motion information provided sufficient number of landmarks can be identified. This process can be difficult for some tissues. However, we have noticed that vessel trees in the lung images may be able to provide sufficient landmarks if bifurcations and vessels can be extracted and matched. Based on this observation, we intend to first explore the feasibility of automatic extraction of physiological landmarks using a random walking method and then match the bifurcation trees in different phases based on parameters associated with the bifurcations.

### ***5.2. Automatic Extraction of Vessel Bifurcations Using a Random Walking Method***

#### **5.2.1. Methods and Materials**

##### **5.2.1.1 Generation of a Bifurcating Tree for Vessel Tree Simulation**

At this stage, it is assumed that the vessel tree can be represented by a bifurcating tree, i.e. each and every branching of a vessel tree is bifurcating. Also, since

the simulated vessel tree is intended for testing a new branching point identification algorithm applied to skeletonized vessel trees, the width of branches is not relevant in the simulated vessel tree. Therefore, in this section, methods for generating a bifurcating tree with skeleton only are presented.

*Determination of the number of bifurcating levels*

For a bifurcating tree, the maximum number of branching points,  $N$ , is calculated by Eq. 1

$$N = 2^{l+1} - 1 \quad [1]$$

where  $l$  is the number of bifurcating levels of the tree, or the depth of the tree.

Therefore, the maximum number of bifurcations is determined by the number of levels. To best satisfy the purpose, the number of level is chosen as 7, such that there will be 255 bifurcations in the simulated tree.

*Generation of a bifurcating tree*

Theoretically, since the algorithm to be tested is not affected by the geometry of the vessel, the orientation of the branches can be unconstrained. However, an unconstrained tree may appear to be very different from real trees, such as vessel trees and botanical trees, and thus be difficult for visual evaluation. In previous studies, it is suggested that the branches and their arrangement in botanical trees can be easily captured by the eyes [60]. To make the simulated vessel tree easy for visual evaluation, the bifurcating tree was generated based on a simplified Holton's model that has been

designed for automatically generating botanical trees and forests for animation purposes [60, 61]. By using this model, the left and right child branches of a node are generated in pair with their orientation controlled by a set of parameters.

The bifurcating tree was generated recursively, starting from the root node, following the work flow shown in Figure 13. In each recursive step, child branches were generated based on the simplified Holton's model and then attached to the current eligible node (eligibility check shown in Figure 13). This process is repeated until the exit criterion was met. The process of child branches generation is elaborated in the next subsection.

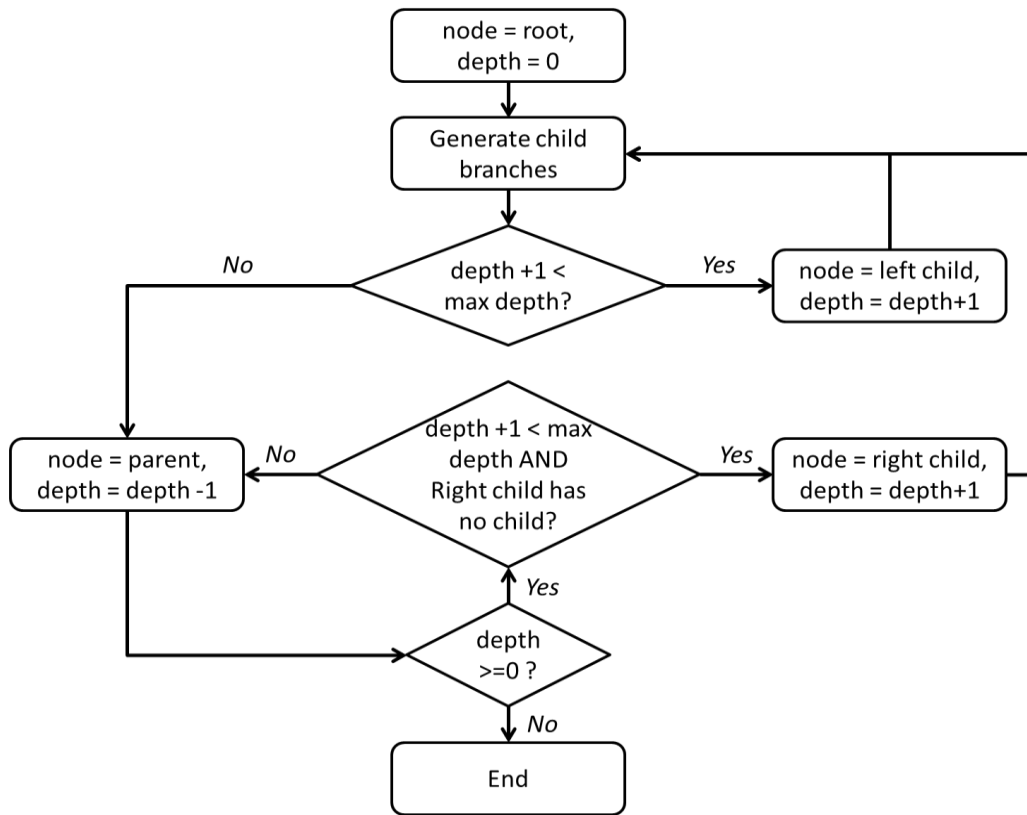


Figure 13: Work flow for generation of a bifurcating tree. **node**: current node in the tree; **root**: the root node, the first node of the tree; **depth**: depth of the current node, depth of the root node set as 0; **max depth**: the number of bifurcating levels of the tree, or the depth of the tree; **left child**: left child node of the current node; **right child**: right child node of the current node; **parent**: parent node of the current node.

*Generation of bifurcations (attach child branches to parent branches)*

In the Holton's model, strands, an analogy of vascular structure in botanical trees, are associated with each branch and used to control various properties of the tree such as branch thickness and length, as well as branching angles between a branch and a sub-branch. In the original Holton's model, the number of strands of the parent branch,  $s$ , is shared by the left and the right branches,  $s_1$ , and  $s_2$ , where  $s = s_1 + s_2$ . The allocation of strands between the two child branches can be random to render a more realistic appearance of the tree. In our simplified version,  $s_1$  and  $s_2$  are constant, as determined by Eq. 2 and 3.

There are two parameters that control the geometry of the simulated bifurcating tree,  $\alpha$ , the angle between the two child branches, and  $\beta$ , the angle at which a bifurcation rotates around its parent branch.

$$s_1 = ps = \frac{1}{2}s \quad [2]$$

$$s_2 = (1 - p)s = \frac{1}{2}s \quad [3]$$

$$\lambda_1 = \lambda \sqrt{\frac{s_1}{s}} = \frac{\lambda}{\sqrt{2}} \quad [4]$$

$$\lambda_2 = \lambda \sqrt{\frac{s_2}{s}} = \frac{\lambda}{\sqrt{2}} \quad [5]$$

$$\alpha_1 = \alpha \frac{s_2}{s} = \frac{\alpha}{2} \quad [6]$$

$$\alpha_2 = \alpha \frac{s_1}{s} = \frac{\alpha}{2} \quad [7]$$

$$\beta = \beta_0 l \quad [8]$$

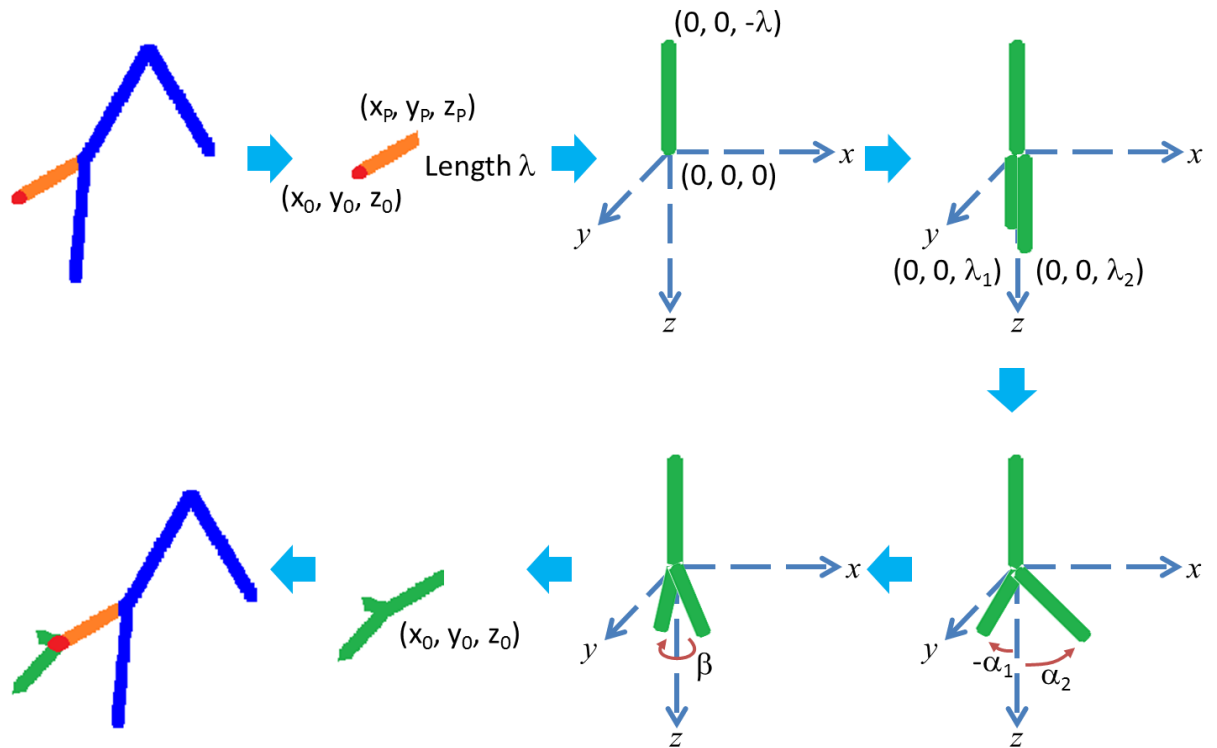


Figure 14: This flow chart shows how a bifurcation is generated, i.e., how two child branches are attached to the current node of interest (red dot) in one recursive step. The existing bifurcating tree is in blue with the parent branch of the current node of interest in orange.  $(x_0, y_0, z_0)$ : the coordinates of the current node of interest.  $(x_P, y_P, z_P)$ : the coordinates of the other end of the parent branch.  $\lambda$ : length of the parent branch.  $\lambda_1$ : length of the left child branch (green).  $\lambda_2$ : length of the right child branch (green).  $\alpha_1$ : magnitude of the angle at which the left child branch rotates about the y axis.  $\alpha_2$ : magnitude of the angle at which the right child branch rotates about the y axis.  $\beta$ : magnitude of the angle at which the entire bifurcation (green) rotates about the z axis.

Generation of a bifurcation, i.e. generating two child branches (green) and attaching them to the current node (red dot), in a recursive step is illustrated in Figure 14. Each newly generated bifurcation is centered at the origin of the coordinate system with the parent branch aligned with the negative z direction. Two child branches were then generated, whose length  $\lambda_1$  and  $\lambda_2$  are calculated by Eq. 2. In the following step, the left and the right branches are rotated about the y axis at angles  $\alpha_1$  and  $\alpha_2$ , calculated by Eq. 3. Next, the entire bifurcation structure is rotated about the z axis at angle  $\beta$ . At last, the entire structure is rotated and translated to the final position by aligning the parent branch of the newly generated bifurcation (green) with the parent branch in the existing bifurcating tree (orange).

To explore how the different parameters  $(\alpha, \beta_0)$  may affect the appearance of the bifurcation tree, four sets of parameters were tested as listed below.

(a)  $\alpha=60^\circ$ ,  $\beta_0=360^\circ/\psi$ , (recommended in previous studies[62, 63]);

(b)  $\alpha=60^\circ+(l-1)*10^\circ$ ,  $\beta_0=360^\circ/\psi$ ;

(c)  $\alpha=30^\circ$ ,  $\beta_0=360^\circ/\psi$ ;

(d)  $\alpha=60^\circ$ ,  $\beta_0=180^\circ/\psi$ , where  $\psi = \frac{1+\sqrt{5}}{2}$

### 5.2.1.2 Vessel Tree Extraction on Patient Images

In Figure 15(a), a workflow of deriving vessel bifurcations is proposed, comprising four operations: segmentation, binarization, skeletonization, and extraction of bifurcations. Segmentation and skeletonization have standard methods and will only

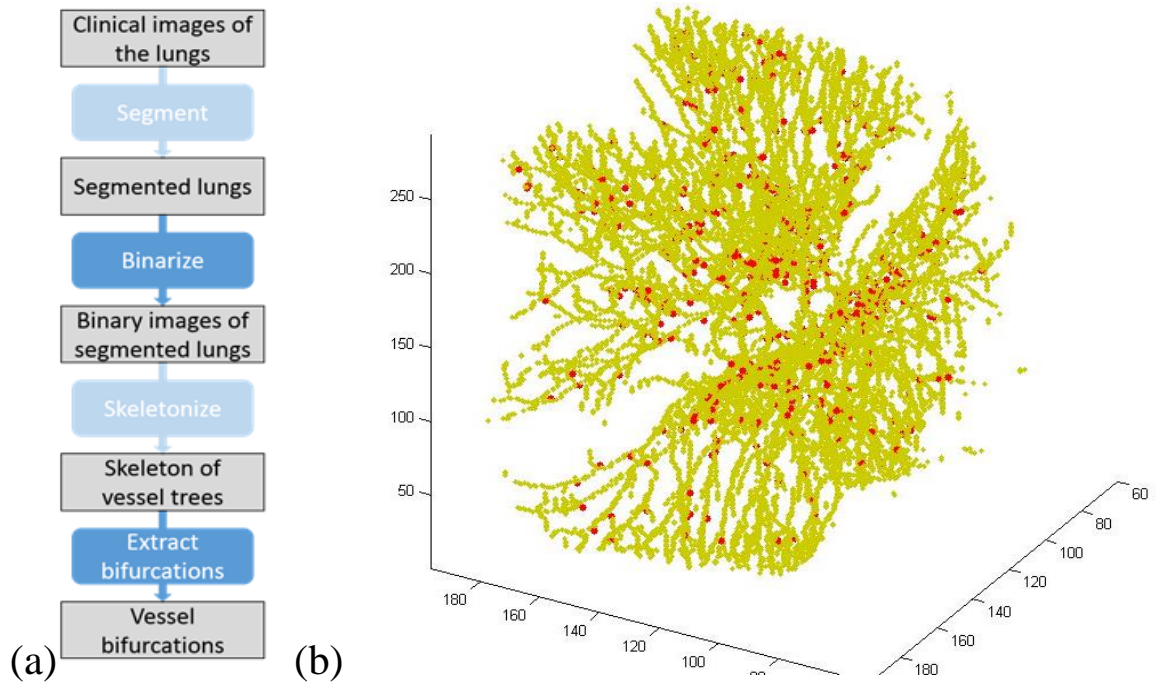
be briefly introduced. Although many binarization methods have been developed, most methods use a single threshold for the entire image. These methods may not provide ideal binary images of vessel trees because of failure to handle intensity heterogeneity commonly seen in clinical images, resulting in missing parts of vessel trees. Therefore, this study aims to improve the binarization method and develop a random walking method for bifurcation extraction. Figure 15(b) shows an example of the extract lung vessel tree with identified bifurcation landmarks from a high quality MR image.

#### *Segmentation*

Segmentation of the lungs can be performed manually or automatically in the Eclipse treatment planning system.

#### *Binarization*

The conversion is performed in ImageJ using the “Make Binary” command. A threshold is automatically determined by analyzing histogram of the image and applied to the entire image for the conversion.



**Figure 15: (a) The workflow of deriving vessel bifurcations from clinical images. Operations of interest are highlighted. (b) Example of extracted lung vessels trees (yellow) with bifurcation landmarks (red) based on a high quality MR image of the lungs (isovoxel 1.75 mm using radio TrueFISP sequence).**

### *Skeletonization*

As we are focused on bifurcation structure and not interested in the thickness of vessels, this operation is to reduce a binary object to its central line. Skeletonization will be performed in ImageJ using the “Skeletonize” command based on thinning algorithms developed by Lee, et al[64].

### 5.2.1.3 Extraction of Bifurcations Using Random Walking

Assuming the image quality is sufficiently good such that there is no disconnection in the vessel tree, a tracker that can move along the skeletonized vessel

tree is possible to reach any voxel on the tree. Before bifurcation detection, end voxels, defined as voxels with only one neighbor, were identified. The tracker would start from end voxels and travel along the vessel skeleton. Once the tracker reaches a bifurcation, the tracker will randomly choose the left or right path. In practice, the tracer will move to a randomly selected neighboring voxel that has not been visited in the current trip. With sufficient repetitions of this process, most of the bifurcation nodes are expected to be visited. A voxel is identified as a bifurcation node if its count of visits is different from the tree voxels around. After  $m$  repetitions, the probability of a bifurcation node at level  $n$  being visited for at least once,  $p$ , can be written as,

$$p = 1 - \left(1 - \frac{1}{2^n}\right)^m \quad [9]$$

The repetition times,  $m$ , can be determined by setting  $p > 95\%$  when  $n$  equals the height of the binary tree, shown as below,

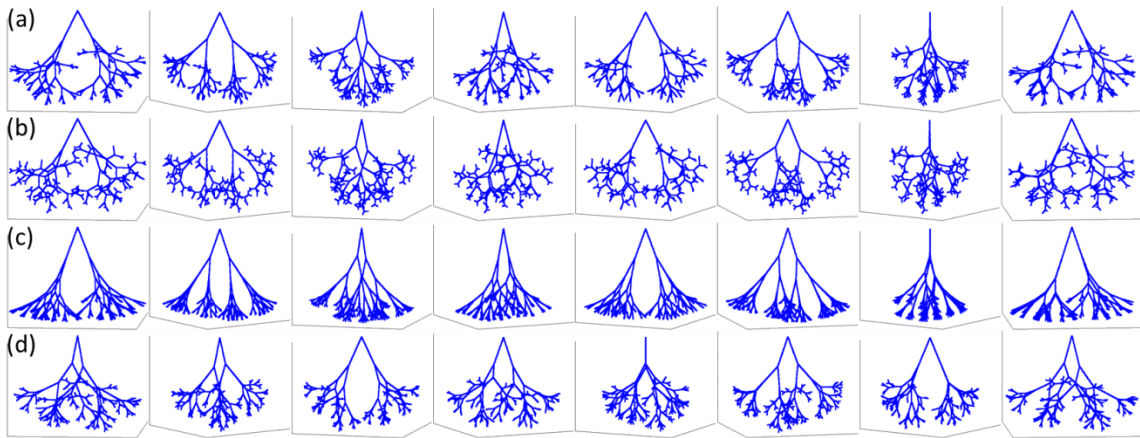
$$m > \frac{-2 + \log 5}{\log\left(1 - \frac{1}{2^n}\right)} \quad [10]$$

The feasibility of the auto-extraction method will be tested on a 7-layer artificial bifurcation structure with 254 ground-truth bifurcations. The algorithm was run different numbers of iterations ranging from 1 to 100 for evaluation of the ability to detect bifurcations. In each iteration, one trip started from each end voxel and ended at another end voxel. After a pre-set number of iterations were met, times each voxel had been visited, or the passing time, were examined for each voxel. A voxel was labeled as a bifurcation if its passing time was greater than all of its neighbors.

To evaluate the accuracy of the bifurcation detection algorithm, the nearest ground-truth bifurcation was identified for each detected bifurcation. Percentages of detected bifurcations within 1 voxel, 1.5 voxels, 2 voxels and 3 voxels from the nearest ground-truth bifurcation were calculated as metrics for bifurcation detection accuracy.

## 5.2.2 Results

### 5.2.2.1 Simulated Vessel Trees

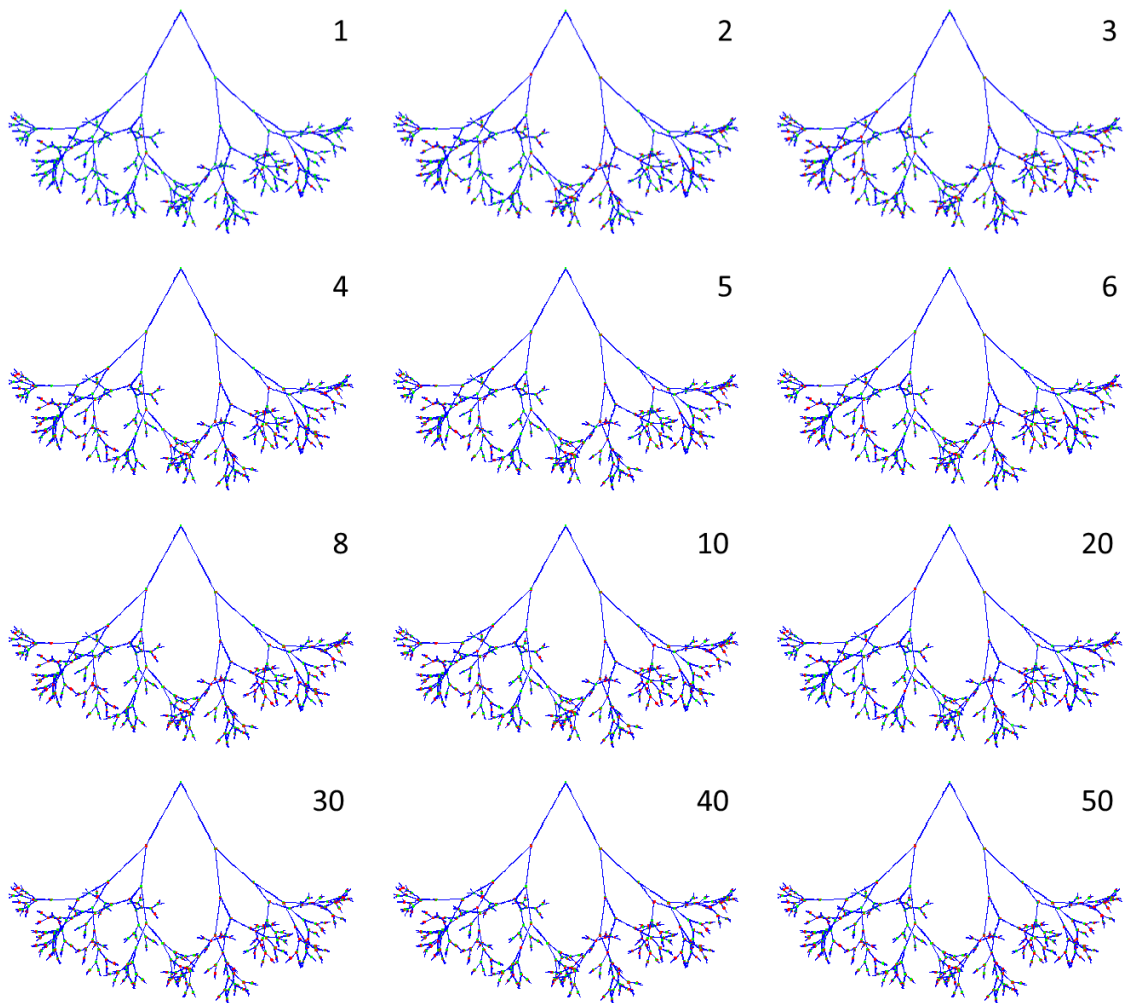


**Figure 16: Snapshots of the vessel trees generated with different sets of parameters, (a)  $\alpha=60^\circ$ ,  $\beta=360^\circ/\psi*\text{level}$ , (b)  $\alpha=60^\circ+(\text{level}-1)*10^\circ$ ,  $\beta=360^\circ/\psi*\text{level}$ , (c)  $\alpha=30^\circ$ ,  $\beta=360^\circ/\psi*\text{level}$ , (d)  $\alpha=60^\circ$ ,  $\beta=180^\circ/\psi*\text{level}$ ,  $\psi=(1+\text{sqrt}(5))/2$ . In each row, snapshots of a simulated vessel tree taken at multiple angles are shown.**

In Figure 16, four bifurcating trees, or simulated vessel trees generated with different sets of parameters were shown. The bifurcating tree (b) differs from the bifurcating tree (a) in that the  $\alpha$  angle is not constant but increasing as the level goes up. The increasing  $\alpha$  angle gives the high level branches a more separated spatial distribution while does not significantly affect low level branches. As a result, the two trees look similar overall. The bifurcating tree (c) was generated with a different  $\alpha$  angle

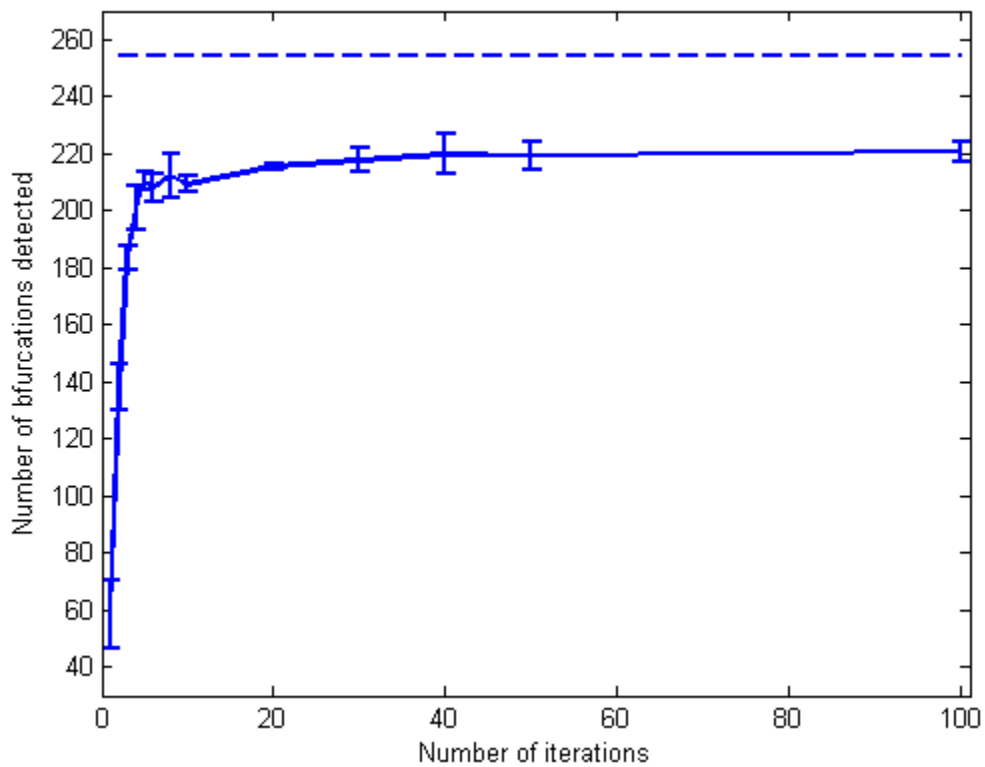
( $30^\circ$  vs  $60^\circ$ ). The branches in bifurcation tree (c) are pointing more forward. The spatial distribution of tree (c) is more constrained, which is expected as the  $\alpha$  angle determines how much a child branch deviates from its parent branch. Tree (d) was generated with a different  $\beta$  angle from tree (a). Visually, although the shapes of the two trees are different, the overall separation of the branches is similar between the two trees.

### 5.2.2.2 Bifurcations Extracted from the Simulated Vessel Tree



**Figure 17: Detected (red dots) and ground-truth bifurcations (green dots) superimposed on the simulated vessel tree**

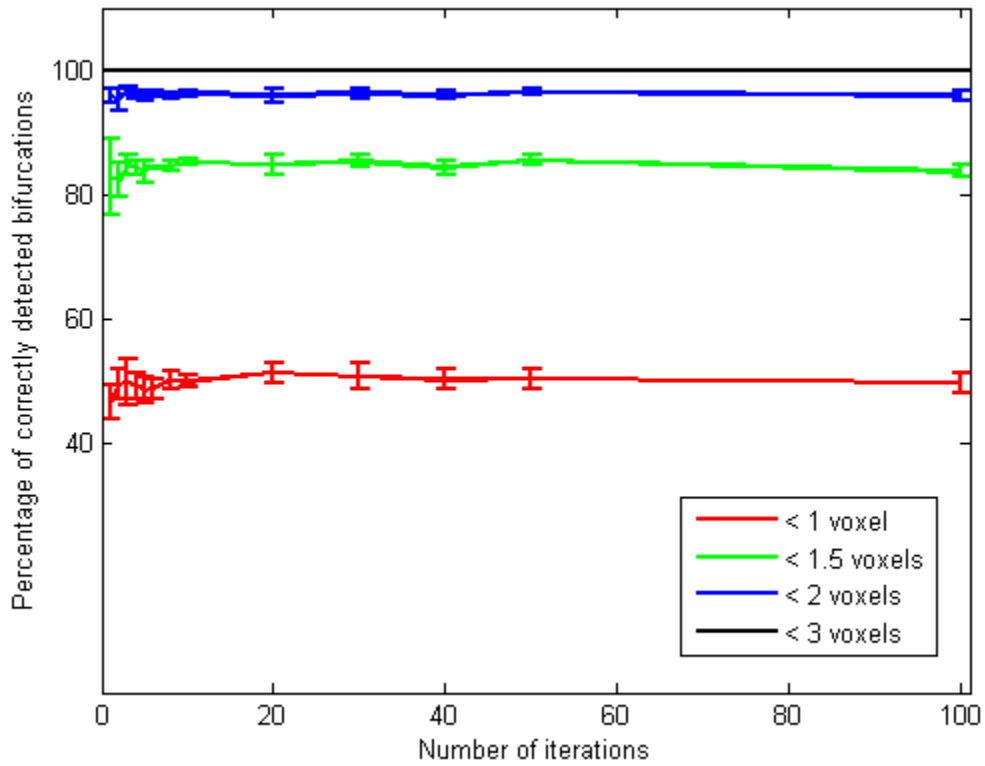
In Figure 17, bifurcations detected (red dots) are superimposed on the simulated vessel tree with ground-truth bifurcations marked with green dots. The number of repetitions is labeled at the upper right corner of each figure.



**Figure 18: Number of bifurcations detected vs number of iterations**

In Figure 18, the numbers of bifurcation detected at different numbers of iterations, ranging from 1 to 100, are shown. The number of bifurcations detected approach stable at around 220 out of the 254 ground-truth bifurcations after more than

20 iterations. In this simulation case, the algorithm is able to detect more than 90% of the maximum detectable bifurcations in 4 iterations.



**Figure 19: Percentage of detected bifurcations within (a) 1 voxel (red), (b) 1.5 voxels (green), (c) 2 voxels (blue), and (d) 3 voxels (black) of the nearest ground-truth bifurcation.**

Accuracy analysis of the detected bifurcations is shown in Figure 19. The percentages of the detected bifurcations within a certain distance to the nearest ground-truth bifurcation are generally stable for different number of iterations. However, variation of the percentage tends to increase toward fewer numbers of iterations.

Overall, approximately 50% of the detected bifurcation falls within 1 voxel around a

ground-truth voxel. Nearly 85% of the detected bifurcations are located within 1.5 voxels from a ground-truth bifurcation. 95% of the detected bifurcations can be found within 2 voxels from a ground-truth bifurcation. None of the detected bifurcations fall more than 3 voxels away from a nearest ground-truth bifurcation.

## 6. Conclusion

In this dissertation, we developed a number of novel methods on the improvement in 4D image quality and feature-based deformable image registration, creating a more solid ground toward physiological plausible respiratory motion modeling. In particular, we have developed a probability-based multi-cycle sorting method for improved 4D image reconstruction and demonstrated the feasibility of a novel probability-based multi-cycle 4D image reconstruction method. It has been shown that the new method can improve the image quality of 4D images, the accuracy of tumor motion PDF and the AIP of 4D images, presenting potential advantages over the conventional phase-based reconstruction method for radiation therapy motion management. We have also developed a hybrid DIR framework which is capable of potentially reducing registration errors by utilizing sparsely available measured motion information. Simulation on a thorax digital motion phantom showed that DIR errors could be considerably reduced with less than 0.5% of total number of voxels. Due to its high flexibility, this hybrid DIR framework can be used to enhance the DIR accuracy in radiotherapy research with a variety of currently available data and tools. Furthermore, we have developed an automatic method to extract vessel bifurcations based on the random walking strategy and demonstrated its high accuracy and efficiency in an evaluation study on a simulated 7-layer bifurcating tree. The method has been shown to detect more than 85% of the actual bifurcations with high accuracy (95% of the detection

errors are within 2 voxels). Vessel bifurcation landmarks automatically detected by this method may further expand the applicability of the DIR improvement method.

Further studies following the probability-based multi-cycle sorting method aim to extend the method for k-space re-ordering and possibly for 4D-PET reconstruction. A study on k-space ordering using the probability-based multi-cycle sorting method has been conducted [65]. Further improvement may be achieved by combining the novel sorting method with accelerated acquisition techniques. The RBF-based hybrid DIR framework is warranted to be evaluated on more scenarios. Despite the lack of ground-truth for DIR evaluation, lung CT data with 300 manually detected landmarks publicly available online ([www.dirlab.com](http://www.dirlab.com)) can be used for point check for the DIR improvement framework. In the current stage, automatic vessel bifurcation extraction is mainly used for providing landmarks for the DIR improvement method. The random walking method can potentially provide more information than just the location of vessel bifurcations, such as the connectivity of vessel bifurcations, length of vessel branches, which may, in combination with the location of bifurcations, be used for parametric analysis-based DIR.

## References

1. Cai, J., et al., *Four-dimensional magnetic resonance imaging (4D-MRI) using image-based respiratory surrogate: a feasibility study*. *Med Phys*, 2011. **38**(12): p. 6384-94.
2. Tian, Y., et al., *Experimental investigation of irregular motion impact on 4D PET-based particle therapy monitoring*. *Phys Med Biol*, 2016. **61**(2): p. N20-34.
3. von Siebenthal, M., et al., *4D MR imaging of respiratory organ motion and its variability*. *Phys Med Biol*, 2007. **52**(6): p. 1547-64.
4. Yamamoto, T., et al., *Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients*. *Int J Radiat Oncol Biol Phys*, 2008. **72**(4): p. 1250-8.
5. Cai, J., et al., *Effects of breathing variation on gating window internal target volume in respiratory gated radiation therapy*. *Med Phys*, 2010. **37**(8): p. 3927-34.
6. Clements, N., et al., *The effect of irregular breathing patterns on internal target volumes in four-dimensional CT and cone-beam CT images in the context of stereotactic lung radiotherapy*. *Med Phys*, 2013. **40**(2): p. 021904.
7. Badawi, A.M., et al., *Optimizing principal component models for representing interfraction variation in lung cancer radiotherapy*. *Med Phys*, 2010. **37**(9): p. 5080-91.
8. Balik, S., et al., *Evaluation of 4-dimensional computed tomography to 4-dimensional cone-beam computed tomography deformable image registration for lung cancer adaptive radiation therapy*. *Int J Radiat Oncol Biol Phys*, 2013. **86**(2): p. 372-9.
9. Mutaf, Y.D., et al., *A simulation study of irregular respiratory motion and its dosimetric impact on lung tumors*. *Phys Med Biol*, 2011. **56**(3): p. 845-59.
10. Mampuya, W.A., et al., *Interfraction variation in lung tumor position with abdominal compression during stereotactic body radiotherapy*. *Med Phys*, 2013. **40**(9): p. 091718.
11. Van den Begin, R., et al., *Impact of inadequate respiratory motion management in SBRT for oligometastatic colorectal cancer*. *Radiother Oncol*, 2014. **113**(2): p. 235-9.
12. Abdelnour, A.F., et al., *Phase and amplitude binning for 4D-CT imaging*. *Phys Med Biol*, 2007. **52**(12): p. 3515-29.
13. Ehrhardt, J., et al., *An optical flow based method for improved reconstruction of 4D CT data sets acquired during free breathing*. *Med Phys*, 2007. **34**(2): p. 711-21.
14. Fitzpatrick, M.J., et al., *Displacement-based binning of time-dependent computed tomography image data sets*. *Med Phys*, 2006. **33**(1): p. 235-46.
15. Lu, W., et al., *A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4D CT*. *Med Phys*, 2006. **33**(8): p. 2964-74.
16. Mutaf, Y.D., J.A. Antolak, and D.H. Brinkmann, *The impact of temporal inaccuracies on 4DCT image quality*. *Med Phys*, 2007. **34**(5): p. 1615-22.
17. Pan, T., X. Sun, and D. Luo, *Improvement of the cine-CT based 4D-CT imaging*. *Med Phys*, 2007. **34**(11): p. 4499-503.

18. Rietzel, E. and G.T. Chen, *Improving retrospective sorting of 4D computed tomography data*. Med Phys, 2006. **33**(2): p. 377-9.
19. Polycarpou, I., et al., *Impact of respiratory motion correction on SPECT myocardial perfusion imaging using a mechanically moving phantom assembly with variable cardiac defects*. J Nucl Cardiol, 2015.
20. Sureshbabu, W. and O. Mawlawi, *PET/CT imaging artifacts*. J Nucl Med Technol, 2005. **33**(3): p. 156-61; quiz 163-4.
21. Colgan, R., et al., *Planning lung radiotherapy using 4D CT data and a motion model*. Phys Med Biol, 2008. **53**(20): p. 5815-30.
22. Li, T., et al., *Model-based image reconstruction for four-dimensional PET*. Med Phys, 2006. **33**(5): p. 1288-98.
23. Brock, K.K. and C. Deformable Registration Accuracy, *Results of a multi-institution deformable registration accuracy study (MIDRAS)*. Int J Radiat Oncol Biol Phys, 2010. **76**(2): p. 583-96.
24. Sonke, J.J., et al., *Respiratory correlated cone beam CT*. Med Phys, 2005. **32**(4): p. 1176-86.
25. Gai, N.D., et al., *Evaluation of optimized breath-hold and free-breathing 3D ultrashort echo time contrast agent-free MRI of the human lung*. J Magn Reson Imaging, 2016. **43**(5): p. 1230-8.
26. Johnson, K.M., et al., *Optimized 3D ultrashort echo time pulmonary MRI*. Magn Reson Med, 2013. **70**(5): p. 1241-50.
27. Roach, D.J., et al., *Ultrashort Echo-Time Magnetic Resonance Imaging Is a Sensitive Method for the Evaluation of Early Cystic Fibrosis Lung Disease*. Ann Am Thorac Soc, 2016. **13**(11): p. 1923-1931.
28. Sheng, K., et al., *A computer simulated phantom study of tomotherapy dose optimization based on probability density functions (PDF) and potential errors caused by low reproducibility of PDF*. Med Phys, 2006. **33**(9): p. 3321-6.
29. Cossmann, P.H., *Video-coaching as biofeedback tool to improve gated treatments: Possibilities and limitations*. Z Med Phys, 2012. **22**(3): p. 224-30.
30. Neicu, T., et al., *Synchronized moving aperture radiation therapy (SMART): improvement of breathing pattern reproducibility using respiratory coaching*. Phys Med Biol, 2006. **51**(3): p. 617-36.
31. Pollock, S., R. Keall, and P. Keall, *Breathing guidance in radiation oncology and radiology: A systematic review of patient and healthy volunteer studies*. Med Phys, 2015. **42**(9): p. 5490-509.
32. Sarrut, D., *Deformable registration for image-guided radiation therapy*. Z Med Phys, 2006. **16**(4): p. 285-97.
33. Kaus, M.R., et al., *Assessment of a model-based deformable image registration approach for radiation therapy planning*. Int J Radiat Oncol Biol Phys, 2007. **68**(2): p. 572-80.

34. Ostergaard Noe, K., et al., *Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy*. Acta Oncol, 2008. **47**(7): p. 1286-93.
35. Low, D., *4D imaging and 4D radiation therapy: a New Era of therapy design and delivery*. Front Radiat Ther Oncol, 2011. **43**: p. 99-117.
36. Paganetti, H., *Four-dimensional Monte Carlo simulation of time-dependent geometries*. Phys Med Biol, 2004. **49**(6): p. N75-81.
37. Qin, Y., et al., *Adaptive stereotactic body radiation therapy planning for lung cancer*. Int J Radiat Oncol Biol Phys, 2013. **87**(1): p. 209-15.
38. Cao, K., et al., *Improving Intensity-Based Lung CT Registration Accuracy Utilizing Vascular Information*. Int J Biomed Imaging, 2012. **2012**: p. 285136.
39. Ding, K., et al., *4DCT-based measurement of changes in pulmonary function following a course of radiation therapy*. Med Phys, 2010. **37**(3): p. 1261-72.
40. Du, K., et al., *Reproducibility of intensity-based estimates of lung ventilation*. Med Phys, 2013. **40**(6): p. 063504.
41. Vinogradskiy, Y., et al., *Use of 4-dimensional computed tomography-based ventilation imaging to correlate lung dose and function with clinical outcomes*. Int J Radiat Oncol Biol Phys, 2013. **86**(2): p. 366-71.
42. Vinogradskiy, Y., et al., *Comparison of 4-dimensional computed tomography ventilation with nuclear medicine ventilation-perfusion imaging: a clinical validation study*. Int J Radiat Oncol Biol Phys, 2014. **89**(1): p. 199-205.
43. Vinogradskiy, Y.Y., et al., *Use of weekly 4DCT-based ventilation maps to quantify changes in lung function for patients undergoing radiation therapy*. Med Phys, 2012. **39**(1): p. 289-98.
44. Castillo, R., et al., *A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets*. Phys Med Biol, 2009. **54**(7): p. 1849-70.
45. Johnson, H.J. and G.E. Christensen, *Consistent landmark and intensity-based image registration*. IEEE Trans Med Imaging, 2002. **21**(5): p. 450-61.
46. Sotiras, A., et al., *Simultaneous geometric--iconic registration*. Med Image Comput Comput Assist Interv, 2010. **13**(Pt 2): p. 676-83.
47. Yin, Y., et al., *A cubic B-spline-based hybrid registration of lung CT images for a dynamic airway geometric model with large deformation*. Phys Med Biol, 2011. **56**(1): p. 203-18.
48. Ehrhardt, J. and C. Lorenz, *4D modeling and estimation of respiratory motion for radiation therapy*. 2013: Springer.
49. Liu, F., et al., *Evaluation of deformable image registration and a motion model in CT images with limited features*. Phys Med Biol, 2012. **57**(9): p. 2539-54.
50. Yeo, U.J., et al., *Performance of 12 DIR algorithms in low-contrast regions for mass and density conserving deformation*. Med Phys, 2013. **40**(10): p. 101701.
51. Cai, J., et al., *Evaluation of the reproducibility of lung motion probability distribution function (PDF) using dynamic MRI*. Phys Med Biol, 2007. **52**(2): p. 365-73.

52. Cai, J., et al., *Reproducibility of interfraction lung motion probability distribution function using dynamic MRI: statistical analysis*. Int J Radiat Oncol Biol Phys, 2008. **72**(4): p. 1228-35.
53. Liu, Y., et al., *T2-weighted four dimensional magnetic resonance imaging with result-driven phase sorting*. Med Phys, 2015. **42**(8): p. 4460-71.
54. Segars, W.P., et al., *4D XCAT phantom for multimodality imaging research*. Med Phys, 2010. **37**(9): p. 4902-15.
55. Pan, T., et al., *4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT*. Med Phys, 2004. **31**(2): p. 333-40.
56. Vedam, S.S., et al., *Acquiring a four-dimensional computed tomography dataset using an external respiratory signal*. Phys Med Biol, 2003. **48**(1): p. 45-62.
57. Cai, J., et al., *MR grid-tagging using hyperpolarized helium-3 for regional quantitative assessment of pulmonary biomechanics and ventilation*. Magn Reson Med, 2007. **58**(2): p. 373-80.
58. Cai, J., et al., *Direct measurement of lung motion using hyperpolarized helium-3 MR tagging*. Int J Radiat Oncol Biol Phys, 2007. **68**(3): p. 650-3.
59. Hu, L., et al., *WE - AB - BRA - 11: Development of Physiological - Based Virtual Thorax Phantoms for Evaluation of Deformable Image Registration*. Medical physics, 2015. **42**(6): p. 3655-3655.
60. Kleiberg, E., H. Van de Wetering, and J.J. Van Wijk. *Botanical visualization of huge hierarchies*. in *InfoVis*. 2001.
61. Holton, M. *Strands, gravity and botanical tree imagery*. in *Computer Graphics Forum*. 1994. Wiley Online Library.
62. Fowler, D.R., P. Prusinkiewicz, and J. Battjes. *A collision-based model of spiral phyllotaxis*. in *ACM SIGGRAPH Computer Graphics*. 1992. ACM.
63. Lintermann, B. and O. Deussen, *Interactive modeling of plants*. IEEE Computer Graphics and Applications, 1999. **19**(1): p. 56-65.
64. Lee, T.-C., R.L. Kashyap, and C.-N. Chu, *Building skeleton models via 3-D medial surface axis thinning algorithms*. CVGIP: Graphical Models and Image Processing, 1994. **56**(6): p. 462-478.
65. Sun, D., et al. *Motion-robust Multi-cycle 4D-MRI Based on Probability-driven K-space Reordering*. in *AAPM 2017*. 2017. Denver, CO.

## Biography

Xiao Liang was born in Beijing, China. He attended Peking University from 2008-2012 and graduated with a Bachelor of Science in Physics. He attended Duke University in the Medical Physics Graduate Program from 2012-2014 and graduated with a Master of Science in Medical Physics. Since 2014, he has continued to work on his doctoral degree in the Medical Physics Graduate Program at Duke University with an anticipated graduation in September 2017.

### Scholarships, Fellowships & Awards

#### **Duke University**

Summer Scholarship for Thesis Concentration, Duke Medical Physics Program

#### **Peking University**

Wusi Scholarship

### Refereed Publications

**Liang, X.**, Yin, F.F., Liu, Y., Cai, J., A probability-based multi-cycle sorting method for 4D-MRI: A simulation study. *Medical Physics*, 2016. 43(12): p.6375

### Conference Abstracts & Presentations

Sun D., **Liang X.**, Yin F.F., Cai J., Motion-robust Multi-cycle 4D-MRI Based on Probability-driven k-space reordering. e-Poster Session at the 59th Annual Meeting of the American Association of Physicists in Medicine (AAPM). July 30-Aug. 3, 2017.

**Liang X.**, Liu Y., Yin F.F., Cai J., A Novel 4D MRI Reconstruction Method Based on Probability-driven Sorting. Poster Session at the ASTRO's 58th Annual Meeting. Sept. 25-Sept. 28, 2016.

**Liang X.**, Wang C., Chang Z., Yin F.F., Cai J., Development of a Deformable Image Registration (DIR) Error Correction Method Employing Kolmogorov-Zurbenko (KZ) Filter. Oral Presentation at the 58th Annual Meeting of the American Association of Physicists in Medicine (AAPM). July 31-Aug. 4, 2016.

**Liang X.**, Yin F.F., Liu Y., Cai J., A Simulation Study on a Novel 4D MRI Reconstruction Method Based on Probability-driven Sorting. Poster Session at the 58th Annual Meeting of the American Association of Physicists in Medicine (AAPM). July 31-Aug. 4, 2016.

**Liang X.**, Yin F.F., Liu Y., Cai J., Improving Patient Setup Accuracy Using a Novel 4D Reconstruction Method Based on Probability-driven Sorting. Poster Session at 2016 Precision Medicine Workshop. June 16-17, 2016.

**Liang X.**, Hu L., Yin F.F., Cai J., Displacement Vector Field (DVF) Error Correction Using Sparsely Distributed Ground-Truth Displacement Vectors. Oral Presentation at the 57th Annual Meeting of the American Association of Physicists in Medicine (AAPM). July 12-16, 2015.

**Liang X.**, Liu Y., Yin F.F., Cai J., Probability-Driven Sorting Method for Reconstructing Variation-Independent 4D-MRI. Oral Presentation at the 57th Annual Meeting of the American Association of Physicists in Medicine (AAPM). July 12-16, 2015.

**Liang X.**, Czito B., Palta M., Bashir M., Yin F.F., Cai J., Evaluation of Liver Deformation During Breathing Using Deformable Image Registration: A Comparison Between 4D CT and 4D MRI. Poster Session at the 56th Annual Meeting of the American Association of Physicists in Medicine (AAPM). July 20-24, 2014.

Cui T., **Liang X.**, Czito B., Palta M., Bashir M., Yin F.F., Cai J., A Novel Method for Evaluating Hepatic Stiffness Based On 4D-MRI and Deformable Image Registration. Oral Presentation at the 56th Annual Meeting of the American Association of Physicists in Medicine (AAPM). July 20-24, 2014.