2013

IMAGE-BASED RESPIRATORY MOTION EXTRACTION AND RESPIRATION-CORRELATED CONE BEAM CT (4D-CBCT) RECONSTRUCTION

Salam Dhou
Virginia Commonwealth University

School of Engineering
Virginia Commonwealth University

This is to certify that the dissertation prepared by Salam Dhou entitled IMAGE-BASED RESPIRATORY MOTION EXTRACTION AND RESPIRATION-CORRELATED CONE BEAM CT (4D-CBCT) RECONSTRUCTION has been approved by her committee as satisfactory completion of the dissertation requirement for the degree of Doctor of Philosophy in Engineering

---

Alen Docef, Ph.D., Dissertation Director, School of Engineering

---

Geoffrey D. Hugo, Ph.D., Committee Member, School of Medicine

---

Ashok Iyer, Ph.D., Committee Member, School of Engineering

Kayvan Najarian, Ph.D., Committee Member, School of Engineering

---

Ruixin Niu, Ph.D., Committee Member, School of Engineering

Supriyo Bandyopadhyay, Ph.D., Graduate Program Coordinator, Electrical and Computer Engineering

---

Barbara D. Boyan, Ph.D., Dean, School of Engineering

F. Douglas Boudinot, Ph.D., Dean, School of Graduate Studies

May 2013

IMAGE-BASED RESPIRATORY MOTION EXTRACTION AND RESPIRATION-CORRELATED CONE BEAM CT (4D-CBCT) RECONSTRUCTION

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

by

Salam Dhou

Director: Alen Docef, Ph.D.

Department of Electrical and Computer Engineering
School of Engineering
Virginia Commonwealth University
Richmond, Virginia
May 2013

To my husband, my daughters, and the family

Acknowledgment

I would like to express my deeply-felt thanks to my advisor, Professor Alen Docef, for his warm encouragement and thoughtful guidance. Your brilliant suggestions and discussions always helped me formulate my ideas and accomplish my tasks. I consider myself lucky being your graduate student. I would like also to express my special appreciation and thanks to my co-advisor, Professor Geoffrey D. Hugo, you have been a tremendous mentor for me. I would like to thank you for encouraging my research on medical physics area and allowing me to have such an interest. Your advice on both research as well as on my career have been invaluable.

I would also like to thank my committee and faculty members, Professor Ashok Iyer, Professor Ruixin Niu, Professor Kayvan Najarian, and Professor Yuichi Motai for serving as my committee members and letting my defense be an enjoyable moment. Thank you for your revision, great comments and suggestions on my dissertation.

I would like to acknowledge the financial, academic and technical support of Virginia Commonwealth University and its staff, particularly the School of Engineering. I would like also to thank the Radiation Oncology Department and Massey Cancer Center for providing us with clinical datasets used for this research.

I would like to thank my beloved husband, Dr. Mamoun Melhem. Words cannot express how grateful I am to you for all of the sacrifices that you've made on my behalf. Thank you for supporting me for everything, and especially I can't thank you enough for encouraging me throughout this experience. To my beloved daughters Layal and Lujayn, I would like to express my thanks for being such good girls, always cheering me up and understanding when I am unable to play all the time. Special thanks to you, Layal, for all

the inspiring letters and drawings you made for me with the title ‘I love you mom’. Special thanks to my family, family-in-law and friends in Jordan, the United States and elsewhere. Your continuous encouragement and prayers for me helped me survive the stress and keep up.

Of all the blessings and happiness in my life, I always thank God for giving me the health, wisdom and ability to pass through all the difficulties and finish my degree. I have experienced His grace and merciful care every day. Thanks God for everything.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

IMAGE-BASED RESPIRATORY MOTION EXTRACTION AND RESPIRATION-CORRELATED CONE BEAM CT (4D-CBCT) RECONSTRUCTION

By SALAM DHOU

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

Virginia Commonwealth University, 2013

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

Accounting for respiration motion during imaging helps improve targeting precision in radiation therapy. Respiratory motion can be a major source of error in determining the position of thoracic and upper abdominal tumor targets during radiotherapy. Thus, extracting respiratory motion is a key task in radiation therapy planning. Respiration-correlated or four-dimensional CT (4DCT) imaging techniques have been recently integrated into imaging systems for verifying tumor position during treatment and managing respiration-induced tissue motion. The quality of the 4D reconstructed volumes is highly affected by the respiratory signal extracted and the phase sorting method used. This thesis is divided into two parts. In the first part, two image-based respiratory signal extraction methods are proposed and evaluated. Those methods are able to extract the respiratory signals from CBCT images without using external sources, implanted markers or even dependence on any structure in the images such as the diaphragm. The first method, called Local Intensity Feature Tracking (LIFT), extracts the respiratory signal depending on feature points extracted and tracked through the sequence of projections.

The second method, called Intensity Flow Dimensionality Reduction (IFDR), detects the respiration signal by computing the optical flow motion of every pixel in each pair of adjacent projections. Then, the motion variance in the optical flow dataset is extracted using linear and non-linear dimensionality reduction techniques to represent a respiratory signal. Experiments conducted on clinical datasets showed that the respiratory signal was successfully extracted using both proposed methods and it correlates well with standard respiratory signals such as diaphragm position and the internal markers' signal. In the second part of this thesis, 4D-CBCT reconstruction based on different phase sorting techniques is studied. The quality of the 4D reconstructed images is evaluated and compared for different phase sorting methods such as internal markers, external markers and image-based methods (LIFT and IFDR). Also, a method for generating additional projections to be used in 4D-CBCT reconstruction is proposed to reduce the artifacts that result when reconstructing from an insufficient number of projections. Experimental results showed that the feasibility of the proposed method in recovering the edges and reducing the streak artifacts.

CHAPTER 1 INTRODUCTION

Lung tumors move as the patient breathes causing a significant change in the tumor's shape and location during radiotherapy. Thus, it is important to adjust the delivered dose to the moving tumor and the surrounding normal tissues to avoid irradiating the nearby healthy tissues. Four-Dimensional or respiration-correlated CT (4D-CBCT) imaging techniques have been recently and rapidly developed in image-guided radiation therapy for patient positioning prior to the treatment. Respiratory motion can effectively indicate the tumor position. It is considered one potential source of error in treatment. A respiratory signal can be generated from external sources (skin markers, abdominal belts, or spirometry) [1] [2] or an image-based estimate of diaphragm position [3] [4] [5]. However, these signals are external or with a limited field of view so they do not correlate robustly with tumor position. Also, the success of respiratory signal extraction using image-based diaphragm position estimation depends on the clear visibility of the diaphragm in the entire projection set, which is not always the case. Implanted radio-opaque tumor markers [6] [7] can help in this issue but has an additional expense which may not be available or may delay the treatment initiation. Thus, the first objective of the dissertation is to propose two alternative methods to extract the respiratory signal using only the Cone-Beam Computed Tomography (CBCT) scan projections without depending on any structure in the projections such as the diaphragm.

The respiratory signal is used in 4D-CBCT reconstruction methods to create reconstructed image volumes. All CBCT projections, acquired from a standard CBCT scan, are sorted into several respiratory phase bins according to a respiratory signal extracted. Then, a standard three-dimensional (3D) CBCT method is used for

reconstruction. Thus, each subset of projections representing one phase of the respiratory cycle is used to reconstruct a 3D CBCT image. However, the insufficient angular sampling of the projections per respiratory phase causes view-aliasing artifacts characterized by high-frequency streaks. Thus, the second objective of the dissertation is to propose a novel method to increase the number of projections in each phase of the respiratory cycle. This method estimates the motion of the organ tissues in every CBCT projection using optical flow and generates intermediate projection views depending on the motion of the tissues and the intensity information of the original projections motion in the same cycle.

Thus, the objectives of this study are twofold. 1) Propose two alternative image-based signal extraction methods (CHAPTER 2 and CHAPTER 3), and 2) analyze the effect of different phase sorting methods on the 4D-CBCT reconstruction quality and propose a new projection generation method to reduce the streaking artifacts (CHAPTER 4). In Chapter 2, an image-based respiration signal extraction method, termed Local Intensity Feature Tracking (LIFT), is proposed. LIFT uses only the CBCT projections of the lung to extract the respiratory signal depending on tissue feature points local to the tumor without dependence on a particular anatomical structure in the CBCT images (such as the diaphragm). LIFT works by extracting and tracking feature points from one projection to another using optical flow [8] to form trajectories of the feature points' correspondences. A selection criterion is applied to select the trajectories that show an oscillating behavior similar to respiration. Using those selected trajectories, the 3D motion is recovered and analyzed to represents the respiration motion. This work has been successfully completed and published as in [9]. In CHAPTER 3, the research work is continued to recover the

respiratory signal from CBCT projections. Another alternative image-based respiration signal extraction method, termed Intensity Flow Dimensionality Reduction (IFDR), is proposed. IFDR method detects the respiration signal by computing a dense optical flow on every pixel of each pair of adjacent CBCT projection images of the patient dataset. Since we know a pattern of a respiration motion exists in the optical flow displacement dataset, we apply linear and non-linear dimensionality reduction techniques to the consecutive optical flow displacement vectors detected to extract this respiratory motion. The principal components of the eigen system resulting from the linear dimensionality reduction method are used to represent the respiratory signal. Similarly, the dimensions of the reduced-dimension dataset resulting from the non-linear dimensionality reduction method serve as the respiratory signal of the lung. This proposed method has been applied to three clinical datasets and the experimental results showed that the respiratory signal extracted using this method correlates with the standard respiratory signals.

In CHAPTER 4, the study aims to address the effect of using different phase sorting techniques on the quality of the 4D-CBCT reconstructed volume. Different phase sorting method has been used and compared including internal markers, external markers and image-based methods (the proposed LIFT and IFDR). Also, a method for generating additional projections to be used in 4D-CBCT reconstruction is proposed to reduce the streaking artifacts that result from reconstructing from an insufficient number of projections. Those new generated projections belong to the same breath but having different projections angles of a gantry rotation of 360º degrees. The generated projections of the same phase are used to reconstruct a 4D-CBCT volume. This method uses the feature extraction and tracking approach discussed in [9] (CHAPTER 2).

Experiments have been conducted to validate the study. Results showed that the proposed method reduced some of the artifacts and blurring in the 4D-CBCT volumes and improved the image quality. CHAPTER 5 summarizes the findings of research work in CHAPTER 2, CHAPTER 3, and CHAPTER 4.

1.1 Respiratory Signal Extraction in Cone-Beam CT (CBCT)

Cone-beam computed tomography (CBCT) is a powerful imaging tool in image-guided radiation therapy (IGRT) [10] [11] [12]. It provides volumetric information for accurate target localization. However, when a moving organ such as the lung or heart is scanned, motion can introduce artifacts in a planning CT scan and blur in a CBCT. This challenge sites that motion often significantly degrade the image quality and restrict the use of CBCT. Respiration-correlated or four-dimensional CT (4D-CT), where projections are sorted according to the respiratory phase, has been shown to reduce respiratory motion artifacts and yield 3D images at different points in the respiratory cycle [13] [14] [15] [16].

Respiratory correlated imaging techniques require a respiratory signal. Respiratory signals can be extracted using external sources such as skin markers, abdominal belts, or spirometry [1] [2]. Those methods require additional equipment such as infrared cameras, detectors, or spirometry that may not be available. Another solution was an image-based measure of diaphragm position directly from the 4DCBCT radiographic projections [3] [4] [5]. The extraction of a diaphragm position based signal requires the diaphragm to be visible in all acquired CBCT projections, which is not possible in some commonly-used CBCT systems that have limited longitudinal fields of view. Also, a number of studies have shown that tumor position is difficult to predict

directly from the diaphragm or external surrogate motion, with reported errors of up to 6 mm in predicting tumor position from external marker position [17] [18] [19] [20] [21] [22] [23] [24] [25] [26]. Another option is using transthoracically or bronchoscopically radio-opaque tumor markers which are implanted near the tumor and tracked to extract the respiratory motion [6] [7]. This method solves the problem, but the additional procedure to implant the markers is invasive and expensive, and any complications may significantly delay treatment initiation.

There are several image-based marker-free respiratory signal extraction methods [3] [4] [27] [28] [29] [30]. Bergner *et al.* [27] proposed a method for measuring motion between 4DCBCT projections based on dense optical flow using a Horn-Schunk [31] implementation. Their method was developed to improve the reconstruction quality in stationary regions of the anatomy. For respiration sorting, they used a diaphragm position-based method similar to [3]. Wachinger *et al.* [32] extract an image-based respiratory signal using manifold learning; however, this method was applied only for fixed view (fluoroscopic) images, as opposed to the rotational view (tomographic) projections used in this study. In [4], thousands of interest points are selected and tracked across projections using a block matching algorithm. Signal processing techniques are then used to acquire a respiratory signal. In [28], Siochi developed a technique that locates a bounding box for the diaphragm motion for all projections based on two pairs of full-inhale and full-exhale views. However, the above methods [4] and [28] require that the diaphragm be present in the field of view. In [29], the respiratory signal is acquired from an analysis of the variation in pixel values between projection images by developing a simple pixel value summation followed by a high pass filtering. This method is

effective, but its efficiency should be demonstrated for various CBCT patient cases and compared to other methods in the literature. Vergalasova *et al.* [30] proposed a markerless method for respiration signal detection based on the principles on the Fourier Transform theory. However, when the entire projections of patient data were used, the results showed a big phase shift between the extracted signal and the ground truth signal. Our proposed image-based respiratory signal extraction methods extract the respiratory signal without the dependence on any structure in the images and the resulted signal correlates with the standard respiration signals (*e.g.* using the implanted markers).

1.2 Respiration-Correlated Cone-Beam CT (4D-CBCT)

Respiration-Correlated or Four-dimensional cone beam computed tomography (4D-CBCT) provides respiratory phase-resolved CBCT images with the 4th dimension being time. 4D-CBCT has been proposed to characterize the breathing motion of tumors before radiotherapy treatment. 4DCBCT using an on board imager [3] [33] [34] [1] has been proposed as a 4D imaging tool for radiotherapy treatment of tumors in the thorax and upper abdomen. Compared with four-dimensional computed tomography (4D-CT) using diagnostic CT scanners [13] [14] [15] [16] [35] [36] [37], 4D-CBCT extends 4D imaging capability from the treatment planning stage to the treatment delivery stage. From 4D-CBCT images, a tumor motion trajectory can be extracted immediately before the treatment delivery and compared with that from the treatment plan so that the treatment can be adapted to optimize the clinical outcome.

4D-CBCT is implemented on the basis of the standard three-dimensional CBCT (3D-CBCT). First, all projection views, acquired from a standard CBCT scan, are sorted into several respiratory phase bins according to a respiratory surrogate. Second, image

reconstruction is performed for each of the phase bins. In such a way, time resolved image volumes are generated which represent the temporal change of the patient anatomy during a respiratory cycle [38]. Recently, CBCT scanners have been integrated with linear accelerators to acquire 3D-CBCT images of the patients for image guidance of radiotherapy [10]. These 3D-CBCT images allow correcting for the target misalignment and adapting the treatment plan [39]. However, respiratory motion causes artifacts in CBCT images of the thoracic and upper abdominal region, such as blur and streaks, which are known to reduce the accuracy of derived information [40].

Respiration-correlated CBCT (4D-CBCT) is the first solution to account for the respiratory motion [1] [3] [34] [41]. It consists of sorting the CB projections depending on their position in the respiratory cycle assessed with a respiratory signal. Each subset of CB projections is then used to reconstruct a 3D-CBCT image representing one phase of the respiratory cycle, thus obtaining a 4D-CBCT image of the respiratory cycle. Although 4D-CBCT is capable of reducing the motion artifacts, it poses another challenge for reconstruction. Reconstruction from one phase of the respiratory cycle leads to insufficient number of x-ray projections in each respiratory phase bin. Reconstructing such an under-sampled data using the standard filtered-back-projection (FBP) [42] algorithm can cause severe view-aliasing and streaking artifacts in the reconstructed images.

Various methods have been proposed to address the under-sampling problem and improve image quality in 4D-CBCT [1] [3] [27] [43] [44] [45] [46] [47] [48] [49] [50] [51] [52] [53] [54] [55]. Most of them can be classified into two categories: (1) slow down the gantry rotation or do multiple gantry rotations while using the standard FBP

reconstruction algorithm [1] [3] [44] [45] or (2) keep the standard gantry rotation time while using advanced methods to mitigate streak artifacts [27] [43] [47] [48] [49] [50] [51] [52] [53] [54] [55]. Among the advanced reconstruction techniques is Compressed Sensing (CS), which was applied to tomographic beam problems by Sidkey *et al.* [50]. It uses a transformation onto the image and tries to find the transformed representation of the image with the least significant entries, while simultaneously bounding the error between the projected image and the measured data. Another CS example was [51] which uses a simple gradient descent for the optimization. Also, a prior image constrained compressed sensing (PICCS)-based algorithm, was proposed to reconstruct each image by regularizing the total variation of the image and its difference from a prior image obtained by using all projections [52]. More work on CS can be found in [53] [54] [55].

Motion compensating algorithms are also available. Those algorithms use further prior information, *i.e.*, deformation fields calculated from the planning CT [56] [43] [57] [58] [59] [60] [61]. The approach consists of two steps: first, it estimates the patient motion during the CB acquisition and second, it uses the estimated motion in the reconstruction algorithm. Thus, a 3D-CBCT image at a reference position is reconstructed from all the CBCT projections. In [43] [57] the computational cost added to the cost of the reconstruction algorithm has prevented the clinical use of motion-compensated CBCT. In [60] [61], an algorithm based on the PCA lung motion model has been proposed and evaluated to reconstruct volumetric images and extract 3D tumor motion information in real-time from a single x-ray projection in a marker-less implantation. It has been also proposed to split the reconstruction region according to volume of interest and treat the reconstructions separately [62]. These motion

compensating approaches are believed to give the best results when accurate deformation information are available as the motion can potentially be almost completely compensated in the reconstruction algorithm [63]. However, the calculation of the deformation maps is costly and the efficacy of these approaches largely depends on the accuracy of the algorithms involved, such as deformable image registration algorithms.

Recently, nonlocal means (NLM) operators have become an effective tool for solving image restoration problems. The underlying assumption is that the image to be restored contains repetitive features that can be utilized to constructively enhance each other. Based on this idea, a generalized (NLM) method, termed temporal nonlocal means (TNLM) [64] [65] has been extended to 4D-CBCT problems. An enhanced version of TNLM algorithm is presented in [66] with a GPU's parallel processing scheme implementation. In this TNLM-based 4D-CBCT enhancement algorithm, 4DCBCT images are first reconstructed by the conventional FDK [42] algorithm and post-processed by utilizing a TNLM approach to remove the streaking artifacts caused by the FDK algorithm due to the insufficient number of projections. Also, a number of research efforts have been made on post-processing of the 4D-CBCT images. For example, a prior image-based approach has been developed by first reconstructing a blurred CBCT images with all projections and then using it to estimate and remove the streaking artifacts [47].

Increasing the sampling density by projections generation may help to reduce the effects of the view aliasing artifacts. The topic of image interpolation has been studied in medical imaging research [67] [68] [69] [70] [71] [72] [27] [73] [74] [75]. For parallel beam geometry, Weiss *et al.* [69] estimated intermediate phantom views using linear interpolation and demonstrated their efficiency in minimizing view aliasing streaks. A set

of interpolation schemes have been described in the literature for medical and nonmedical applications, such as optimized interpolation kernels [71] [72] [76], shape-adaptive image interpolation algorithms [77] [78] [79] [80] and other advanced interpolation methods [67] [68] [81]. In [67] [68], a shape-driven directional interpolation algorithm based on a structure tensor approach is developed. The algorithm is locally adaptive to the orientation of gray value structures to be interpolated. Quantitative evaluation shows that this method outperforms conventional scene-based interpolation schemes and reduces streak artifacts and noise in the reconstructed images. However, this method was applied to a non-moving anatomy, *i.e.* human head, as opposed to the human lung datasets used in this study.

CHAPTER 2 IMAGE-BASED RESPIRATORY SIGNAL EXTRACTION USING LOCAL INTENSITY FEATURE TRACKING AND MOTION MODELING IN CONE-BEAM CT PROJECTIONS

Accounting for respiration motion during imaging can help improve targeting precision in radiation therapy. This chapter presents Local Intensity Feature Tracking (LIFT), a novel marker-less breath phase sorting method in Cone Beam Computed Tomography (CBCT) scan images. LIFT extracts the respiratory signal from the CBCT projections of the thorax depending only on tissue feature points that exhibit respiration. The extracted respiratory signal using LIFT is shown to correlate with standard respiration signals.

2.1 INTRODUCTION

Respiratory motion extraction from CBCT images is an important task in biomedical engineering research [82] [83] [84]. It can be used to study the influence of organ motion on CBCT imaging [1]. Motion modeling can be used also for measuring the position and orientation of objects [85] [86] or for respiratory image sequence segmentation techniques [87]. Four-dimensional or respiration-correlated CT (4DCT) imaging techniques have become a basic task in radiation therapy planning. As the respiratory motion can be a major source of error in determining the position of thoracic and upper abdominal tumor targets during radiotherapy, extracting respiratory motion is a key task in reconstructing 4DCT. Volumetric image guidance techniques, such as 4D Cone Beam CT (4DCBCT) have been recently and rapidly integrated into the clinic for verifying tumor position during treatment and managing respiration-induced tissue motion [1] [3] [34]. An acquired respiratory signal serves as a surrogate for the tumor position. This surrogate is used to assign each projection to its appropriate breathing phase bin, in a process termed "sorting," prior to 4D image reconstruction.

Respiratory signals can be extracted using external sources such as skin markers, abdominal belts, or spirometry [1] [2]. Those methods require additional equipment such as infrared cameras, detectors, or spirometry that may not be available. Another solution was an image-based measure of diaphragm position directly from the 4DCBCT radiographic projections [3] [4]. The extraction of a diaphragm position based signal requires the diaphragm to be visible in all acquired CBCT projections, which is not possible in some commonly-used CBCT systems that have limited longitudinal fields of view. Also, a number of studies have shown that tumor position is difficult to predict

directly from the diaphragm or external surrogate motion, with reported errors of up to 6 mm in predicting tumor position from external marker position [17] [18] [19] [20] [21] [22] [23] [24] [25] [26]. Another option is using transthoracically or bronchoscopically radio-opaque tumor markers which are implanted near the tumor and tracked to extract the respiratory motion [6] [7]. This method solves the problem, but the additional procedure to implant the markers is invasive and expensive, and any complications may significantly delay treatment initiation.

Thus, Local Intensity Feature Tracking (LIFT), an image-based respiration signal extraction method is proposed. The contributions of this study are twofold. First, the proposed method uses only the CBCT projections of the lung to extract the respiratory signal depending on tissue feature points local to the tumor. Second, the extracted respiration signal using LIFT correlates to the standard respiration signals. The specific novel contributions of this work are: 1) LIFT extracts the respiratory signal without dependence on a particular anatomical structure in the CBCT images (such as the diaphragm), and can be tuned to focus on a particular region of the anatomy (say, near a tumor). 2) No other group has used 3D motion modeling in recovering the 3D motion of the lung and used it as a respiratory signal. LIFT works as follows. First, the CBCT projections of the entire dataset are grouped into arcs. Then, in each arc, feature points are extracted in the lung. Those feature points are tracked from one projection to another using optical flow [8] to form trajectories of the feature points' correspondences. A selection criterion is applied to select the trajectories that show an oscillating behavior similar to respiration. Using those selected trajectories, the 3D motion is recovered. Then, the 3D rotation around the Z-axis of the patient represents the respiration motion in this

Table 1: Comparison of Respiratory Motion Detection Methods

Method name	Procedure	Requirements
Diaphragm position [3] [4]	Measuring the position of the diaphragm in subsequent projections and using it as a surrogate to the respiration motion	Diaphragm should be visible in all projections
Internal Marker-based [6] [7]	Tracking the markers in subsequent projection images using image-based tracking techniques	Internal markers/ implanting procedure costs
Proposed LIFT	Tracking interest points local to the tumor position and 3D motion modeling of the 2D motion of the best selected motion trajectories	CT projections only

study. CBCT projections are sorted into phases according to the respiration signal detected.

The advantage of LIFT over the external resources methods is that no external equipment is required. It also has the advantage over the diaphragm position method in that the diaphragm is not required to be visible in all CBCT projections. Moreover, no internal markers are required to be implanted as the signal is generated from tissue features local to the tumor position. LIFT is considered widely applicable as only the CBCT projections are required to generate the respiratory signal. There are several works related to our study. Bergner et al. [27] proposed a method for measuring motion between 4DCBCT projections based on dense optical flow using a Horn-Schunk [31] implementation. Their method was developed to improve the reconstruction quality in stationary regions of the anatomy. For respiration sorting, they used a diaphragm position-based method similar to [3]. Wachinger *et al.* [32] extract an image-based respiratory signal using manifold learning; however, this method was applied only for fixed view (fluoroscopic) images, as opposed to the rotational view (tomographic) projections used here.

The works in [4] [28] [29] [30] are also related. In [4], thousands of interest points are selected and tracked across projections using a block matching algorithm. Signal

processing techniques are then used to acquire a respiratory signal. In [28], Siochi developed a technique that locates a bounding box for the diaphragm motion for all projections based on two pairs of full-inhale and full-exhale views. However, the above methods [4] and [28] require that the diaphragm be present in the field of view. In [29], the respiratory signal is acquired from an analysis of the variation in pixel values between projection images by developing a simple pixel value summation followed by a high pass filtering. This method is effective, but its efficiency should be demonstrated for various CBCT patient cases and compared to other methods in the literature. Vergalasova *et al.* [30] proposed a markerless method for respiration signal detection based on the principles on the Fourier Transform theory. However, when the entire projections of patient data were used, the results showed a big phase shift between the extracted signal and the ground truth signal. Thus, our proposed method is the first method to use feature tracking and 3D motion modeling for respiratory signal extraction. Table 1 shows a comparison summary of respiration signal extraction methods.

The remainder of this chapter is organized as follows. Section 2.2 describes in detail the proposed method. Section 2.3 presents the experimental results on four clinical datasets and a discussion of the results. Section 2.4 summarizes the findings.

2.2 Proposed Respiespiratory Motion Extraction Method

2.2.1 Feature Extraction and Tracking

Feature points are extracted and tracked through projection sequences to detect the optical flow motion in the following three steps:

Step 1: Feature points are extracted at pixel locations equally spaced by a constant number of pixels, not based on image intensity (like corners). That's because of the nature of transmission tomography in which corners visible in 2D projection images may not correspond to actual high contrast boundaries in the 3D anatomy. Also, choosing features at equally spaced locations allows the extraction of feature points in any CT projection image. An extracted feature point is represented by $p_{f,p}=(x_{f,p}, y_{f,p})$, where f is the projection number and p is the point number. The aim is analyzing the motion between any pair of consecutive projections. The inferior portion of projections containing the diaphragm was excluded to simulate common CBCT acquisition systems that have smaller longitudinal fields of view. Fig. 1 shows feature points extraction in one projection.

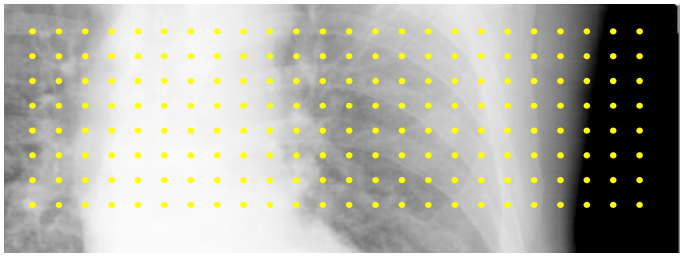


Fig. 1. Feature points extraction used in this study
Feature points are extracted at specific pixel locations spaced by a number of pixels (20 pixels in this figure). The inferior part of the projection image is excluded to simulate the CBCT systems that have a small field of view.

Step 2: Feature points are tracked through subsequent projection frames. Due to the respiration motion and different projection angles, a feature point has different locations (x and y coordinates) in every subsequent projection. Those locations of the feature point are called "correspondences". To find the displacement $\delta(\Delta x,\Delta y)$ that a point $p_{f,p}$ with intensity $I(x_{f,p},y_{f,p},\tau)$ makes when it moves from one image to the next, a single pixel cannot be tracked. That is because the value of the pixel can change due to noise, and be confused with adjacent pixels [88]. Thus, windows of pixels are tracked instead of single pixels. Since adjacent projections refer to the same scene taken from slightly different viewpoints at τ and $\tau+\Delta\tau$, intensity patterns move in the image sequence satisfying the constraint property: $I(x_{f,p},y_{f,p},\tau) = I(x_{f,p}+\Delta x, y_{f,p}+\Delta y, \tau+\Delta\tau)$.

The next image can be defined as $J(p)=I(p-\delta)+\eta(p)$, where η is some noise. The displacement $\delta(\Delta x,\Delta y)$ that minimizes the sum of squared intensity differences between a past and current window w is found in the following cost function:

$$\varepsilon = \int_w [I(p-\delta)-J(p)]^2 dp, \qquad \textbf{(1)}$$

which can be written as $\varepsilon = \int_w (h-g\delta)^2 dp$, where $h=I(p)-J(p)$. The residue is minimized by differentiating ε with respect to δ and setting the result equal to zero as: $\int_w (h-g\delta)gdA = 0$, where $(g\delta)g=(gg^{\mathrm{T}})\,\delta$, and δ is assumed to be constant within w. Therefore, the resulting system is $\left(\int_w gg^T dA\right)\delta = \int_w hgdA$. This system has two scalar equations and two unknowns which can be written as: $G\,\delta=e$, where G can be computed from one image by estimating gradients and computing their second order moments, and e can be computed from the difference between the two images along with the gradient computed above. The displacement δ is then the solution of system [88].

Drift problem is a very classical issue when tracking feature points through long sequences. To overcome this problem, the quality of feature points is monitored during tracking by measuring the dissimilarity of the features between the first image in the sequence and the current image. The feature is abandoned when dissimilarity grows too large. Affine image changes are used to calculate dissimilarity as in [89].

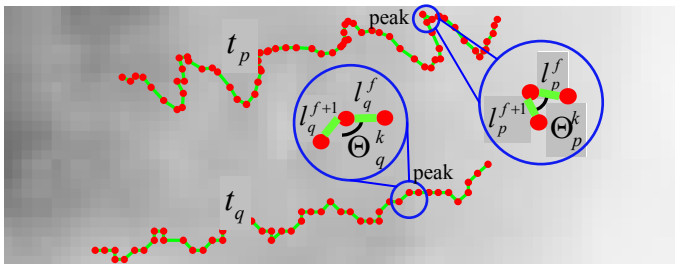


Fig. 2. Two trajectories formed by tracking a sequence of 47 frames
Both trajectories t_p and t_q are selected to be used in the reconstruction of the respiration signal. The first two segments of both trajectories are illustrated.

Step 3: A trajectory is formed as sequence of point correspondences through F frames and is defined by t_p as: $t_p=\{p_{1,p}, p_{2,p}, \dots, p_{F,p}\}$. A set of trajectories is represented by $T=\{t_1, t_2, \dots, t_P\}$, where P is the total number of trajectories which equals the total number of points. Trajectories are represented by a list of line-angle vectors (l,Θ), where l is the line between two adjacent feature points $p_{f,p}$ and $p_{f+1,p}$, and Θ is the value of the angle between two adjacent lines. The line-angle vector representation of a trajectory t_p of F point correspondences is: $t_p=\left((l_p^1,\Theta_p^1),\dots,(l_p^{F-1},\Theta_p^{F-2})\right)$. In a trajectory of F points, the number of lines l equals F-1 while the number of angles Θ equals F-2. This representation of trajectories is defined for the purpose of clustering. The details of two trajectories are shown in Fig. 2. Algorithm 1 shows a summary of the steps used for feature extraction and tracking.

Algorithm 1: Feature Extraction and Tracking
Step 1: Extract feature points $p_{f,p}=(x_{f,p}, y_{f,p})$ at pixel locations equally spaced by a constant number of pixels.
Step 2: Track feature points through a projection sequence by finding the displacement $\delta(\Delta x,\Delta y)$ that minimizes the cost function: $\varepsilon = \int_w [I(p-\delta)-J(p)]^2 dp$, (1)
Step 3: Form trajectories t_p and represent them as a list of line-angle vectors as: $t_p = \left((l_p^1,\Theta_p^1),...,(l_p^{F-1},\Theta_p^{F-2})\right)$.

2.2.2 SHAPE-BASED TRAJECTORY CLUSTERING

Trajectories resulting from tracking the feature points are clustered based on their motion behavior. Respiratory-induced motion in the thorax and upper abdomen is quasiperiodic and directed mainly in the superior-inferior direction (along the patient longitudinal axis). Thus, motion trajectories exhibiting this oscillating behavior may imply respiration motion (*i.e.* thoracic tissue areas). Other trajectories have only orbital motion due to the scanner rotation (*i.e.* bony areas). In order to detect a true respiration signal out of this mixed signal, the effect of the orbital trajectories should be minimized by using only the ones with the highest superior-inferior motion signals in the motion detection process. The following three steps show the process of trajectory clustering.

Step 1: A set of metrics for trajectory clustering is formulated to describe the shape of trajectories. Table 2 describes the set of metrics used and their formula. Those metrics compare trajectories based on their shapes using the number of peaks in each trajectory, angle measurements, and the average number of lines between peaks. A peak in a trajectory, as seen in Fig. 2, is the point that has the largest y coordinate within a breath cycle (3-5 seconds). It is detected by traversing every trajectory to find the points with the largest y coordinate within a breath cycle.

Those metrics are combined in one similarity measure to apply to every pair of trajectories. This similarity measure detects the similarity in the shape of trajectories and is insensitive to their location in the projection and length in pixels. The following is the definition of the similarity measure that uses the distance metrics defined in Table 2.

$$S(t_p, t_q) = d - \begin{pmatrix} \alpha k_{p,q} \\ +\beta \left| \overline{\Theta}_p - \overline{\Theta}_q \right| / \pi \\ +\psi(1-\rho_{p,q}) \end{pmatrix}, \qquad \textbf{(2)}$$

where t_p and t_q are two trajectories of the same length. d is the maximum similarity possible between any two trajectories. α, β, and Ψ are weights to adjust the significance of one metric to the other metrics. All metrics are represented as ratio or normalized difference. Each metric in (2) compares the pairs of trajectories based on one specific aspect of each trajectory's shape. The first metric compares trajectories based on the number of peaks. The second one compares the trajectories based on average angle measurement between the lines forming the peaks. The third one compares the trajectories based on the average number of lines between each consecutive peak. Each metric has its own influence on the result of similarity. The metrics are combined in a weighted summation to determine the overall similarity between each pair of trajectories.

Table 2: Trajectories Clustering Metrics

Metric Formula	Description
$k_{p,q} = \frac{\lvert k_p - k_q \rvert}{\max\{k_p, k_q\}}$	The difference in the number of peaks k_p and k_q in trajectories p and q divided by the maximum number of peaks. $k_{p,q} \leq 1$.
$\overline{\Theta}_p = \frac{\sum_{\alpha=1}^{p} \Theta_p^{\alpha}}{F-2}$	Average measurement value of angles in trajectory p. $\Theta_p^{\,k}$ is the angle existing between the lines $l_p^{\,k}$ and $l_p^{\,k+1}$ forming a peak in trajectory p.
$r_p = \frac{\sum_{k=1}^{n_k} l_p^k}{(n_p^k - 1)}$	The average number of lines $l_p^{\,k}$ between every consecutive peaks in trajectory p, where $n_p^{\,k}$ is the number of lines $l_p^{\,k}$ between two peaks in p.
$\rho_{p,q} = \frac{\min\{r_p, r_q\}}{\max\{r_p, r_q\}}$	The ratio of the average number of lines between peaks in trajectories p and q: r_p and r_q. $\rho_{p,q} \leq 1$.

Those metrics are also weighted using three weighs (α, β, and Ψ) to increase or decrease the influence of a specific metric on the overall similarity.

When each of the weights α, β, and Ψ is set to 1, the maximum similarity possible d should be set to a numeric value that is equal to or greater than 3 in order to have a positive overall similarity $S(t_p,t_q)$. If d is set to 3, the overall similarity measure $S(t_p,t_q)$ will range from 0 (minimal similarity) to 3 (maximal similarity). The weights can be adjusted to give more significance to one metric than the other. For example, to emphasize on the average angle difference, such that similar trajectories should have very similar angles, β should be given a value greater than other weights α and Ψ, and the maximum similarity possible d should be changed accordingly.

Step 2: The hierarchical agglomerative clustering method [90], which is a type of hierarchical clustering, is used to classify trajectories. The similarity is computed for every pair of trajectories based on the similarity measure defined in (2). Clustering is achieved by first finding the closest pairs of trajectories and placing them into a cluster. The similarity between a single trajectory and the new cluster is computed as the average similarity between the single trajectory and all trajectories belonging to the cluster. Then, the most similar pair of trajectories/clusters is combined again in a new cluster until having two clusters eventually. One cluster contains "breath-like oscillating" trajectories and is denoted by T_b. The other one contains "orbital" trajectories and is denoted by T_o. This clustering method is meant to work on regular and irregular breathing patterns. Due to the irregular breathing, trajectories may have "abnormal" shapes. Since the whole anatomy is affected by the same regular/irregular respiration motion, most of the oscillating trajectories will have similar motion behaviors and thus similar shapes. So,

they will be clustered in T_b, the cluster with breathing trajectories, and the other cluster T_o will contain either orbital or non-breathing trajectories. Fig. 3 shows an example of trajectories clustering using our method.

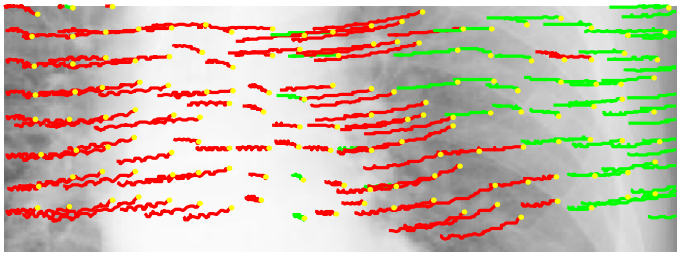


Fig. 3. Trajectories clustering in the projection images
Red trajectories (cluster T_b) show a motion similar to the respiration motion so they are selected to be used in the signal extraction. Green trajectories (cluster T_o) show an orbital motion, so they are discarded. The range of images used is #250–300 of Patient 1 dataset.

Step 3: The result of the clustering process has been evaluated after the process of clustering is done. Two validation criteria described in [91] are used in this study. The first criterion used is the compactness of a cluster which measures the average similarity of trajectories in a cluster. The compactness in cluster T_b is computed as:

$$\overline{S}_b = \sum_{i=1}^{n_b} \sum_{j=i+1}^{n_b} S(t_i, t_j) \Big/ n_b \,, \qquad (3)$$

where n_b is the number of trajectories in T_b. The standard deviation between trajectories in the same cluster T_b is defined as: $\sqrt{\sum_{i=1}^{n_b} \sum_{j=i+1}^{n_b} (S(t_i, t_j) - \overline{S}_b)^2 \Big/ n_b}$.

The second criterion used is the isolation. It measures the separation of the two clusters by estimating the highest similarity to a trajectory outside the cluster. The isolation of the two clusters T_b and T_o is defined as:

$$D(T_b, T_o) = \max(S(t_{i,b}, t_{j,o})), \forall i = 1,..n_b, j = 1,..n_o \,. \quad (4)$$

The smaller the similarity between clusters, the greater the isolation. Algorithm 2 summarizes the steps of clustering.

Algorithm 2: Shape-based Trajectory Clustering
Step 1: Formulate a similarity measure and apply to every pair of trajectories as in (2).
Step 2: Use hierarchical agglomerative clustering based on (2) to cluster trajectories into two clusters T_b and T_o.
Step 3: Validate clustering results by computing the compactness of T_b as in (3), and the isolation of T_b and T_o as in (4).

2.2.3 Respiratory Motion Detection

For the detection of respiratory motion, 3D motion modeling of the selected trajectories is used. The following three steps show the detection of the respiratory motion.

Step 1: The subsequent projections are grouped into overlapping arcs. An *arc* is defined as a sequence of projections captured from a unique and continuous record of X-ray radiation. Projections are grouped into overlapping arcs because corresponding feature points are not visible in all projections due to scanner rotation. Overlapping the arcs allows for a breathing signal to be correlated between two arcs in this region without having the same feature points in both arcs. The process of grouping projections into arcs is done manually based on their projection angle, regardless of their contents and/or the breathing signal they may carry. Projection grouping has no relation to the regularity or periodicity of breathing. The goal of grouping projections into arcs is to overcome the problem of feature points that become invisible because of the scanner rotation regardless of the breathing status. Fig. 4 shows dataset groupings into overlapping arcs.

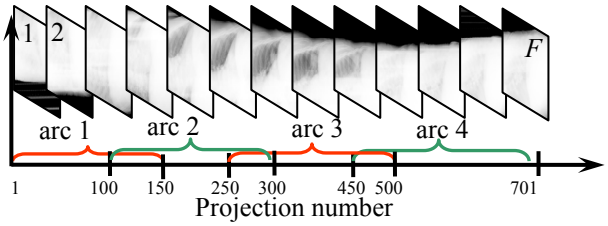


Fig. 4. Grouping consecutive projection images into arcs
Patient # 1 dataset images grouped into four arcs. Arc #1: 100-150, Arc #2: 100- 300, Arc #3: 250-500, Arc #4: 450-701. The process is done manually.

Step 2: The 3D motion of the lung is detected in each arc using the selected trajectories T_b clustered in Section 2.2.2. Fig. 5 shows the coordinate systems used in this study. This figure shows one 2D point tracking in three consecutive projection frames. The x coordinates displacement Δx corresponds to the displacement caused by the orbital motion. Δy corresponds to the cranio-caudal (up-down) position of the lung. To detect the 3D motion of the lung, structure-from-motion technique is used [92]. The 2D points $p_{f,p}=(x_{f,p},y_{f,p})$, of the set of trajectories T_b, are filled in the measurement matrix $W:2F\times P$ as in (5). Then, the mean m_f and n_f for each measurement type is subtracted off from W to yield the registered measurement matrix W^* as in (6):

$$W=\begin{bmatrix} x_{1,1} & \cdots & x_{1,P} \\ \vdots & \ddots & \vdots \\ x_{F,1} & \cdots & x_{F,P} \\ \hline y_{1,1} & \cdots & y_{1,P} \\ \vdots & \ddots & \vdots \\ y_{F,1} & \cdots & y_{F,P} \end{bmatrix}=\begin{bmatrix} x_{f,p} \\ \hline y_{f,p} \end{bmatrix} \quad (5), \qquad W^*=\begin{bmatrix} x_{f,p} \\ \hline y_{f,p} \end{bmatrix}-\begin{bmatrix} m_f \\ \hline n_f \end{bmatrix}. \quad (6)$$

To recover the 3D motion, the registered measurement matrix W^* is decomposed by (SVD) into $W^*=U'D'V'^T$. The respective motion matrix M' is extracted as $M'=U'\sqrt{D'}$ and the true value of the motion matrix $M:2F\times 3$ is recovered as $M=M'\Lambda$, where Λ is a

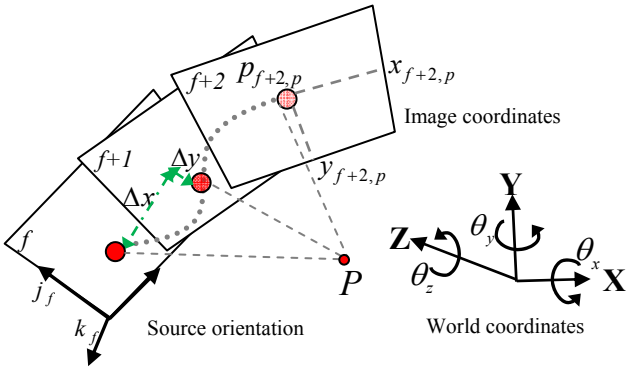


Fig. 5. The coordinate systems used in this study
One feature point tracked in three projections *f*, *f+1,* and *f+2*. Real world frame is represented by *XYZ* coordinates and CBCT projection frame is represented by *xy*. Source orientation is represented by the orientation vectors i_f, j_f, and k_f.

3×3 matrix that satisfies the three metrics constrains: $|i_f|^2 = |j_f|^2$, $i_f . j_f = 0$, and $i_1 = 1$. The motion matrix is defined as: $M = [i_1 \ \ldots \ i_1 \ | \ j_1 \ \ldots \ j_F]^T$.

Step 3: The 3D rotation angles θ_x, θ_y, and θ_z at every projection image from the motion matrix M are recovered. Entries of motion matrix M for each projection are defined to equal the first two rows of the arbitrary rotation matrix $R(\theta_z,\theta_y,\theta_x)$, where

$$R(\theta_z,\theta_y,\theta_x) = R(\theta_z).R(\theta_y).R(\theta_x)$$

$$= \begin{bmatrix} \cos\theta_Z & -\sin\theta_Z & 0 \\ \sin\theta_Z & \cos\theta_Z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta_Y & 0 & \sin\theta_Y \\ 0 & 1 & 0 \\ -\sin\theta_Y & 0 & \cos\theta_Y \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{bmatrix},$$

where $R(\theta_z)$, $R(\theta_y)$, and $R(\theta_x)$ correspond to the rotation matrices about the Z-axis, Y-axis, and X-axis, respectively.

So, $[i_f \quad j_f]^T = R_2(\theta_z,\theta_y,\theta_x)$

$$= \begin{bmatrix} \cos\theta_Z\cos\theta_Y & \cos\theta_Z\sin\theta_Y\sin\theta_x - \sin\theta_Z\cos\theta_x & \cos\theta_Z\sin\theta_Y\cos\theta_x + \sin\theta_Z\sin\theta_x \\ \sin\theta_Z\cos\theta_Y & \sin\theta_Z\sin\theta_Y\sin\theta_x + \cos\theta_Z\cos\theta_x & \sin\theta_Z\sin\theta_Y\cos\theta_x - \cos\theta_Z\sin\theta_x \end{bmatrix}$$

The rotation angle θ_y is interpreted as the gantry (orbital) motion of the X-ray source around the patient. θ_x is interpreted as the rotation around the X-axis which is not our

concern in this study. θ_z is interpreted as the respiration signal because the lung anatomy is seen by the X-ray source as rotating around Z-axis as shown in Fig. 5. Algorithm 3 summarizes of the steps used for respiratory motion detection.

Algorithm 3: Respiratory Motion Detection
Step 1: Group consecutive projections into overlapping arcs to overcome the problem of feature points' invisibility.
Step 2: Detect the 3D motion of the lung in each arc. The motion matrix found is: $M = [i_1 \quad \dots \quad i_1 \quad | \quad j_1 \quad \dots \quad j_F]^T$.
Step 3: Recover the 3D rotation angles θ_x, θ_y, and θ_z at every projection using M. The rotation angles are interpreted and θ_z is chosen to represent the respiratory signal.

2.2.4 PROJECTIONS PHASE SORTING

To sort the projections into phases, the respiratory signal θ_z is used as the input of the phase sorting method. The following steps show the process of breath phase sorting:

Step 1: The respiratory signal θ_z extracted is smoothed before phase sorting. Savitzky Golay smoothing filter [93] is used to clear out the noisy respiratory signal. The technique use a set of weighting coefficients (w_{-m}, $w_{-(m-1)}$, …, w_{m-1}, $w_{m)}$ to carry out the smoothing operation. The use of these weights C is equivalent to fitting the data to a polynomial. Thus, the smoothed data point $\tilde{\theta}_{z,f}$ at frame f is:

$$\tilde{\theta}_{Z,f} = \sum_{i=-m}^{m} w_i \theta_{Z,f+i} \Big/ \sum_{i=-m}^{m} w_i \quad . \tag{7}$$

Fig. 6 shows the respiratory signal smoothing and breath phase interpretation in one breath cycle.

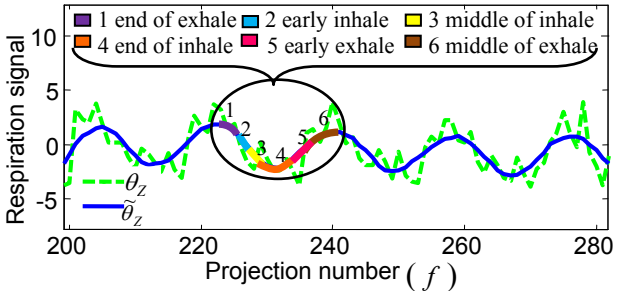


Fig. 6. Smoothing the respiratory signal and sorting the projections into phases.
The noisy (θ_z), smoothed respiratory signal ($\tilde{\theta}_z$) and phase sorting into six bins using the smoothed respiratory motion. The six colors in the oval correspond to each bin as shown on top of the figure.

Step 2: The smoothed respiratory signal $\tilde{\theta}_z$ is used in recovering breath phases. Let $H(f)$ denote the breath phase. Projections are sorted according to their breath phase by setting all peak projections to a phase of "1": $H(f)$=1. The remaining projections between the two peaks are assigned to the other number of phase bins wanted by taking the total number of projections divided by the number of phases, and then sort the projections so each bin contains roughly the same number of projections. Fig. 7 shows the breath phases extraction from the smoothed respiratory signal $\tilde{\theta}_Z$ in multiple breath cycles.

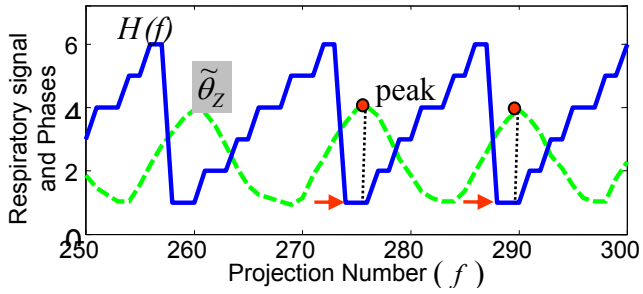


Fig. 7. Breath phase sorting of the projections into six bins based on the respiratory signal in multiple respiratory cycles.
Projections are sorted into phases based on the respiration signal $\tilde{\theta}_z$ (dashed green). Phases $H(f)$ (solid blue) are interpreted as: 1 as the end of exhale, 2 as early inhale, 3 as middle of inhale, 4 as the end of inhale, 5 as early exhale, and 6 as middle of exhale.

Step 3: Breath phase sorting quality is measured by comparing the sorted projections using LIFT and the ground truth signal. The phase shift, defined as $\Delta H(f)$, is measured between every corresponding phase "1" in both signals. The average phase shift is estimated for each arc and is defined as:

$$\overline{H} = \sum_{c=1}^{n_c} \Delta H(f) / n_c , \tag{8}$$

where n_c is the number of cycles. The standard deviation of phase shift is defined as $\sqrt{\sum_{c=1}^{n_c} (H(f) - \overline{H})^2 / n_c}$. Also, the breathing amplitude error is computed by dividing the average number of phase-shifted projections by the average number of projections in each cycle as follows: $\overline{H} / (F / n_c)$. (9)

Algorithm 4: Projection Sorting
Step 1: Smooth the noisy breath signal θ_Z using Savitzky Golay filter by fitting the θ_Z values to a polynomial as (7).
Step 2: Recover breath phases $H(f)$ from the smoothed respiratory signal $\tilde{\theta}_z$.
Step 3: Evaluate phase sorting by estimating the average phase shift as (8) and the breath amplitude error as (9).

2.3 Experimental Results

2.3.1 Dataset Specifications

Four datasets were used to validate LIFT. The first dataset used has been taken under the following characteristics: The imaging system used consists of a radiation source and a detector panel which orbit in the XZ plane around the fixed point in space (which is placed at the world coordinate system origin). The actual distance of the source to origin is always fixed 1000 mm, and the virtual distance is also 1000 mm in this setup. The detector-center-to-origin is also fixed 536 mm. The system is calibrated to provide a virtual image of known size at the origin by calibrating the physical pixel size. With this calibration, the virtual panel dimensions are 265.2×265.2 mm, the pixel size is 0.518 mm/pixel. Patient 2-4s' datasets are similar, but acquired on a different vendor's imaging system. The geometry is similar, but the virtual panel dimensions are 198.5×264.7 mm, and the pixel size is 0.258 mm/pixel. Patient 4's dataset showed some irregular breathing patterns, while patient 1-3's datasets have generally regular breathing patterns. The ground truth used for the datasets for patient 1 was the result of the diaphragm position-based method, while the ground truth for patient 2-4s' datasets was the internal markers trajectories. Four markers were used in patient 2's and patient 4's datasets and the averages of their trajectories through the projection images were estimated. One marker was used in patient's 3 dataset. Table 3 shows the specifications of the datasets used in this study.

Table 3: Dataset Specifications for Multiple Patient Datasets

Patient number	Number of projections	Projection size (pixel)	Pixel size (mm/pixel)	Source-Origin distance (mm)
1	701	512×512	0.518	1000
2	2396	768×1024	0.258	1000
3	2436	768×1024	0.258	1000
4	2300	768×1024	0.258	1000

2.3.2 FEATURE TRACKING AND TRAJECTORY SELECTION

Feature points were extracted in the first projection of the sequence on locations equally spaced by a specific number of pixels, 20 pixels in this experiment. Fig. 8 shows feature points tracking in Arc 4 of patient 1's dataset (between projection #250 and #500). The displacement shown is represented in pixels. In (a), yellow dots show the extracted feature points and green lines show the trajectories of those tracked points through the arc projections. The shape of the trajectories determines the motion of the area in which those trajectories reside. For example, trajectories residing in thoracic tissue regions tend to have a shape similar to a respiration curve, while trajectories residing in bony areas look like an orbital trajectory or unorganized shape. In (b), selected trajectories of the tracked points are shown.

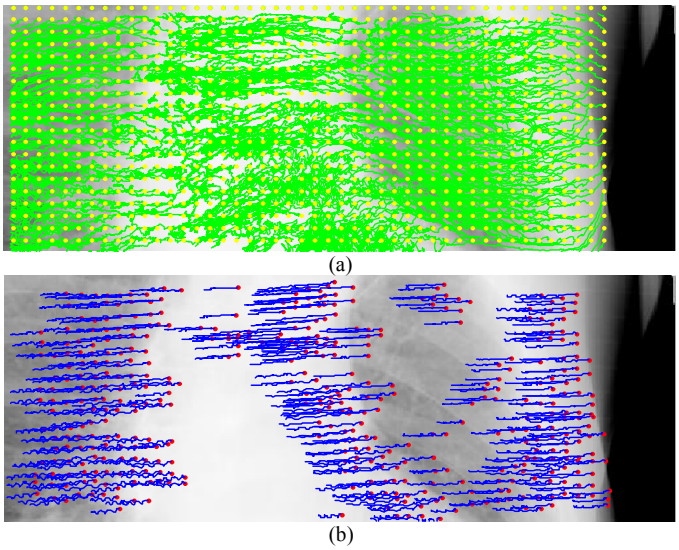
The average optical flow displacement was 0.51 pixels in patient 1's dataset, 0.37 pixels in patient 2's dataset, 0.23 pixels in patient 3's dataset, and 49.5 in patient 4's dataset. The displacement vectors estimated for patient 1's datasets in arc 4 have an average horizontal displacement larger than the vertical one. Those values were compared to Fig. 8 in which it appeared that the horizontal motion was larger than the vertical motion. Our evaluation of optical flow performance has been limited to a qualitative judgment as [94] due to the lack of true displacement vectors. Also, the results

of the 3D motion recovery, shown in Section 2.3.3, proved that the optical flow algorithm used performed well and produced good results.

As previously mentioned, trajectories showing cyclical superior-inferior motion representing respiratory motion were selected according to the criteria followed in (2). Table 4 presents the compactness and isolation measures (3) and (4) applied to the resulting clusters in four patient datasets. As shown, the average compactness of the selected trajectories in the four patient datasets was around 87%, and the average isolation was around 44%.

Table 4: Average Compactness and Isolation of the Chosen Cluster of Trajectories in Multiple Patient Datasets

Patient #	AVG Compactness (3)	AVG Isolation (4)
1	88.01 ± 10.19	43.20 ± 9.27
2	86.24 ± 12.53	47.53 ± 10.96
3	90.15 ± 9.84	45.14 ± 7.67
4	85.14 ± 9.54	42.52 ± 11.86



(a)

(b)

Fig. 8. Results of feature points selection and tracking
(a) Feature points detected in the first projection image of Arc 4 and trajectories through the arc. Yellow dots show the extracted pixels and green lines show the trajectories. (b) Selected trajectories that represent breathing in blue.

2.3.3 Respiratory Motion Detection and Phase Sorting

To model the 3D motion of the lung, structure-from-motion is used as described in Section 2.2.3. Fig. 9 shows the extracted respiratory signal in selected arcs. The respiratory signal is the rotational angle about the Z-axis measured in degrees. LIFT based respiratory signal was compared to the diaphragm position-based signal in (a) and to the markers-based signal in (b), (c), and (d). The X-axis of the figure corresponds to the projection number and the Y-axis corresponds to the degree of the rotation at each projection.

As shown in Fig. 9, LIFT respiratory motion matches the ground truth. To compare the accuracy of LIFT signal, the horizontal shift exiting between the peaks of the signals extracted using LIFT and the ground truth is taken into account. Smaller shifts

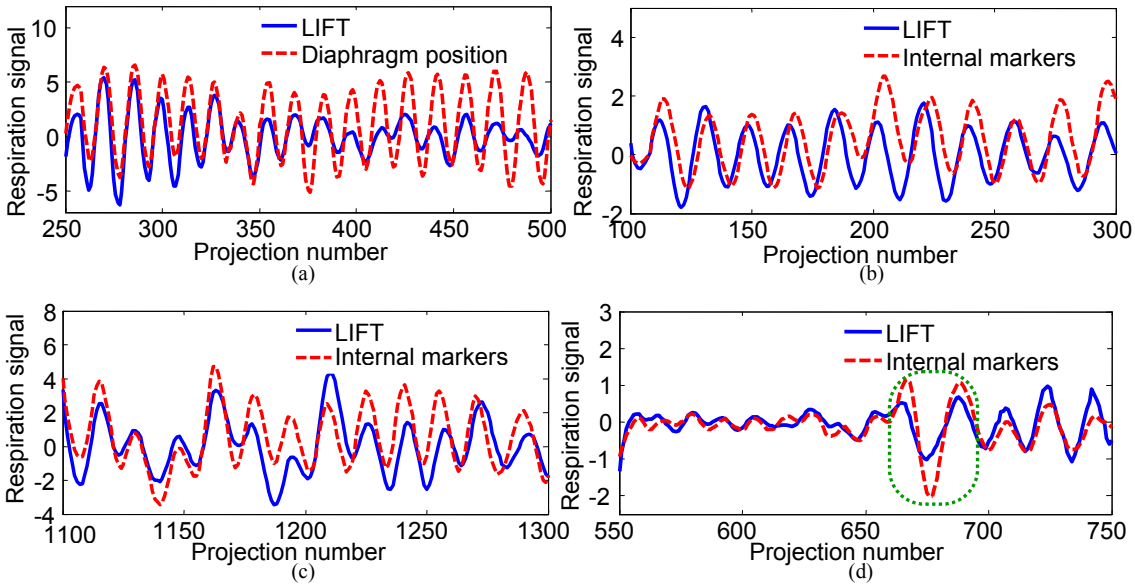


Fig. 9. Modeled respiratory motion in four clinical datasets
Datasets of (a) patient 1, (b) patient 2, (c) patient 3, and (d) patient 4 are used. Patient 4 had changes in period and amplitude of breathing in this portion of the breathing trace, and the green rectangle shows a particularly large change in amplitude. The respiratory motion is the modeled 3D rotation around the *Z* axis measured in degrees. Peaks in LIFT respiratory signal and other methods are compared.

imply stronger correlation between the two signals. The curves are scaled to have similar amplitudes. The difference in the height or the location of the curves in the plot is not important, as this information is not used to phase-sort the projections. To calculate the phase shift, projections are sorted into breath phases according to the respiration signal as described in Section 2.2.4.

Patient 1-3's datasets generally represented regular breathing. LIFT is be able to recover, or at least identify, irregular breathing patterns, as it is able to recover the actual 3D motion through the sequence in any form, given accurate 2D trajectories. Similar methods have been used to recover free-form 3D motion [8]. In patient 4's dataset, an irregular portion of the breathing trace was detected using LIFT as shown in Fig. 9 (d). A deep breath around projection 680 was detected. This breath cycle is surrounded by a green rounded rectangle.

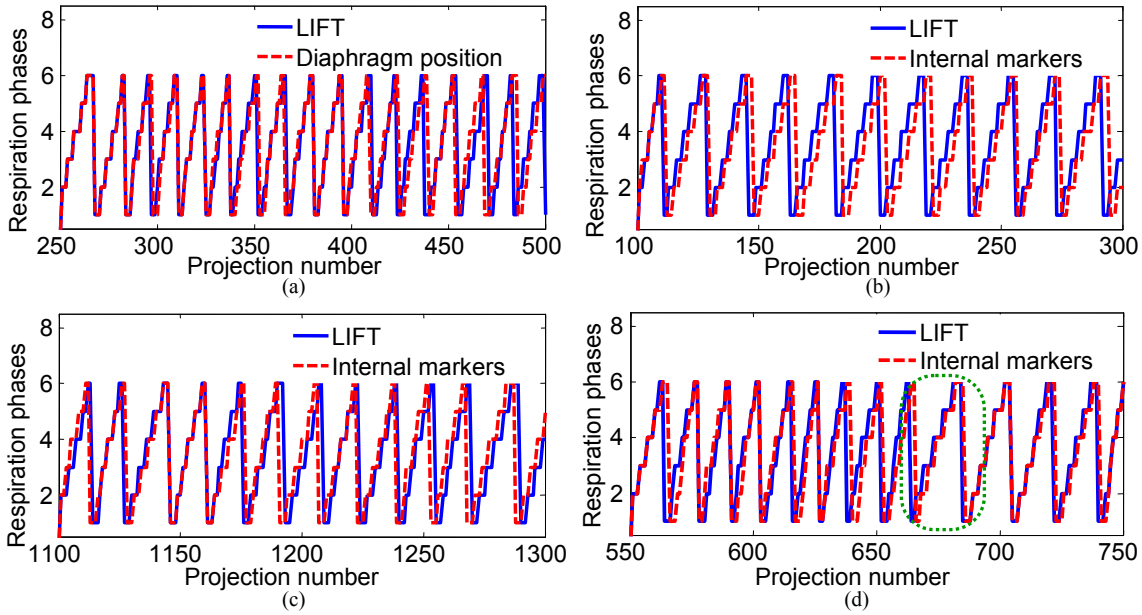


Fig. 10. Extracted respiration phases in four datasets
Datasets of (a) patient 1, (b) patient 2, (c) patient 3, and (d) patient 4 are used. The number of bins used is 6. The green rounded rectangle in (d) corresponds to a cycle with large change in breathing amplitude. The X-axis is the projection number and the Y-axis is respiration phases range from 1 to 6. Respiration phases match when the respiration signal matches.

Fig. 10 shows respiratory phase sorting in the four patients' datasets. As shown, sorted projections using LIFT signal match the sorted projections using the ground truth signal especially in (a), (c) and (d). The average phase shift using LIFT and ground truth was quantitatively estimated as described in Section 2.2.4.

Table 5 shows the quantitative accuracy of the breath phase sorting of the respiratory signal extracted of patient 1 using LIFT compared to the diaphragm position-based signal. The average and standard deviation of phase shift is used as described in (8). The phase shift between LIFT based signal and the diaphragm-based signal was big in the first arc compared to the other arcs. This was because projection images in this arc have high contrast and the respiration signal extracted was less accurate compared to the other arcs.

Table 5: Error in Breath Phase Sorting For Patient 1

Arc #	Frame#	AVG $\pm$ STD phase shift (8)
1	1-150	2.60 ± 0.84
2	100-300	1.29 ± 0.76
3	250-500	1.00 ± 0.82
4	450-701	1.86 ± 1.92

Table 6 shows the quantitative measurement of phase shift between LIFT respiratory signal and the ground truth signal in all arcs for the four datasets. The criteria used for error measurement is discussed in Section 2.2.4. The average and standard deviation of phase shift is used as in (8) and the average breathing amplitude error as in (9). As shown, the average phase shift measured in respiratory motion was around 1.68 compared to the diaphragm-based signal estimated in patient 1's dataset. Comparing to the internal markers-based signal, the average phase shift was 3.04 projections in patient 2, 1.25 in patient 3 and 1.14 in patient 4's datasets. As the average breathing amplitude error was estimated using the average phase shift, patient 3's case had the smallest error of 7.19% as an average, while it is 15.96% in patient 2's case. This shows that the computed signal using LIFT correlates to the diaphragm position-based signal in patient 1's case and to the internal marker's signal in patient 3's and patient 4's cases more than patient 2's case, if LIFT is performed on the entire projection. One of the reasons behind this is that the respiration signals extracted using markers may depend on the location of the markers. The respiration signal derived using LIFT is the 3D rotation, about the Z-axis, of the lung tissues appearing in the raw CBCT projections. The locations of the feature points used in extracting the respiration signal affect the respiration signal extracted. Feature points in the lower part of the lung close to the Diaphragm showed a strong oscillating motion more than the feature points in the upper part or edges of the

lung. As LIFT finds a global 3D motion of the flow intensity in the entire sequence, the effect of the stronger motion in the lower part of the lung dominates.

Table 6: Average Error in Breath Phase Sorting in Multiple Datasets

Patient #	AVG phase shift (8)	AVG Breathing amplitude error (9)
1	1.68 ± 1.09	11.20 ± 7.27 %
2	3.04 ± 1.52	15.96 ± 7.98 %
3	1.25 ± 0.83	7.19 ± 4.77 %
4	1.14 ± 1.05	8.35 ± 7.96 %

In patient 2's dataset, markers exist in the middle of the lung close to the bronchioles. The average phase shift between LIFT respiratory signal and markers' signal was 3.04 ± 1.52 projections. When the signal was extracted from a Region of Interest (ROI) surrounding the markers position, the phase shift was 1.6±1.9 projections which is less than the phase shift when extracting the respiratory motion from the entire projection. These results demonstrate that there were phase differences between different parts of the lung, and a single respiratory signal may not be optimal to completely characterize breathing motion. One advantage of LIFT is that it can be applied to a custom ROI surrounding the part of the lung containing the tumor, allowing more accurate characterization of the respiratory signal for this region.

Also, LIFT was applied on the inferior portion of the CBCT images that include the diaphragm. The average phase shift between the respiration signal using LIFT and the diaphragm position based method was 1.1 ±0.57 projections in arc 1 of patient 1's dataset, which was less than the phase shift excluding the diaphragm area (2.60 ± 0.84 projections). This result demonstrates phase shifts in the anatomy at different superior/inferior locations in the lung, which implies breathing phase obtained at the diaphragm may not be appropriate to represent phase at other locations.

Fig. 11 shows the reconstructed 4DCBCT images for patient 2 using the FDK algorithm [42]. Projections that were used for reconstruction are the ones sorted in the end of inhalation phase using the respiratory signal extracted by the implanted markers in (a1)-(a2), and using the respiratory signal extracted by LIFT in (b1)-(b2). The difference of the axial images (a1) and (b1) is shown in (c1) and difference of the coronal images (a2) and (b2) is shown in (c2). As shown in (c1) and (c2), most of the differences were not anatomical, but rather due to differences in the streaking (view-aliasing) artifact due to too few projections. Since each sorting algorithm selects slightly different projections for reconstruction, the streaks appear in different locations in the two images.

2.3.4 Improved Reliability and Applicability

Here the reliability and applicability of LIFT comparing to the other standard methods is discussed. Table 7 shows the acquired data ratio using the diaphragm-based method [3] [4] the implanted markers method [6] [7] and LIFT. The acquired data ratio is

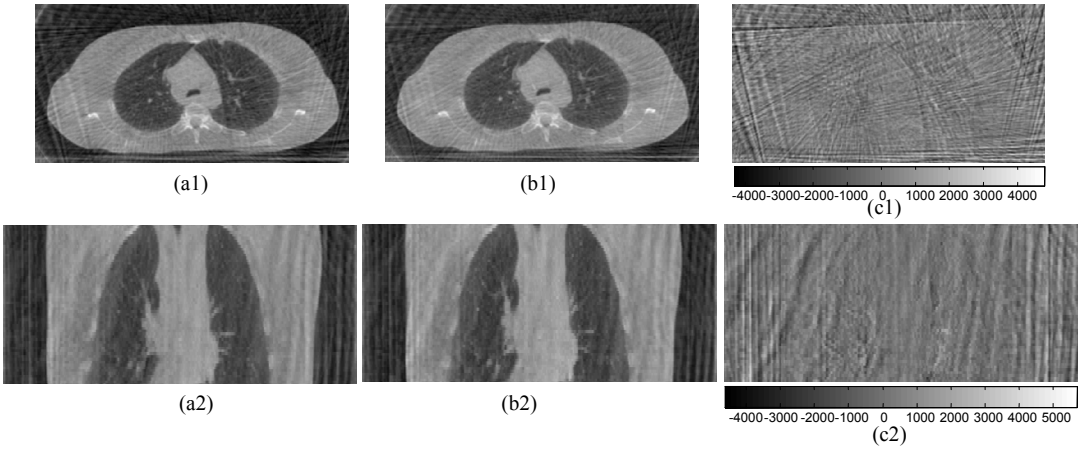


Fig. 11. Reconstructed 4DCBCT images of sorted in the end of inhalation phase
Reconstruction used the respiratory signal extracted using (a1)-(a2) implanted markers, (b1)-(b2) LIFT of patient 2 dataset. The difference of the reconstructed images is taken in (c1) between the axial images (a1) and (b1), and in (c2) between the coronal images (a2) and (b2).

Table 7: Acquired Data Ratio in LIFT Compared to Other Existing methods

Method	Samples	Acquired data ratio (%)
Diaphragm Position [8],[11]	Patient 1	100
Internal Markers [22], [23]	Patient 2: Marker 1, 2	100
	Marker 3	72.9
	Marker 4	70.0
	Patient 3: Marker 1	49.9
	Patient 4: Marker 1, 2	100
	Marker 3	68.8
	Marker 4	0
Proposed LIFT	Patients 1-4	100

the number of projections of the dataset from which the method was able to extract the respiratory signal, divided by the total of number of projections.

As shown in Table 7, the diaphragm-based method applied to patient 1's dataset was able to extract the respiratory signal in all projections. For the internal markers method, some of markers were not visible in some of the dataset projections. In patient 3's case, only one marker was used and was visible in around 50% of the projections, which prevented the respiration signal from being extracted in the rest 50% of the dataset projections. When having multiple markers implanted, visible markers' traces can be used as replacements to the invisible ones such as in the cases of patients 2 and 4. As LIFT extracts the respiratory motion from the organ tissues of the lung in any CBCT projection, it was able to generate a respiratory signal in 100% of the projection sets of the four patients. These results showed that LIFT was more reliable than the internal marker's method in terms of acquiring the respiration signal from projections.

The average time for running the entire LIFT procedure is around 26 minutes on a set of 1000 projections (of approximately 550×650 pixels as the analysis region) using a PC of Intel Core 2 Duo 2.4 GHz CPU and 2 GB of RAM. Running speed can be improved if LIFT is implemented on GPU with C++ platform or on a multiprocessor

computer using Matlab parallel processing. For implementation purposes, LIFT follows feature tracking and 3D reconstruction approach which is well-known and easy to implement. There are many freely-available, optimized implementations that a developer can start from, such as OpenCV for optical flow computation [95]. For 3D motion reconstruction, Tomasi and Kanade method [92] is used which is a mathematical approach consisting of well described series of linear algebra computations. Agglomerative hierarchical clustering is described in Mathworks [96]. Thus, LIFT can be implemented and applied in the clinic to extract the respiratory signal from CBCT projections.

2.4 SUMMARY

A novel method for respiratory motion extraction and breath phase sorting using CBCT projections was proposed. Modeling of the respiratory motion of the patient's lung was accomplished on multiple arcs. On each arc, feature points were extracted and tracked to find point trajectories. Trajectories with shapes similar to breathing curve were selected to be used in the 3D motion modeling module to recover the 3D motion of the lung. The 3D rotation around the Z-axis of the patient represented the respiratory motion and the CBCT projections were then sorted according to the respiration signal.

Experimental results were conducted on datasets exhibiting regular and irregular breathing. The respiratory motion extracted using LIFT was compared to the ones extracted using other standard methods. An average phase shift of 1.78 projections was estimated between LIFT based signal and markers based signal, and of 1.68 projections between LIFT based signal and the diaphragm-based signal. The average breathing amplitude error of LIFT compared to the diaphragm-based method was 11.2% while it is 10.68% compared to the internal markers method. LIFT was able to extract the respiration signal in all projections of all datasets without the dependence on a particular anatomical structure (such as the diaphragm).

CHAPTER 3 IMAGE-BASED RESPIRATORY SIGNAL EXTRACTION USING INTENSITY FLOW DIMENSIONALITY REDUCTION IN CONE BEAM CT PROJECTIONS

This chapter presents an algorithm that detects an image-based respiratory signal automatically in each Cone Beam Computed Tomography (CBCT) projection for respiratory phase sorting. The proposed method, termed Intensity Flow Dimensionality Reduction (IFDR), has been developed and experimentally validated on clinical datasets.

3.1 INTRODUCTION

Respiratory motion detection from CT scan images is important because it can be a major source of error in determining the position of thoracic and upper abdominal tumor targets and critical normal tissues during radiotherapy planning and delivery. Four-dimensional or respiration-correlated CT (4DCT) imaging techniques have become the mainstay for imaging of respiratory-influenced tissues for guiding radiation therapy planning. Volumetric image guidance techniques, such as cone beam CT (CBCT), for verifying tumor position during treatment have also been enhanced with 4D techniques to account for respiration-induced tissue motion [1] [3] [34]. Current methods to reconstruct 4DCT images retrospectively require a one-dimensional signal related to respiration ("respiratory signal") to sort each 4D image or projection into a respiratory bin.

The respiratory signal serving as a surrogate of the tumor position is acquired in temporal synchrony with the raw cone beam CT imaging data acquisition. The surrogate is used to assign each projection to the appropriate breathing phase bin in a process termed 'sorting'. Respiratory signals are currently generated from external sources (skin markers, abdominal belts, or spirometry) [1] [2] [97] or by an image-based measure of diaphragm position [3] [4] [5]. The disadvantage with using an external signal is not only the requirement of additional hardware, but also the external signal and tumor motion trajectories are often phase-shifted in time, with the phase shift varying from cycle to cycle and day to day [17] [18] [19] [20] [21] [22] [23] [24] [25] [26]. Projection sorting based on diaphragm position may have similar issues [20], and in some vendors' systems the limited longitudinal field of view may prohibit use of this method in the middle and upper lung. Another option is to use an implanted fiducial marker near the tumor to aid in

extracting the respiratory motion [6], [7]. This solves the problem, but at the expense of an additional invasive and often morbid procedure which could significantly delay treatment initiation.

In this chapter, we propose a novel Intensity Flow Dimensionality Reduction (IFDR) method; an image-based respiration signal extraction method based on the organ tissues motion local to the tumor. This method has the advantages over the previous mentioned methods in that neither external equipment nor internal markers are required, and visibility of high-contrast objects such as the diaphragm is not necessary. Moreover, this method is fully automatic in that no prior information about the anatomy or training data are required. Since the signal can be extracted from a pre-defined region near the tumor, the respiration signal should ideally correlate better with the tumor position.

IFDR detects the respiration signal by computing a dense optical flow on every pixel of each pair of adjacent CBCT projection images of the patient dataset. Since we know a pattern of a respiration motion exists in the optical flow displacement dataset, we apply linear and non-linear dimensionality reduction techniques to the consecutive optical flow displacement vectors detected to extract this respiratory motion. The principal components of the eigen system resulting from the linear dimensionality reduction method are used to represent the respiratory signal. Similarly, the dimensions of the reduced-dimension dataset resulting from the non-linear dimensionality reduction method serve as the respiratory signal of the lung. Both linear and non-linear dimensionality reduction methods are applied to the clinical datasets and a comparison between them is accomplished.

Previous work using optical flow analysis of cone beam CT projections has guided our work. Bergner *et al.* [27] developed an optical flow-based algorithm for measuring motion between 4DCBCT projections using a Horn-Schunk implementation to obtain the dense optical flow (of each pixel in each projection) [31]. The purpose of their algorithm was to improve reconstruction quality in stationary regions of the anatomy. Optical flow was used to classify regions of pixels as moving or stationary, but they did not explicitly use the dense optical flow to measure a respiratory signal. For respiration sorting, they used an image-based measure of diaphragm position similar to Sonke *et al.* method [3]. Dimensionality reduction methods, such as manifold learning [32], have been used for image-based respiratory signal detection; however, these techniques were applied only for fixed view (fluoroscopic) images, as opposed to the rotational view (tomographic) projections used here. The works [4] [28] [29] [30] are also related. In the work [4], thousands of interest points are selected in the projection images. These points are tracked across projections using a Block Matching Algorithm. Signal processing techniques are then used to acquire a breath signal. However, the results showed that the accuracy of the sorting of projections depends on the number of desired phase bins. In [28], Siochi developed a technique that locates a bounding box for the diaphragm motion for all projections based on two pairs of full-inhale and full-exhale views. However, this technique requires the diaphragm to be present in the field of view. In [29], the breathing pattern is acquired from analyzing the variation in pixel values between projection images by developing a simple pixel value summation followed by a high pass filtering. This method is powerful, but its efficiency should be demonstrated for various CBCT patient cases and compared to other methods in the literature. Vergalasova *et al.* [30]

proposed a markerless method for respiration signal detection based on the principles on the Fourier Transform theory. However, when they use the entire projections of patient data, their results showed a big phase shift between their extracted signal and the ground truth signal.

3.2 PROPOSED METHODS

3.2.1 INTENSITY FLOW DIMENSIONALITY REDUCTION (IFDR) ALGORITHM

The IFDR algorithm developed here automatically extracts a breathing signal that directly corresponds to the motion of the internal tissues of the lung. This method estimates the optical flow motion of objects in the projections, near the target volumes. Then, dimensionality reduction techniques are used to extract the motion patterns in the optical flow displacement vectors. The reduced dimensional datasets represent the respiration signal of the patient. Fig. 12 shows the modules that work to estimate the respiratory motion of the patient. In the motion estimation model, the only input required

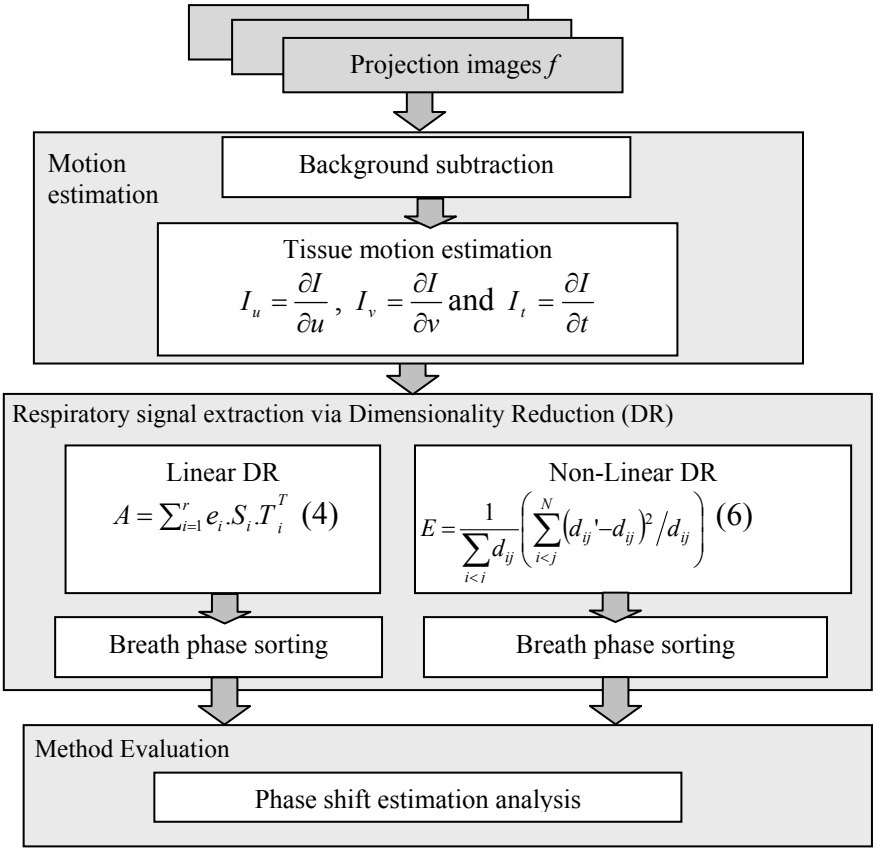


Fig. 12. Schematic of the Intensity Flow Dimensionality Reduction (IFDR)

is the CBCT projection images. The motion of the lung internal tissues is estimated using a dense optical flow estimation of every pixel in each pair of the projection images. The dense optical flow displacement dataset is used by the motion analysis model to extract the respiration signal and sort projections according to their respiration state. This model uses linear and non-linear dimensionality reduction methods to detect the motion patterns in the optical flow dataset which serve as a respiration signal. The respiration signal extracted is used to sort projections into numbered bins by breathing phase in a process called "phase sorting". In the methods evaluation model, IFDR algorithm is experimentally evaluated and compared to other algorithms in the literature.

3.2.2 Algorithm Details

Soft Tissue Motion Estimation

IFDR follows a tracking and motion analysis process. The optical flow of the moving tissues was extracted for every pixel in the CBCT scan projections. Tissue motion from one view to another is estimated through the whole dataset. As shown in Fig. 12, the motion analysis model consists of two phases. The first phase is the background subtraction phase. We subtract the background to eliminate the effect of the noisy motion existing in the background intensity. We use thresholding to subtract the background by setting the pixel intensity value to 0 if it is less than a threshold *T*. The second phase used is tissue motion estimation. We use dense optical flow to find the displacement of the intensity in each pair of adjacent projection images. We used the original dense optical flow algorithm proposed by Horn-Schunck [31] with improvements by Sun, Roth, and Black [98] in which the accuracy of the "classical" OF such as median filtering of intermediate flow fields and optimization has been improved.

Let *I(u,v,t)* represent the brightness of a pixel at *(u,v)* at time *t*. The respiratory motion estimated on images using optical flow is derived from the continuity equation:

$$\frac{\partial I}{\partial u}\delta u + \frac{\partial I}{\partial v}\delta v + \frac{\partial I}{\partial t}\delta t, \qquad (10)$$

which results in

$$I_u U + I_v V + I_t = 0,$$

where U and V are the horizontal and vertical components of velocity or optical flow of $I(u,v,t)$ respectively, and $I_u = \frac{\partial I}{\partial u}$, $I_v = \frac{\partial I}{\partial v}$ and $I_t = \frac{\partial I}{\partial t}$ are the derivatives of the image at (u,v,t) in the corresponding directions. A constrained minimization problem can be formulated to calculate optical flow vector for the frame at *t+1*.

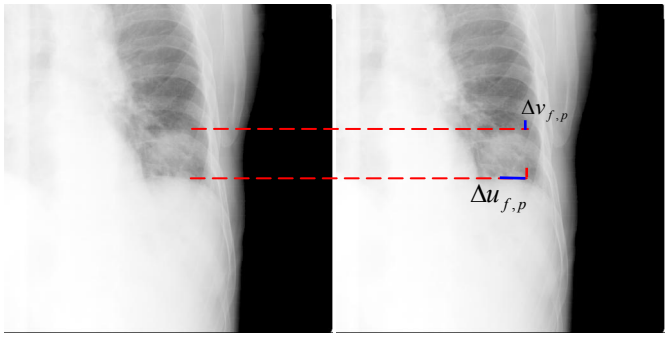
Using perspective projection in cone beam CT, mapping between voxel (x,y,z) and the projections is described by:

$$u = \frac{x.\lambda}{z}, v = \frac{y.\lambda}{z}, w = \lambda, \quad (11),$$

where λ is the focal length which is fixed.

Dense optical flow is computed between each pair of successive frames in the entire dataset. The displacement between two corresponding points p_f in projections *f* and p_{f+1} in projection *f+1* is defined as $(\Delta u_{f,p}, \Delta v_{f,p})$, so the 2D point of the next projection frame *f+1* can be written as: $(u_{f+1,p}, v_{f+1,p}) = (u_{f,p}, v_{f,p}) + (\Delta u_{f,p}, \Delta v_{f,p})$. Our algorithm finds the optical flow displacement vectors in all pixels and does not discriminate between pixels that show a respiratory motion through projections (e.g., tumor features, bronchioles, or bronchiole vessels) and static pixels (e.g., bony regions, background regions). Static pixels show only orbital movement (due to the gantry rotation of the

system), while respiratory pixels are shown having sinusoidal component of motion. Fig. 13 shows two projections from a patient's dataset that shows a shift in the tumor position during CT scan image acquisition. This figure shows the displacement vector for one pixel $(\Delta u_{f,p}, \Delta v_{f,p})$ in a pair of CBCT images.



(a) (b)

Fig. 13. Two consecutive projections from a patient CBCT scan showing a slight difference in the tumor position
This difference is due to the patient breathing during the CT scan image acquisition. The circle shows the tumor and the illustrative horizontal red lines shows the change of the position of the tumor in both images: (a) Projection #408 in Patient 1 dataset, (b) Projection #409 in Patient 1 dataset.

Respiratory Signal Detection using Linear Dimensionality Reduction

The optical flow step results in a set of 2D displacement vectors for each pixel in a projection. This is a large dataset for each CBCT scan, which must be distilled into a 1D signal for each projection. In the respiratory signal detection phase, the respiratory signal is extracted and analyzed from this set of trajectories. Pixels containing tissues of the lung have motion related to respiration, while others (e.g., bony regions) are stationary and only exhibit the rotational trajectory of the scanner. Because of this fact, we aim to discover the motion patterns of the moving tissues. So, we use the dimensionality reduction approaches to discover motion patterns related to respiration

that exist in the dataset. First, a linear dimensionality reduction approach using Principal Component Analysis (PCA) is used as follows:

Optical flow displacement components $(\Delta u_{f,p}, \Delta v_{f,p})$ are written in data matrix $A : F \times 2P$ so that the optical flow components of each pixel exist in one row. F is the number of projections and P is the number of points.

$$A = \begin{bmatrix} \Delta u_{1,1} & \Delta v_{1,1} & \cdots & \Delta u_{1,P} & \Delta v_{1,P} \\ \vdots & \vdots & & \vdots & \vdots \\ \Delta u_{F,1} & \Delta v_{F,!} & \cdots & \Delta u_{F,P} & \Delta v_{F,P} \end{bmatrix} , \qquad (12)$$

We obtain the eigensystem of the displacement data and extract the Principal Components of this eigensystem by using the Singular Value Decomposition (SVD) of the optical flow displacement vectors matrix *A* as described:

$$A = \sum_{i=1}^{r} e_i . S_i . T_i^T \qquad (13)$$

where *r* is the rank of *A*. SVD decomposes the matrix A into two sets of eigenvectors; S_1, S_2,…, S_i defined in space and T_1, T_2,…, T_i defined in time *i.e.* the equivalent direction as the number of projections. The eigenvalues e_1, e_2, …, e_i are sorted in decreasing order and so are their associated eigenvectors. Eigenvectors are assumed to be scaled to unit length. The principal eigenvector specifies the main changes in the values of the optical flow displacement vectors which may indicate the motion patterns in the optical flow components. The eigenvectors associated with the smallest eigenvalues may indicate the motion patterns with less significance. The magnitude of the eigenvalues may indicate a quantitative measure of optical flow values variation along the directions of the eigenvectors. So, the lung motion state can be approximated by a linear combination of eigenvectors that correspond to the first *i* largest eigenvalues as follows:

$$A \approx \sum_{i=1}^{I} e_i . S_i . T_i^T \quad (14)$$

The first few principal components are analyzed for signatures of the respiratory motion. Because a respiration motion signal will have a quasi-periodic and oscillatory nature, we can represent it by one of the principle components that correspond to one of the largest eigenvalues.

Respiratory Signal Detection using Non-Linear Dimensionality Reduction

While a linear method such as PCA maximizes the amount of the original variance present in the transformed dataset, it does not (in general) preserve "complex" structures [99]. Formally, for PCA to work properly, the *N*-dimensional space formed by the variables must be linear. Consider an *N*-dimensional space formed by all pixels in a projection as the variables in this space. Each projection image then represents a point in this *N*-dimensional space. A linear space would require all points to be valid, which cannot be true otherwise this would produce random images of noise and of all other possible images. Each set of projections should generate a reduced dimension manifold in the *N*-dimensional space, consisting of only points that represent images of only the patient. We use non-linear mappings between the original space and the reduced one to be able to describe the data with greater accuracy and/or by fewer factors than linear mapping such as PCA, given that there are sufficient data to support the formulation of this more complex mapping function.

To reduce the dimensionality of datasets non-linearly, multi-dimensional scaling is used [100]. A new lower-dimensional dataset is constructed which has structure that is as similar to the original dataset as possible. We map the original data matrix $A : F \times 2P$

defined in (12) into a lower N-dimensional space $A': F \times N$ as N is the number of dimensions in the desired space.

$$A' = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ \vdots & \vdots & \vdots & \vdots \\ a_{F,1} & a_{F,2} & \cdots & a_{F,N} \end{bmatrix},$$

where $a_{1,1}, \ldots, a_{F,1}$ are the first dimension and $a_{1,N}, \ldots, a_{F,N}$ is the last dimension in the reduced space.

The matrix $B=AA^T$ is computed. This matrix contains distance information. It is different than the covariance matrix $C=A^TA$. The matrix B is factorized into eigenvectors to yield the reduced dimensionality data. The N largest eigenvalues give the N-dimensional representation of the dataset A.

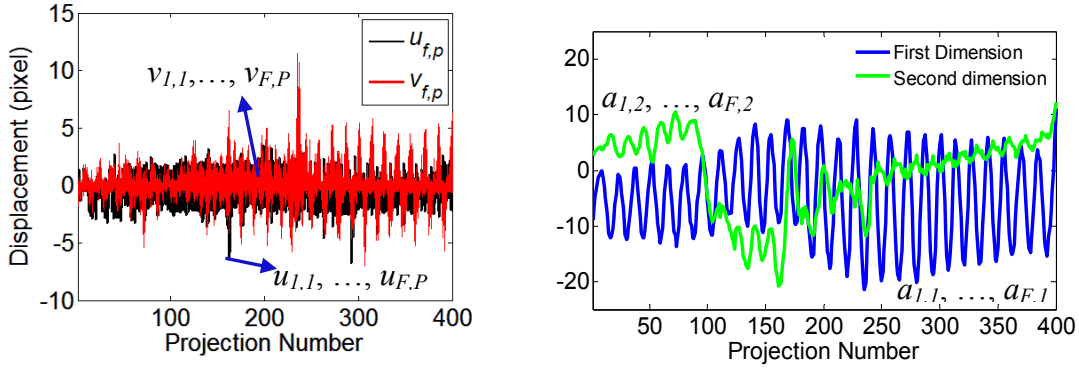
We start from an initial matrix A' filled with random values. The distance between two points in the original matrix A is denoted by d_{ij}, while the distance between two points in A' is denoted by d_{ij}'. Let E be the mapping error, which represents how well the present configuration of the points in the original dataset A fits the points in the N-dimension matrix A'. The error E represents the amount of structure present in the original dataset but lost in the transformed one. E is defined as follows according to the non-linear mapping by Sammon [100]:

$$E = \frac{1}{\sum_{i<j} d_{ij}} \left(\sum_{i<j}^{N} \left(d_{ij}' - d_{ij} \right)^2 / d_{ij} \right) \qquad (15)$$

The error E is calculated by summing up the squared differences (before versus after mapping) in pairwise distances between points. The summations are over the range $i < j$ so that each pairwise distance is counted once (i and j are not swapped). We used Euclidean distance. The existence of the factor of d_{ij} in the denominator of the main

summation ensures that if the original distance between two points is small, then the weighting given to their squared difference is greater.

Each time we find new values for d that minimize E. This is repeated until convergence. This will give this method its tendency to preserve topology of the dataset. The points in A' iteratively are adjusted to change the reduced space configuration and then to decrease the error. Fig. 14 shows the original space and the reduced one.



(a) (b)

Fig. 14. Original dataset and reduced dimensionality data using non-linear reduction.
(a)The original multi-dimension dataset A with displacement vectors has been reduced to (b) two-dimensional dataset A'.

Phase sorting based on Respiratory Signal

Here, we describe how the respiratory signal is extracted using linear and non-linear mapping methods. We sort the output respiratory signal from both methods individually for comparison purpose. This output signal R' may contain some noise. In order to get accurate breath phases for the respiratory motion, we first smooth the respiratory signal detected using Savitzky Golay smoothing filter [93]. Then, the smoothed respiratory signal R is used for breath phase extraction. The respiration motion phases are determined in a 'phase sorting' manner as follows: All peak projections are set to a phase of '1'. The remaining projections between two peaks are assigned to the other

phase bins by taking the total number of projections divided by the number of phases, and then sort the projections so each bin contains roughly the same number of projections. Fig. 5 shows the breath phases extracted from the smoothed respiratory signal.

3.2.3 CLINICAL EVALUATION

The motion analysis algorithm was evaluated on three clinical CBCT datasets, each from a different patient. Table 8 lists the imaging geometry for each clinical dataset. Each dataset had a gold standard respiratory signal, although the standard was different for the datasets. The gold standard used for patient 1 and 3 datasets is the result of the diaphragm-position based method, while the gold standard for patient 2 dataset is the position of implanted fiducial markers located in and near the tumor. Four markers were used and their trajectories through the projection images are estimated. For comparison purposes, we are using the average trajectory over the four markers. For experiments, we used MATLAB installed on 16 GB RAM Dell PC with Intel Core i7 CPU 3.07 GHz.

Table 8: Multiple Patient Datasets Specifications

Patient number	Number of projections	CBCT system	Projection size (pixel)	Pixel size at isocenter (mm/pixel)	Gold standard signal
1	701	XVI 3.5 (Elekta)	512x512	0.518	Diaphragm position
2	2396	OBI (Varian)	768x1024	0.258	Internal markers
3	2436	OBI (Varian)	768x1024	0.258	Diaphragm position

3.3 Experimental Results

3.3.1 Respiratory Signal Detection using the Proposed Dimensionality Reduction Methods

Fig. 15 shows the eigenvalues spectrum for the three patients datasets used in this study recovered using the linear method, with the eigenvalues sorted in descending order by variance. The eigenvectors associated with the largest eigenvalues are used in extracting the respiration motion in the dataset.

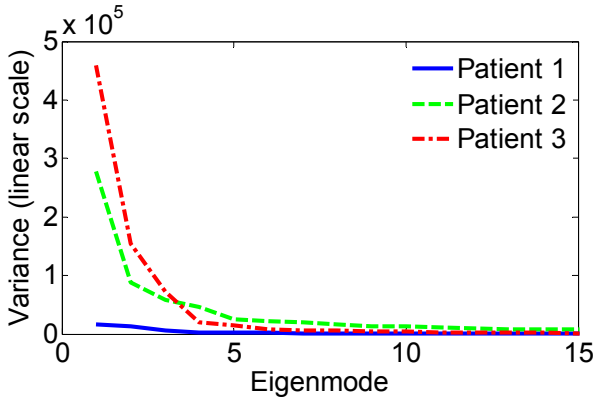


Fig. 15. Eigenvalues spectrum using linear dimensionality reduction method for the three patient datasets

Fig. 16 shows the stress value through the number of iterations in the three patients' datasets for the non-linear mapping. The stress value decreases when the number of iterations increases. As shown, the mapping converges after 50 iterations in the three datasets.

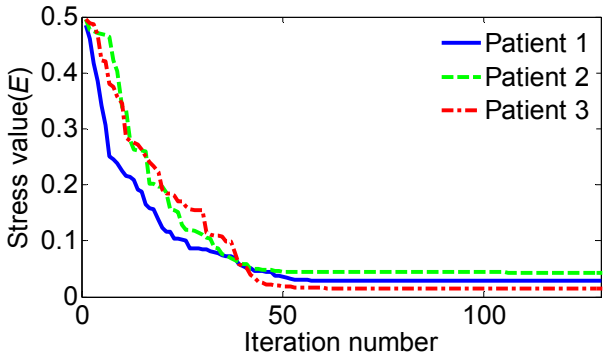
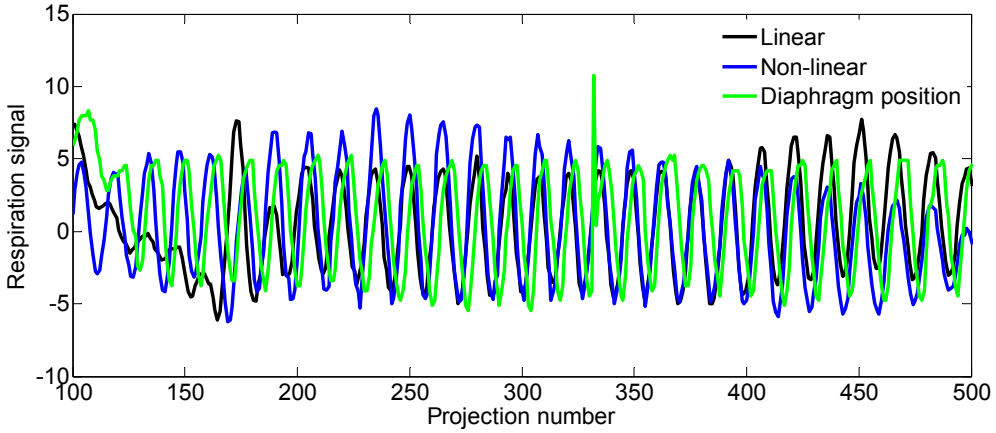
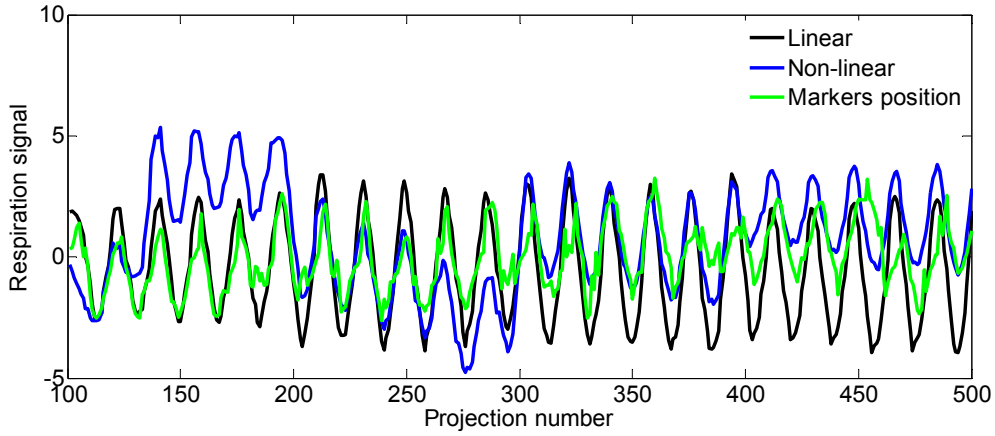


Fig. 16. Non-linear mapping stress value *E* vs. the number of iterations for three patient datasets.

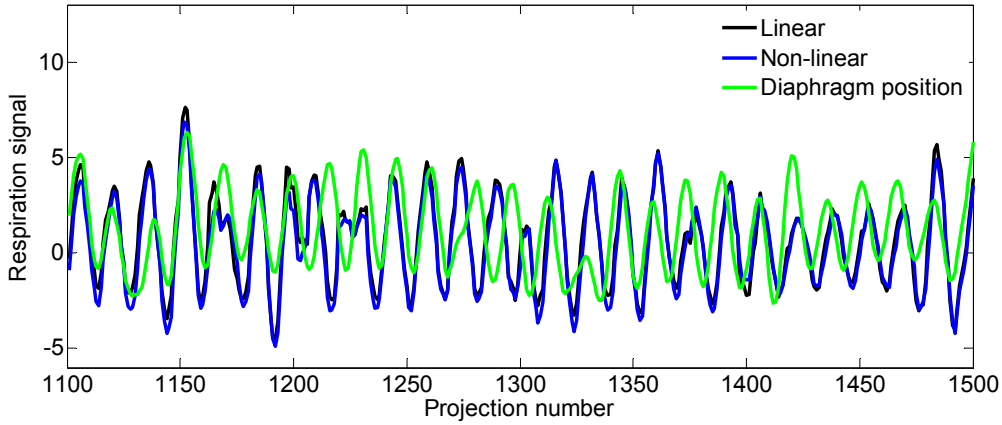
Fig. 17 shows the extracted signal using the linear and non-linear methods for each patient dataset compared to the gold standard. For the linear method, one of the first two principal components (coefficients) is used to represent the respiration signal. For the non-linear method, one of the first two dimensions of the reduced dataset is chosen to represent the respiratory signal. The extracted signal is compared to the respiratory signal detected using the diaphragm position based method in the datasets of Patient#1 and Patient#3 and to the signal extracted using the internal markers method in the dataset of Patient# 2. For Patient# 1 dataset, the second coefficient was chosen to represent the respiration signal, while in Patient#2 and Patient#3 datasets, the first coefficient is chosen. Using the non-linear method, the second dimension was used to represent the respiration signal for Patient#1 and Patient #3 datasets and the first dimension for Patient#2 dataset.



(a)



(b)



(c)

Fig. 17. Computed breath signal using Linear and non-linear DR method versus measured breath signal for (a) Patient 1, (b) Patient 2, and (c) Patient 3

As you can see in Fig. 17, the respiration signal extracted using the linear and the non-linear method correlate with the gold standard signal in Patient#2 which is based on the internal markers position. However, it has a small phase shift with the gold standard signal in Patient#1 and Patient#3 datasets which is based on the diaphragm position. This happens because the respiratory signal extracted using our method uses the motion of the organ tissues of the lung as its input, which is the same location of the implanted markers. However, the diaphragm-based method extracts the respiration signal depending on the diaphragm position which doesn't always correlate exactly to the internal tissues of the lung or the tumor position.

Fig. 18 is an illustrative figure showing selected projections from one breathing cycle extracted using Non-linear method. Those projections are associated with their location in the breath signal extracted. An auxiliary line exists to help the reader to observe the lung movement.

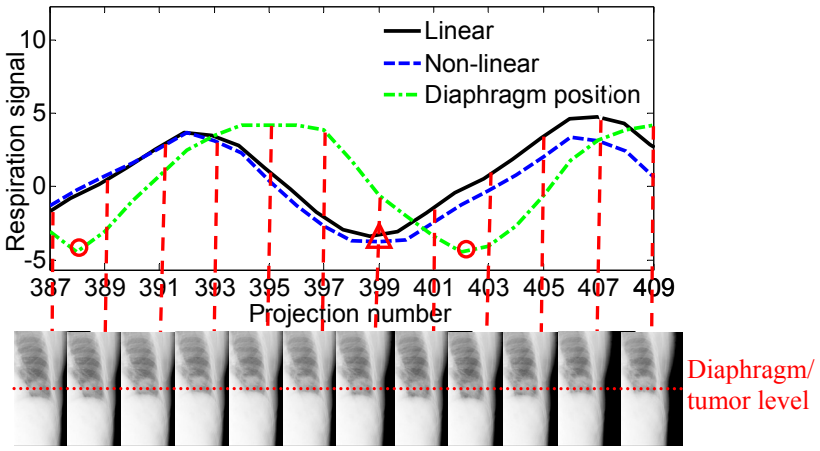


Fig. 18. Selected projections from one breathing cycle
Dashed vertical lines connect the projection images to their location in the signal. The horizontal auxiliary (diaphragm/tumor level) line on the projection images helps the reader to observe the lung movement. The red circle shows the end of inhalation in the diaphragm-position based signal and the red triangle shows the end of inhalation in the IFDR-based signal.

As shown in Fig. 18, projection #402 represents the end of inhalation in diaphragm position based respiration signal, while using the IFDR-based signal, this phase happens in projection #399 in both the linear and the nonlinear implementations which indicate a phase shift of 3 projections. The diaphragm is in its lowest position through the cycle in projection #402, while using IFDR the organ tissues of the lung, like the tumor, appear to have their lowest position in projection # 399.

3.3.2 Projection Phase Sorting

Table 9 shows the accuracy of the breath signal estimation and sorting in terms of overall average and standard deviation phase shifting between the computed signal and the other methods as gold standard. Average phase shift is calculated by sorting the projections into phases according to the respiration signal extracted using each compared method. Then, the difference in terms of the number of projections between phase '1' (end of exhale) in the corresponding breath cycles of the compared methods is computed. The average 'projection difference' is computed for all breathing cycles of each method as the average phase shift. Also, the breathing amplitude error is computed by dividing the average number of projections that are phase shifted, by the average number of projections in each cycle. We applied those criteria on linear and non-linear methods with "six" as the number of bins used.

Table 9: Overall Average and Standard Deviation of Phase Shift in Multiple Patient Datasets

Patient #	Method	Average phase shift	Average STD phase shift	Breathing amplitude error (%)
1	Linear	3.77	1.31	25.13 ± 8.73
	Non-Linear	3.08	0.81	20.53 ± 5.40
2	Linear	3.59	2.44	18.82 ± 12.82
	Non-Linear	2.72	1.99	14.28 ± 10.44
3	Linear	3.83	2.65	22.02 ± 15.24
	Non-Linear	2.79	1.71	16.04 ± 9.83

As shown in Table 9, the phase shift in the Patient#2 dataset is the minimal among other datasets. This is because the gold standard we are comparing to was generated by using trajectories of the internal markers which are implanted in the organ tissues of the lung local to the tumor position. Also, in all patient datasets, non-linear method correlates more with the gold standard more than the linear method. It has a less phase shift and breath amplitude error than the linear method.

3.3.3 Comparison to Other Methods

We compare the IFDR algorithm to three approaches existing in the literature. First, we compare to the diaphragm-based approach [3] [5] [4] which detects the diaphragm position directly in the 4DCBCT radiographic projections and uses it as a surrogate of the respiratory signal. Second, we compare to the implanted marker-based approach [6] [7] in which implanted radio-opaque markers in and near the tumor, are tracked through the projections and the position of those markers is estimated and used as respiration signal. The third approach is an image based Local Intensity Feature Tracking (LIFT) [9] which extracts feature points that are local to the tumor and tracks them through consecutive lung CT scan projections resulting in a set of feature correspondences. Feature correspondences that have a motion pattern similar to the

breathing motion are selected to be used in a 2D/3D mapping procedure to recover the 3D motion of the lung. The respiratory signal is a component of the 3D motion of the lung.

In Table 10, we discuss the results of the comparison of LIFT and IFDR (using the non-linear implementation) with respect to the diaphragm-based method in Patient 1 and Patient 3 datasets, and with respect to implanted markers method in Patient 2 dataset. The average and standard deviation phase shift between them and the average breath amplitude error is presented.

Table 10: Average Phase Shift in Projections Using LIFT and IFDR Compared to Other Existing methods

Patient #	Method	Average phase shift	Average STD phase shift	Average Breathing amplitude error (%)
1	LIFT	1.68	1.09	11.20 ± 7.27
	IFDR	3.08	0.81	20.53 ± 5.40
2	LIFT	3.04	1.52	15.96 ± 7.98
	IFDR	2.72	1.99	14.28 ± 10.44
3	LIFT	1.25	0.83	7.19 ± 4.77
	IFDR	2.79	1.71	16.04 ± 9.83

As you can see in Table 10, our proposed IFDR method has less phase shift and breathing amplitude error than LIFT when compared to internal markers method used in Patient # 2 dataset. However, comparing to the diaphragm-position based method, LIFT has less phase shift and breathing amplitude error than the proposed IFDR.

Also, we compare IFDR to the other mentioned methods based on their limitations such as missing data. Missing data represents the number and percentage of projections in which the method was unable to produce a respiration signal. Table 11 compares IFDR with all other methods based on missing data.

Table 11: Average Error in Respiratory Motion Compared to Other Existing methods

Method	Samples	# Projections with Missing data (%)
Diaphragm position based	Patient 1 dataset	0
	Patient 3 dataset	1235 out of 2436 (50.70)
Internal Markers	Marker 1	2 out of 2395 (0.08)
	Marker 2	2 out of 2395 (0.08)
	Marker 3	648 out of 2395 (27.05)
	Marker 4	719 out of 2395 (30.02)
Local Intensity Feature Tracking (LIFT)	N/A	0
Proposed IFDR	N/A	0

Table 11 shows that the diaphragm-based method was unable to estimate the respiratory signal in nearly 50% of the projections in Patient#3 dataset. This is because the diaphragm was not visible. For the internal markers method, two markers (#3 and #4) were invisible in around 30% of Patient #2 dataset. The other two markers represent the breath signal with almost 0% missing respiration information. Respiration extraction methods depending on the local intensity tracking like our proposed IFDR and LIFT were able to detect the respiration signal with 0% missing respiration data. In this study, IFDR and LIFT were more reliable than other methods as they were able to generate a respiration signal depend only on the CBCT projections.

Eventually, the goal of developing this algorithm is to use an image-based surrogate to the respiratory signal without the dependence on any structure in the images. This algorithm has a limitation that it cannot be applied in real time due to its high computational cost.

3.4 SUMMARY

An image-based respiratory motion extraction method for breath phase sorting in cone beam CT images was developed and evaluated in three clinical research subjects. The respiratory motion extracted was based on the intensity flow of patient's organ tissues existing in the CBCT scan images. Because we seek a respiratory motion pattern existing in the CT images of the lung, we applied dimensionality reduction methods to the dataset of the intensity flow displacement vectors to recover those motion patterns. Experimental results conducted showed that Non-linear dimensionality reduction method showed less phase shift and breathing amplitude error than the linear method. IFDR using the non-linear implementation has an average phase shift of 2.94 ± 1.26 projections with the diaphragm position-based signal and an average breathing amplitude error of 18.04 ± 7.62. Compared to the implanted-markers based signal, IFDR under the non-linear approach has an average phase shift of 2.72 ± 1.99 projections and an average breathing amplitude error of 14.28 ± 10.44. IFDR using the linear and non-linear approaches were able to extract the breath signal in all projections of the patients' dataset.

CHAPTER 4 PROJECTION GENERATION BASED ON RESPIRATION MOTION FOR 4D-CBCT RECONSTRUCTION

In cone-beam computed tomography (CBCT) and related application, 4D-CBCT reconstruction accuracy depends highly on the phase sorting technique used. In this chapter, we study the quality of 4D-CBCT reconstructed images based on different phase sorting techniques: internal markers, external markers, and image-based phase sorting methods (LIFT and IFDR discussed in CHAPTER 2 and CHAPTER 3, respectively). Also, the quality of the reconstruction is affected by the number of projections used in the reconstruction. Reconstructing a 4D-CBCT image with too few projections in each respiratory phase leads to characteristic streak artifacts in the reconstructed image. To overcome this problem, generating additional projections by means of interpolation has been suggested. For this purpose, a new interpolation algorithm based on the respiratory motion existing in the CT projections of a human lung is developed. Experiments were conducted to generate additional projections via interpolation and quantitative evaluation evaluates the reconstruction result.

4.1 INTRODUCTION

The reconstruction of Cone Beam Computed Tomography (CBCT) projections of a moving anatomy, such as a lung, results in blurring volume due to the respiratory motion artifacts [27]. Four-dimensional or respiration-correlated CBCT (4D-CBCT) has been developed to provide respiratory phase resolved volumetric imaging in image guided radiation therapy. Phase-correlated (PC) reconstruction is achieved by correlating the reconstruction with this motion phase signal. Filtered backprojection-type Feldkamp–Davis–Kress (FDK) algorithm is the conventional PC reconstruction method used to reconstruct an image from projections of only a single motion phase [3] [42] [44]. The resulting image will have angular gaps corresponding to the other projections of the motion cycle that are not used for the reconstructions. Thus, reconstruction with too few projections results in low quality reconstructed images with obvious streaking artifacts and high noise levels [3] [45] [46]. Autoadaptive Phase Correlation (AAPC) reconstruction algorithm was proposed to reduce those artifacts by estimating the motion within the projections and reconstructing an image from unaffected projection areas regardless of the motion phase [27]. Algebraic reconstruction (ART) techniques [101] are well suited when there is a small number of projections. However, they are rarely used in practice for CBCT applications because of their computational cost.

Different strategies have been proposed in the literature to alleviate the effects of the view aliasing artifacts. Compressed Sensing (CS) was applied to tomographic beam problems by Sidkey *et al.* [50]. It uses a transformation onto the image and tries to find the transformed representation of the image with the least significant entries, while simultaneously bounding the error between the projected image and the measured data.

Another CS example was [51] which uses a simple gradient descent for the optimization. More work on CS can be found in [53] [54] [52] [55]. Motion compensating algorithms are also available. Those algorithms use further prior information, *i.e.*, deformation fields calculated from the planning CT [59] [56] [58] [43] [60] [61]. In [60] [61], an algorithm based on the PCA lung motion model has been proposed and evaluated to reconstruct volumetric images and extract 3D tumor motion information in real-time from a single x-ray projection in a markerless implantation. These motion compensating approaches are believed to give the best results when accurate deformation information are available as the motion can potentially be almost completely compensated in the reconstruction algorithm [63]. However, the calculation of the deformation maps is costly and the results depend on the accurate registration of the deformation map to the current patient position of each treatment session.

Increasing the sampling density by projections generation may help reducing the effects of the view aliasing artifacts. The topic of image interpolation has been studied in medical imaging research [67] [68] [69] [70] [71] [72] [27] [73] [74] [75]. For parallel beam geometry, Weiss *et al.* [69], estimated intermediate phantom views using linear interpolation and demonstrated their efficiency in minimizing view aliasing streaks. A set of interpolation schemes have been described in the literature, for medical and nonmedical applications, such as optimized interpolation kernels [71] [72] [76], and shape-adaptive image interpolation algorithms [77] [78] [79] [80]. Few advanced interpolation methods have been proposed in the literature [67] [68] [81]. In [67] [68], a shape-driven directional interpolation algorithm based on a structure tensor approach is developed. The algorithm is locally adaptive to the orientation of gray value structures to

be interpolated. Quantitative evaluation shows that this method outperforms conventional scene-based interpolation schemes and reduces streak artifacts and noise in the reconstructed images. However, this method was applied to a non-moving anatomy, *i.e.* human head, as opposed to the human lung datasets used here.

This study investigates the effect of sorting the projections using different respiratory signals on the reconstruction quality. Also, a motion-based interpolation technique is proposed to generate 2D projections of a moving anatomy in order to increase the number of projections used in reconstruction. Those new generated projections belong to the same breath phase of the desired reconstruction phase. Thus, the number of projections in desired reconstruction phase is increased and streaking artifacts can be reduced. A dense motion analysis of image pixels using an optical flow approach is conducted for every image in the sequence [88]. The intensity of every pixel is tracked through the projection sequence and a set of trajectories are formed. The location of the anatomy in the original projections in the desired reconstruction phase is used to compute the corresponding location of the anatomy in the interpolated projections. Pixel intensity values in the original corresponding projections are used to compute the intensity values of the interpolated projections.

The remainder of this chapter is organized as follows. In Section 4.2, details on 4D-CBCT reconstruction are presented. Section 4.3 presents the proposed respiration-based projection generation. Section 4.4 discusses the experimental results. The findings are summarized in Section 4.5.

4.2 Respiratory-Correlated Cone-Beam CT (4DCBCT) Reconstruction

The purpose of this section is to reconstruct respiratory-correlated (4D) CBCT volumes from on-board CBCT scans using different phase sorting techniques. We study the reconstruction quality using different phase sorting techniques and compare between the reconstructed volumes based on different assessment criteria [102].

Phase sorting techniques used are:

- Internal markers
- External markers
- LIFT (discussed in CHAPTER 2)
- IFDR (discussed in CHAPTER 3)
- Proposed projection generation method (discussed in 4.3)

Assessment criteria used to evaluate the reconstruction quality are:

- *Difference*

The signed difference of the reconstructed images using the ground truth and other phase sorting techniques are computed to show the difference between the reconstructed images.

- *Noise level in regions of interests (RIOs)*

The noise level in selected regions of interests is estimated to evaluate the quality of the reconstructed images. Several regions of interests (ROIs) in the reconstructed images are chosen to be evaluated. The noise is measured by computing the standard deviation of the intensity in the ROI.

- *Edge profiles*

A profile is a one-dimensional image extracted from a two-dimensional image along a line segment. Computing edge profiles is used to evaluate the spatial resolution of an

image in some regions. For our assessment, several positions are selected in the reconstructed images and the edge profiles are computed and compared among all images reconstructed using different phase sorting techniques.

- *Histogram entropy and normalized mutual information*

To compare the similarity between different 4D-CBCT reconstructed images using different phase sorting methods, two measures has been used, the histogram entropy and the normalized mutual information. Those measures quantify the impact of artifacts, e.g., streaks and motion blurring, and evaluate the similarity of the reconstructed volume with the ground truth.

a. Histogram Entropy

The entropy measure using the gray-level histogram *h(q)* over all intensities *q* of a given volume *X* with volume size *N* is calculated by

$$H(X) = -\sum_q \frac{h(q)}{N} \ln \frac{h(q)}{N} \qquad (16)$$

In case of all volume voxels contain the same value, the histogram has one single peak resulting in the minimum entropy of $H = 0$. On the other hand, a volume with an equal number of voxels for all intensities has a constant histogram producing the maximum entropy. The entropy measure has already been proven to be a suitable criterion to qualitative measure image artifacts in the case of a misalignment correction and is also used for motion correction.

b. Normalized mutual information (NMI)

The definition of entropy can be extended to the joint entropy $H(X,Y)$ to gain information on the similarity of two volumes X and Y by replacing the probability of one value occurring in a single volume with the probability of two values occurring together in two

different volumes. Based on both entropy and joint entropy the mutual information measures the information that two volumes share and is given by:

$$I(X,Y) = -\sum_{q_X,q_Y} \frac{h(q_X,q_Y)}{N_X,N_Y} \ln \frac{h(q_X,q_Y)}{h(q_X)h(q_Y)}$$

Mutual information is widely used as a similarity measure in the field of image registration [103]. Applying the normalized mutual information rather than the mutual information *I(X,Y)* avoids the dependence on the amount of volume overlap. The normalized mutual information is given by:

$$NMI(X,Y) = \frac{I(X,Y)}{H(X,Y)} \qquad (17)$$

Those four criteria discussed above are used in 4.4 to evaluate the 4D-CBCT reconstruction quality based on different respiratory signal extraction method.

4.3 Respiration-based Projections Generation

In this section, the details of the respiration-based projection generation are discussed. The proposed projection generation algorithm comprises of the following three steps.

Step 1: *Respiratory motion estimation and projection sorting*- An image-based respiratory signal extraction algorithm has been used in this study to extract the respiratory signal and sort the CBCT projections into phases based on their respiratory signal. A specific phase is chosen for reconstruction, *i.e.* the end of inhale. In this study we consider the respiratory signal extraction algorithm described in CHAPTER 2 to be used.

Step 2: *Dense intensity flow tracking:* Optical flow motion is computed for every point in each pair of adjacent projections in the sequence 1,…, *F*. Each point is tracked in

the sequence of projections forming point trajectories as shown in Fig. 19. The optical flow motion through the sequence of projections is regularized from projection to another based on the motion of the adjacent pixels. The regularization of optical flow smoothes the motion of the anatomy and enhances the interpolation results.

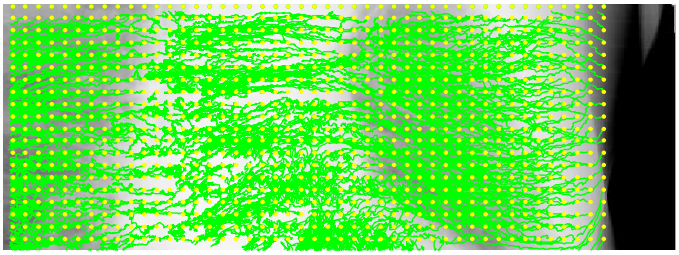


Fig. 19. Dense intensity flow tracking
Yellow dots show the points and green lines show the trajectories through the sequence of projections

Step 3: *Motion-based image interpolation-* In this procedure, we aim to generate projections based on the respiratory motion and the intensity information existing in the corresponding original projections. The coordinate system of the CBCT imaging equipment that is used for generation is shown in Fig. 20. The position of the kilovoltage source rotates in the *XY* plane, with the *Z* direction corresponding to the patient cranio-caudal axis, with positive *Z* pointing away from the linac gantry. The *xyz* coordinate

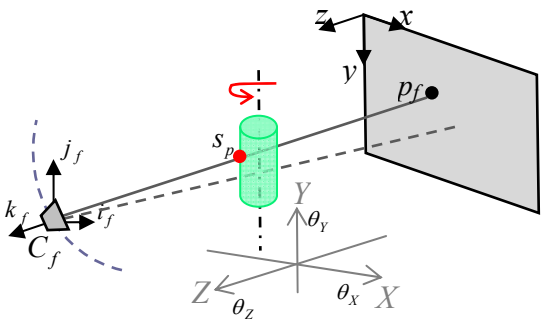


Fig. 20. Coordinate systems of the Cone Beam CT.
The world coordinates of the object point s_p are *XYZ* and the homogeneous coordinates of the image points *p* are *xvz*.

system is fixed to the imaging system, and rotates with the source and detector. The *xyz* coordinate system is the world coordinate system and is stationary.

Generating the in-between projections requires the following: (1) two images I_f and I_{f+n} , representing views of the same patient's lung and are in the same phase, (2) the original in-between image I_i, where $f < i < f+n$, and (3) a dense pixel correspondence between images I_f, I_i and I_{f+n}. Fig. 21 is illustrative figure showing the process of respiration-based image interpolation. In this procedure, point locations in the generated image I'_i are estimated based on the locations of their flow correspondences in the other original images in the desired phase I_f and I_{f+n}. To calculate point locations in the generated image I'_i, we assume that the location of the anatomy has been changed due to the gantry rotation only and we subtract the location change due to respiration motion.

The interpolated projection at location i : I'_i, is considered as the same phase of the two images I_f and I_{f+n}. The generated projection I_i is calculated according to the

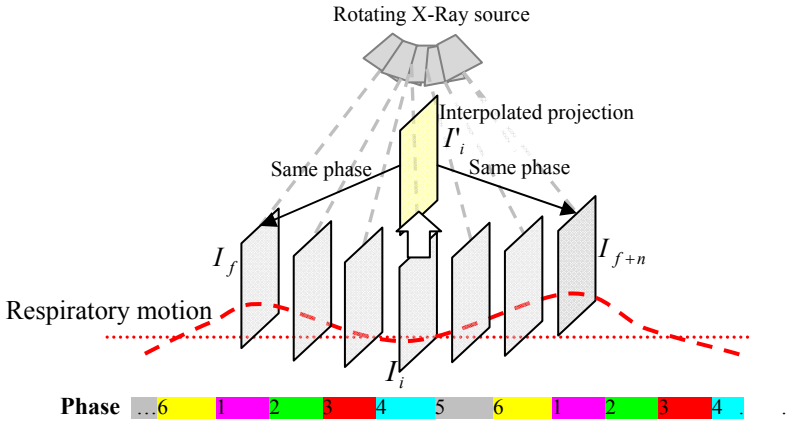


Fig. 21. Interpolation process of projections in the same breathing phase.
I_f and I_{f+n} exist in the same breath phase (i.e. phase "1" of pink color). The projection I'_i (shown in yellow) is the in-between interpolated projection which is supposed to be in the same phase as the projections used in the interpolation (i.e. phase "1"). The intensity information used in generating the projection I'_i is taken from projection I_i.

equation :

$$I'_i \left[(f+n-i)\, x_{f,p} + (i-f)\, x_{f+n,p} , (y_{f,p} + y_{f+n,p})/2 \right] = I_i \left[x_{i,p} , y_{i,p} \right], \qquad (18)$$

where p is the point number, $x_{i,p}$, $y_{i,p}$ are the image coordinates at projection i.

Using equation (18), point locations in the generated image I'_i are estimated based on the locations of their correspondences in the original images in the desired phase I_f and I_{f+n}. Once the location of the point has been computed, the intensity value of the point is copied from the corresponding original image in the sequence I_i. The same procedure is applied to all points of to generate the interpolated image I'_i. Thus, the generated images in one respiratory cycle all appear as having a static anatomy projection from different projection angles.

4.4 EXPERIMENTAL RESULTS

In this section, the quantitative evaluation of the 4D-CBCT reconstruction based on different phase-sorting methods is conducted in 4.4.1. Also, the results of the proposed respiration-based projections generation method are discussed in 4.4.2.

4.4.1 4D-CBCT RECONSTRUCTION BASED ON DIFFERENT PHASE-SORTING METHODS

In this section, the reconstruction quality is discussed using different phase sorting methods as mentioned in 4.2. Table 12 shows the set of phase sorting methods used in this section and datasets available for this study.

Table 12: Ground truth and phase sorting methods using in Multiple Patient Datasets

Method / Datasets	Number of projections	Ground Truth	Internal markers	External markers	LIFT	IFDR	Projection generation
Phantom	720	known			✓		✓
Patient 2	2396	Internal markers	✓		✓	✓	
Patient 4	3517	Internal markers	✓	✓	✓		

Patient 2:

For patient 2 dataset, the internal markers' respiratory signal is used as the ground truth. Fig. 22 shows the reconstructed images using phase sorting based on the internal markers trajectories. The figure shows the reconstructed images in the end of exhale (a) and end of inhale (b) in three views: axial, coronal and sagittal views. A difference image is also presented in (c) to show the difference is shown in the first column from right.

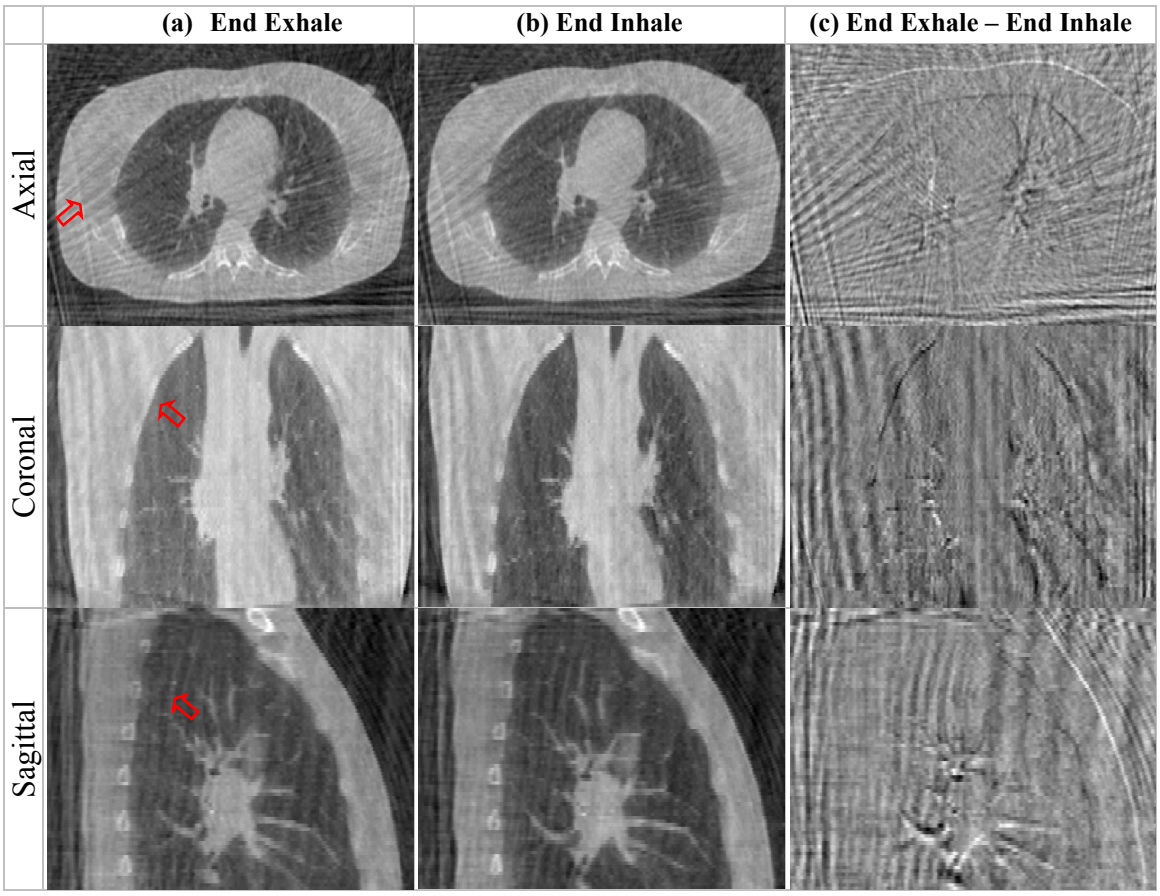


Fig. 22. 4D-CBCT Reconstruction from projections of Patient 2 in two different phases
Ground truth (internal markers' result) of the patient 2 in end-exhale (a) as well as end-inhale (b) phase bin, and the difference of the corresponding images (c). The bin width is 17%. All images are displayed at a grayscale window of *min* = −6711 HU, *max* = 9917 HU. There are some streak artifacts (shown in red arrows) in these images due to the small number of projections used in the reconstruction.

As you can see from Fig. 22 (c), the difference between the images in different respiratory phases appear mainly on motion affected regions, such as chest wall, diaphragm and soft tissues. Fig. 23 showd the reconstruction result using different phase sorting techniques: Internal markers result which acts as the ground truth (a), LIFT (b), and IFDR (c).

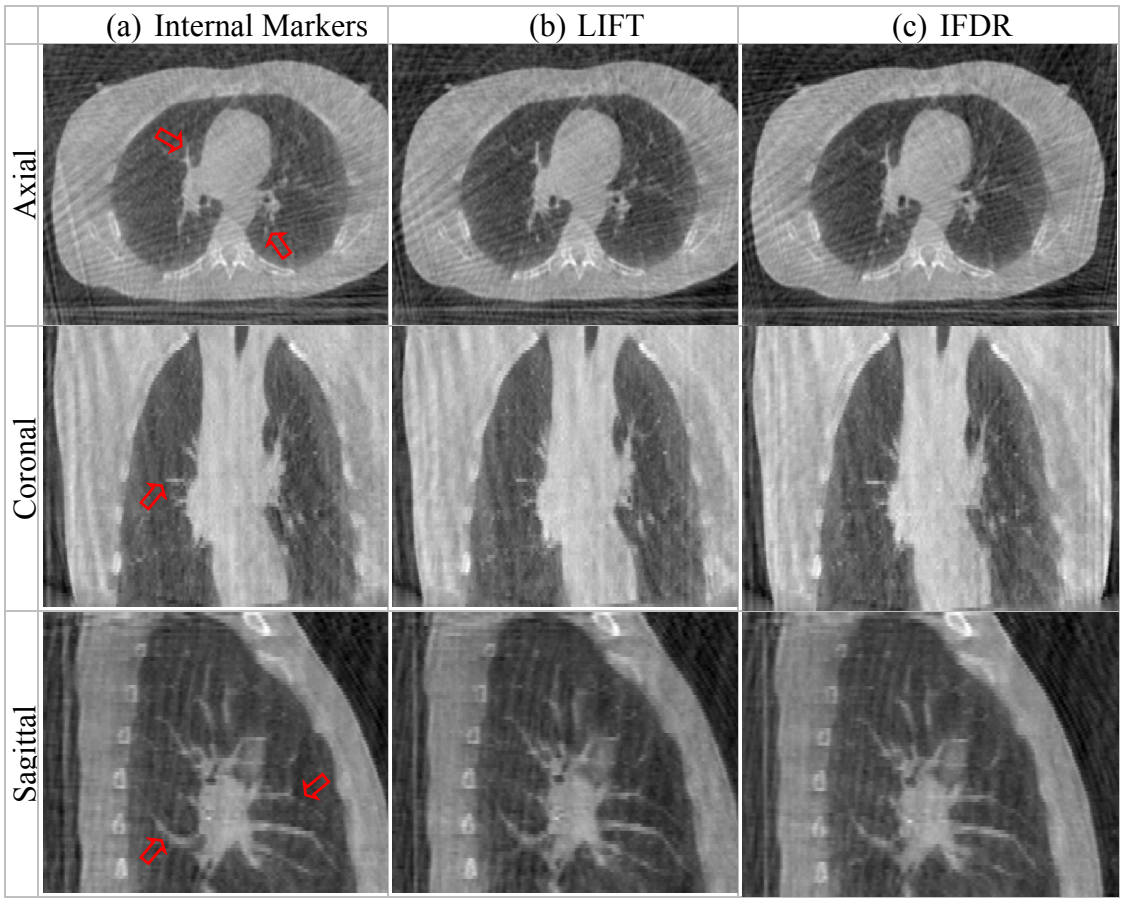


Fig. 23 4D-CBCT Reconstruction from projection of Patient 2 sorted using different methods
Ground truth (marker's result) as well as reconstruction of patient 2 using FDK with different phase sorting methods: An axial (top row), a coronal (middle row), and a sagittal view (bottom row) of the end-inhale phase bin is shown here. The different phase sorting methods used are: Internal markers (a), LIFT (b), and IFDR (c). All images are displayed at a grayscale window of $min = -6711$ HU, $max = 9917$ HU.

Difference

Difference is the first criterion used to compare the 4D reconstructed images. Fig. 24 shows the signed difference images of the conventional reconstructed images using different phase sorting. The difference between LIFT-based image and internal markers-based image is shown (a) and the difference between IFDR-based image and internal markers-based images is shown in (b). Those different images are dominated by streak artifacts due to the different projections used in each bin.

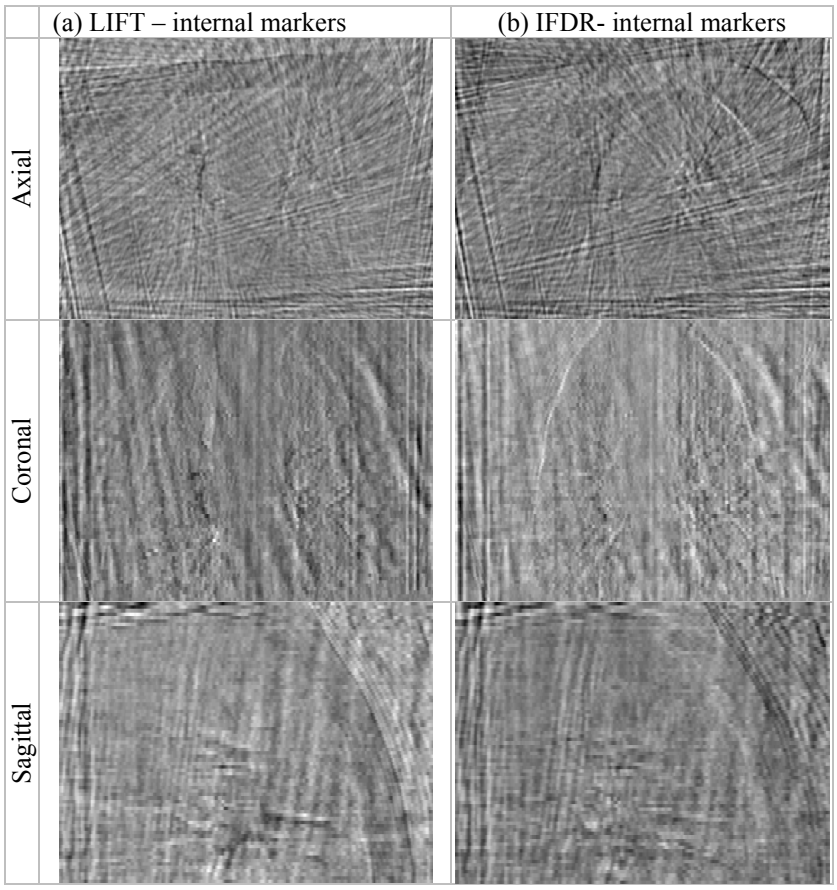
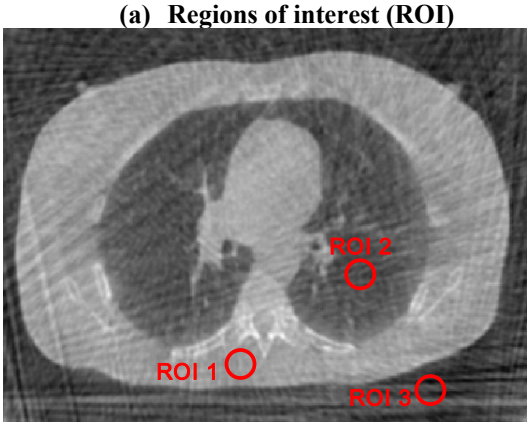


Fig. 24. Signed difference images of 4DCBCT reconstructed images of Patient 2 using different phase sorting methods compared to the ground truth (internal markers result)
Axial (top row), a coronal (middle row), and a sagittal view (bottom row) of differences for the end-inhale phase bin is shown here. All images are displayed at a grayscale window of *min* = -4348 HU, *max* = 3885 HU.

Noise measurement

Noise measurement is the second criterion used to evaluate the quality of the 4D-CBCT reconstruction. In Fig. 25, the noise level is evaluated in three regions of interest (ROI). A soft tissue region (ROI 1), a lung tissue region (ROI 2), and an air region (ROI 3) are considered as depicted in Fig. 25 (a). The table in Fig. 25 (b) shows the corresponding standard deviations in a single reconstructed image for each ROI. Noise measurement results show that the 4D images reconstructed using the image-based phase sorting methods, have nearly similar noise measurements.

(a) Regions of interest (ROI)



(b) the standard deviation of the intensity in a ROI

	Internal markers	LIFT	IFDR
ROI 1	388.21	351.5	370.91
ROI 2	616.39	482.70	420.02
ROI 3	1552.4	1234.9	1276.9

Fig. 25. ROI positions and noise values
Noise was measured in reconstructed images using FDK, with different phase sorting techniques.

Edge Profiles

Edge profile analysis is the third criterion used to evaluate the quality of the 4D-CBCT reconstructed images. Several edge profiles were examined for axial, coronal and sagittal views of patient 2 dataset in the ROIs shown in Fig. 26. The regions of interest are placed in various places, e.g., in motion-affected and motionless regions. The edge profiles are measured for the reconstructions based on phase sorting using LIFT, IFDR and the ground truth (internal markers result). The edge profiles for Fig. 27 (a) show the change of CT values for a part of the moving chest wall, for Fig. 27 (b) a moving pulmonary vessel and for Fig. 27 (c) a stationary edge between the spine and the lung. In addition, Fig. 27 (d) is a part of a soft tissue and a moving pulmonary blood vessel. Fig. 27 (e) is two moving pulmonary blood vessels. Fig. 27 compares these measured edge profiles to demonstrate that effect of different phase sorting methods on the spatial

resolution. These profiles all indicate that there is no loss of edge information due to different phase sorting methods compared to the internal markers' sorting method. In particular, the edge profiles for Fig. 27 (b), (d), (e) demonstrate that the position of the moving blood vessels is detected in the almost same position compared to the ground truth.

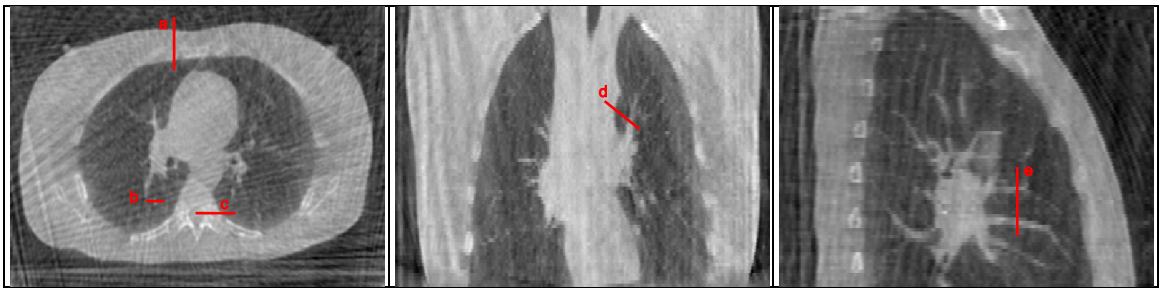
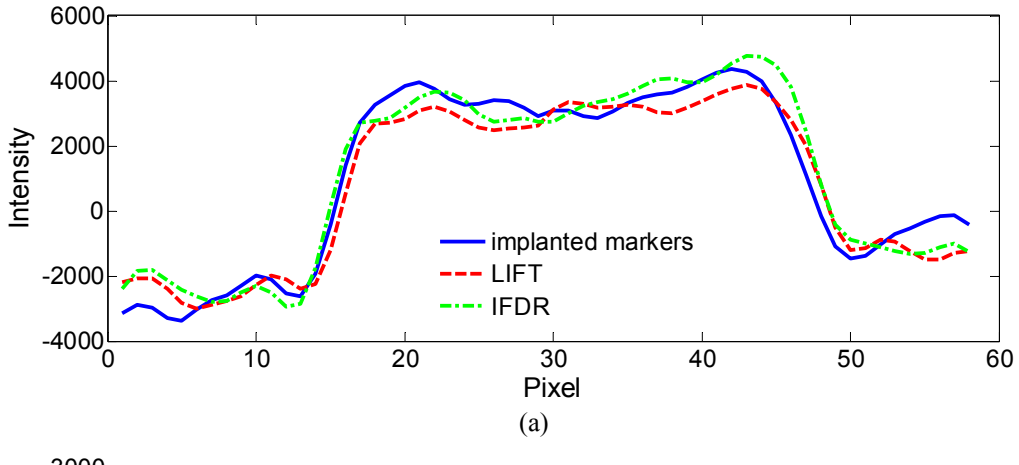
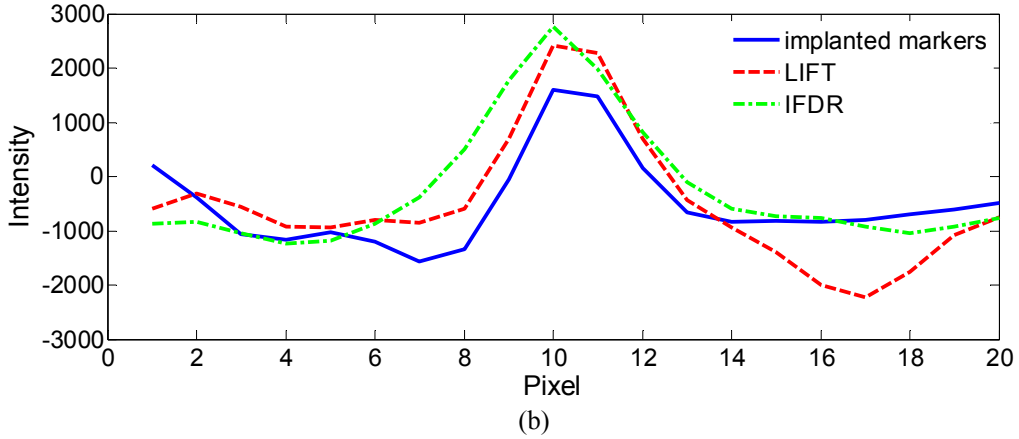


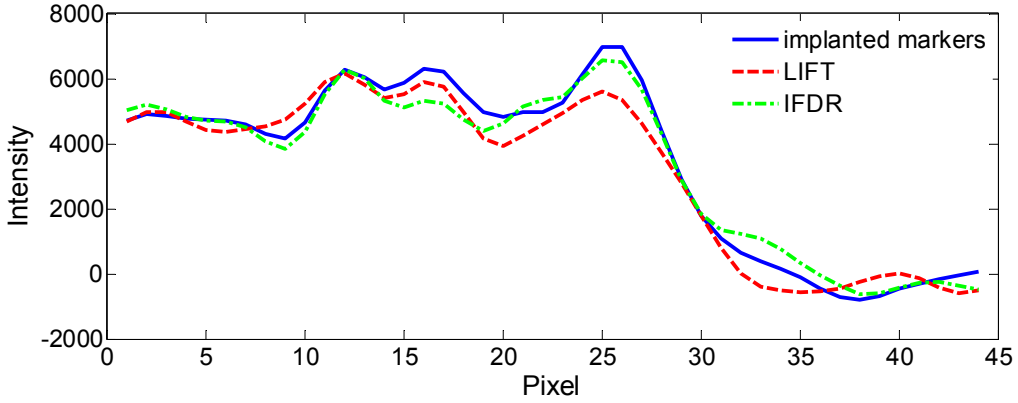
Fig. 26. ROI positions for edge profiles for Patient 2
Five different regions of interest (a)–(e) in three different views of the reconstructed images of patient 1. The ROIs are selected for the evaluation of spatial resolution.



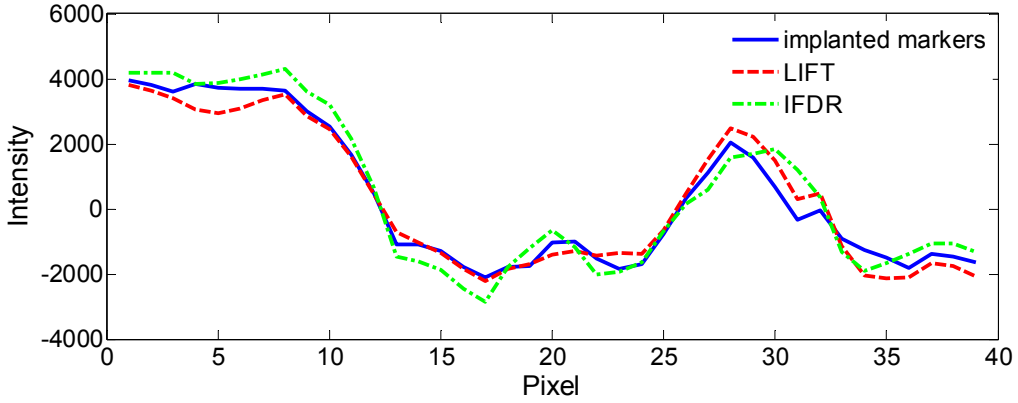
(a)



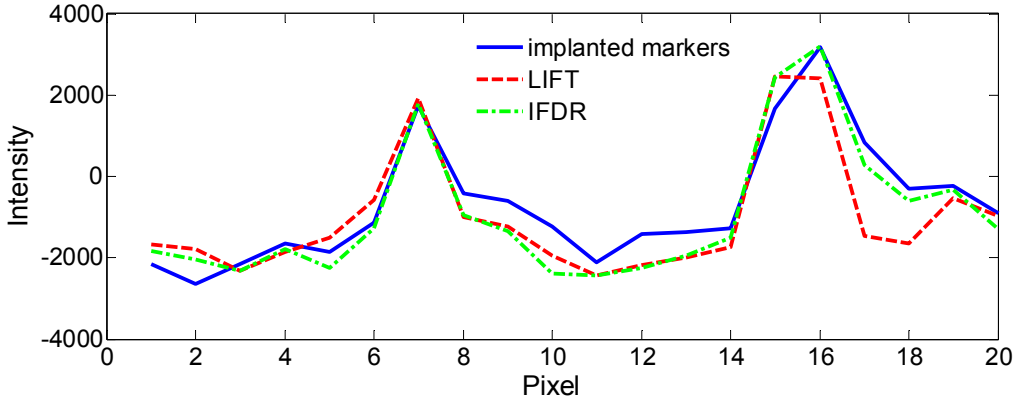
(b)



(c)



(d)



(e)

Fig. 27. Edge profiles in reconstructed images of Patient 2

Edge profiles in images reconstructed using ground truth (implanted markers) and different phase sorting algorithms (LIFT, IFDR) for five different regions of interest (a)–(e) as depicted in **Fig. 26**. The respective edge profiles are given for internal markers (solid blue line), LIFT (dashed red line), and IFDR (dashed-dotted green line). The first profile (a) shows a part of the moving chest wall, the second profile (b) a moving pulmonary blood vessel, the third profile (c) a stationary edge next to the spine, the fourth profile (d) a part of a soft tissue and a moving pulmonary blood vessel, and finally the fifth profile (e) two moving pulmonary blood vessel again.

Histogram entropy and normalized mutual information (NMI)

This is the forth criterion used in this study to evaluate the 4D-CBCT reconstruction result. Table 13 shows quantitative measurements for image quality. The entropy, defined in equation (16), is used as an image quality measure to determine the impact of artifacts such as streak artifacts or blurry edges on the reconstruction result. 4D-CBCT results shown earlier have streaks and artifacts due to the small number of projections used in the reconstruction. Regarding the entropy, all methods have close results. The values of the normalized mutual information, defined in equation (17), show that LIFT and IFDR have similar reconstructed images as the ground truth.

Table 13. Entropy and normalized mutual information measurements for different phase sorting techniques of Patient 2. The volumes are cropped before calculating those results.

X	E(X)	NMI(X, markers)
Markers	6.9330	0.9992
LIFT	6.8341	0. 2622
IFDR	6.8608	0.2700

Patient 4:

For patient 4 dataset, the internal markers-based respiratory signal is used as the ground truth. Fig. 28 shows the reconstructed images using phase sorting based on the internal markers results. The figure shows the reconstructed images in the end of exhale (a) and end of inhale (b) in three views: axial, coronal and sagittal views. A difference image is also presented to show the difference is shown in (c).

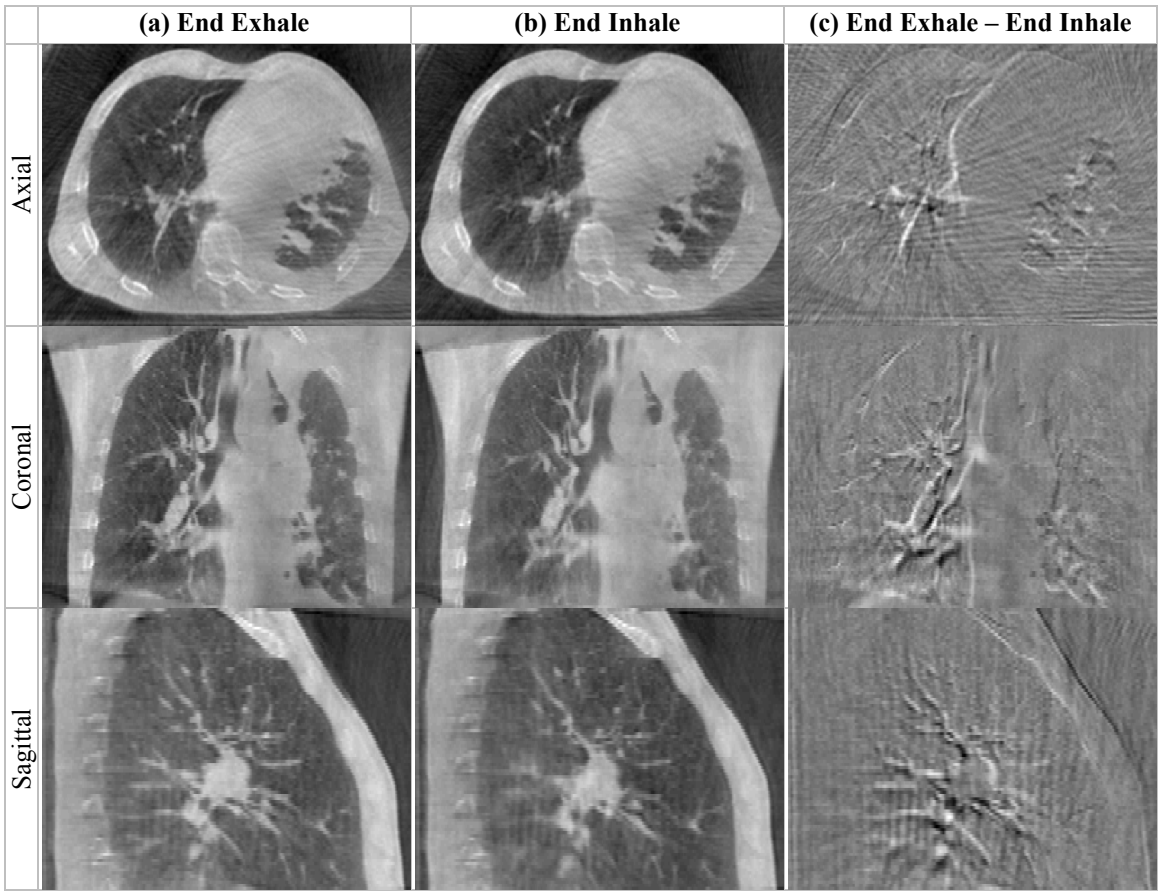
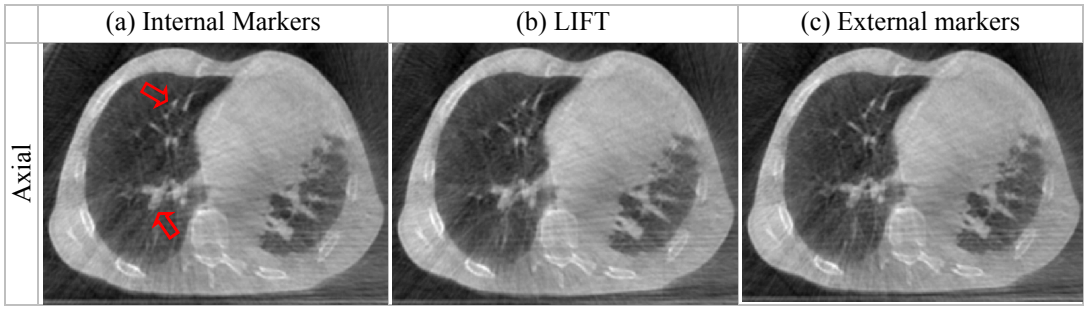


Fig. 28 4D-CBCT Reconstruction from projection of Patient 4 in two different phases
Ground truth (internal markers' result) of the patient 4 in end-exhale (a) as well as end-inhale (b) phase bin, and the difference of the corresponding images (c). The bin width is 17%. All reconstructed images are displayed at a grayscale window of *min* = −12067 HU, *max* = 14416 HU and the difference images are displayed at a grayscale window of *min*= -10297 HU, *max*= 10812 HU.

Difference

Fig. 29 shows the reconstruction result using different phase sorting techniques: Internal markers which acts as the ground truth (a), LIFT (b), and external markers (c).



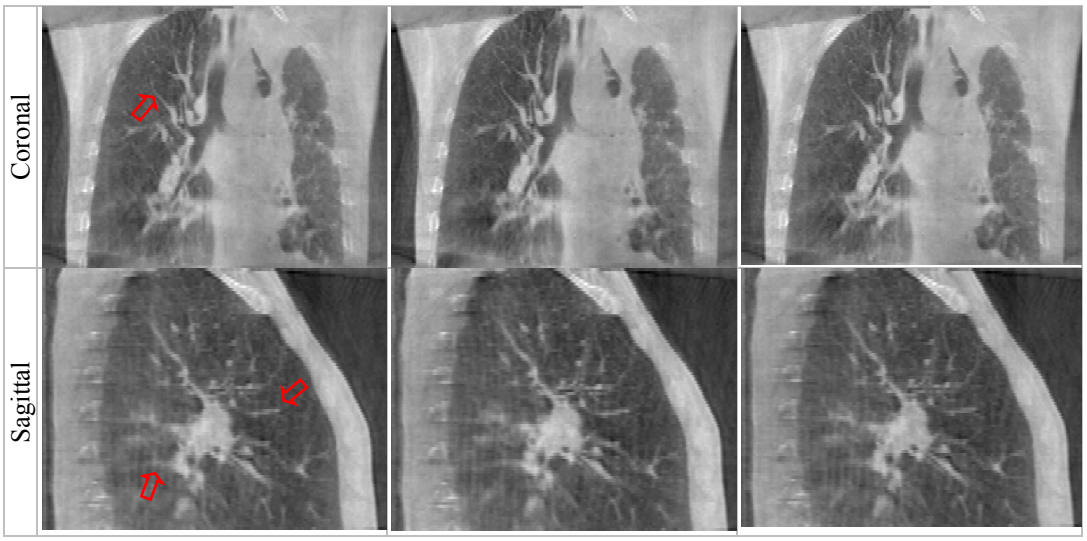
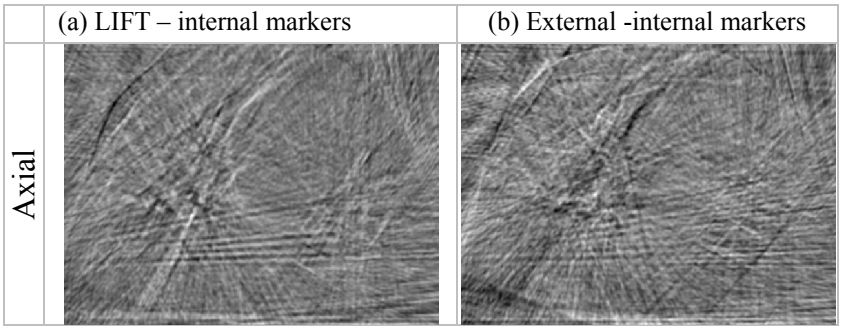


Fig. 29. 4D-CBCT Reconstruction from projection of Patient 4 sorted using different methods
Ground truth (marker's result) as well as reconstruction of patient 4 using FDK with different phase sorting methods: An axial (top row), a coronal (middle row), and a sagittal view (bottom row) of the end-inhale phase bin is shown here. The different phase sorting methods used are: Internal markers (a), LIFT (b) and external markers (c). All images are displayed at a grayscale window of *min* = −12067 HU, *max* = 14416 HU.

Fig. 30 shows the differences of the conventional reconstructed images using different phase sorting methods compared to the ground truth (internal markers signal). The difference between LIFT and markers is shown in (a) and between the external and the internal markers is shown in (b). The signed difference images are dominated by streak artifacts due to the different projections used in each bin.



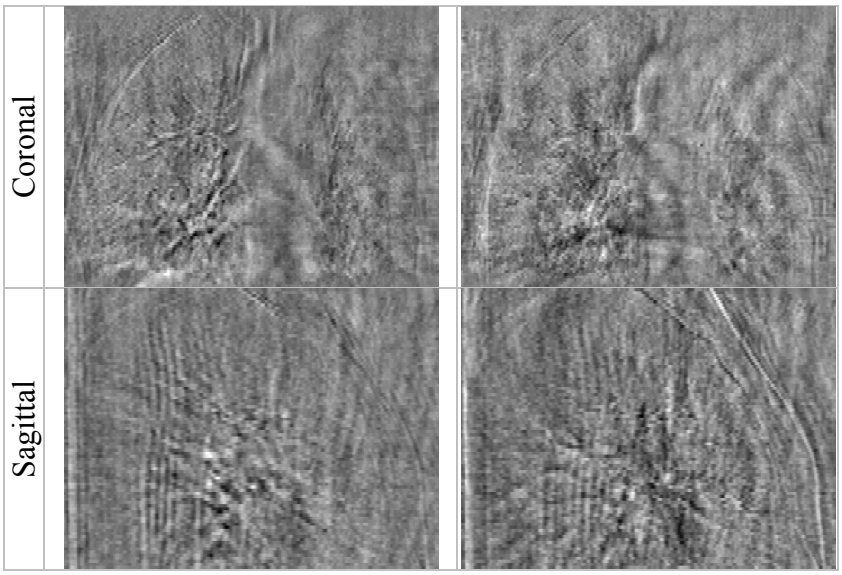
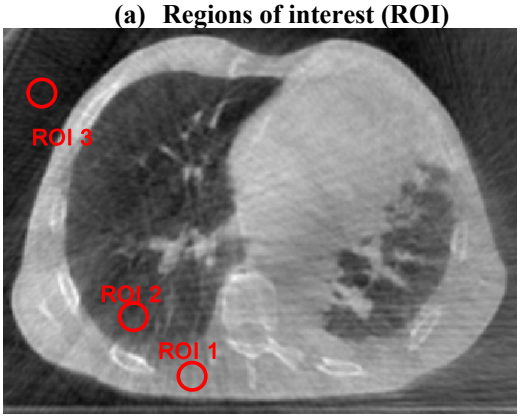


Fig. 30. Signed difference images of 4DCBCT reconstructed images of Patient 4 using different phase sorting methods compared to the ground truth (internal markers result)
An axial (top row), a coronal (middle row), and a sagittal view (bottom row) of differences for the end-inhale phase bin is shown here. All images are displayed at a grayscale window of *min*= -10297 HU, *max*= 10812 HU.

Noise measurement

In Fig. 31, noise level is evaluated in three regions of interest (ROI). A soft tissue region (ROI 1), a lung tissue region (ROI 2), and an air region (ROI 3) are considered as depicted in Fig. 31 (a). The table in Fig. 31 (b) shows the corresponding standard deviations in a single reconstructed image for each ROI. The noise level of the phase sorting methods used here is similar and LIFT has better results for ROI1 and ROI2.

(a) Regions of interest (ROI)



(b) the standard deviation of the intensity in a ROI

	Internal markers	LIFT	External markers
ROI 1	659.24	583.079	616.76
ROI 2	1026.9	886.21	985.74
ROI 3	394.04	441.57	421.56

Fig. 31. ROI positions and noise values
Noise was measured in reconstructed images using FDK, with different phase sorting techniques.

Edge Profiles

Several edge profiles were measured for axial, coronal and sagittal views of patient 4 in the positions shown in Fig. 32. The regions of interest are placed in various places, e.g., in motion-affected and motionless regions. The edge profiles are measured for the reconstructions based on phase sorting using LIFT, external markers and the ground truth (internal markers result). The edge profiles for Fig. 33 (a) show the change of CT values for a part of the moving chest wall, for Fig. 33 (b) a moving pulmonary vessel and for Fig. 33 (c) a stationary edge between the spine and the lung. In addition, Fig. 33 (d) is a part of a soft tissue. Fig. 33 (e) is two moving pulmonary blood vessels. Fig. 33 compares these measured edge profiles to demonstrate that effect of different phase sorting methods on the spatial resolution. Those profiles all indicate that there is no loss of edge information due to different phase sorting methods compared to the internal markers' sorting method. In particular, the edge profiles for Fig. 33 Fig. 33(b), (e) demonstrate that the position of the moving blood vessels is detected in the almost same position compared to the ground truth.

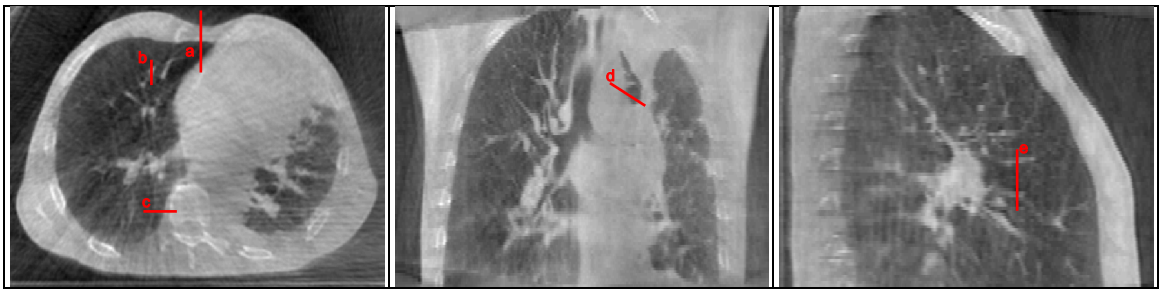
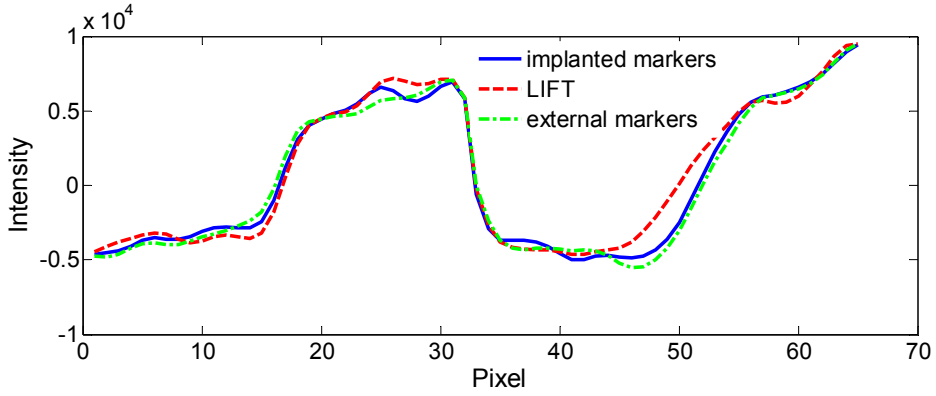
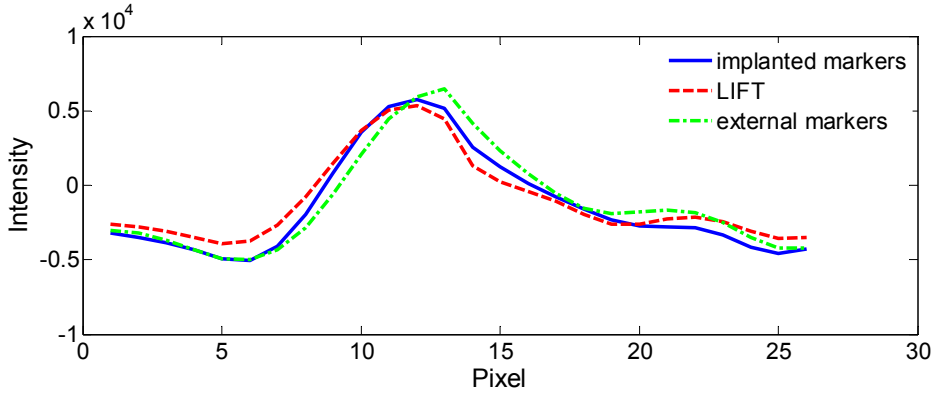


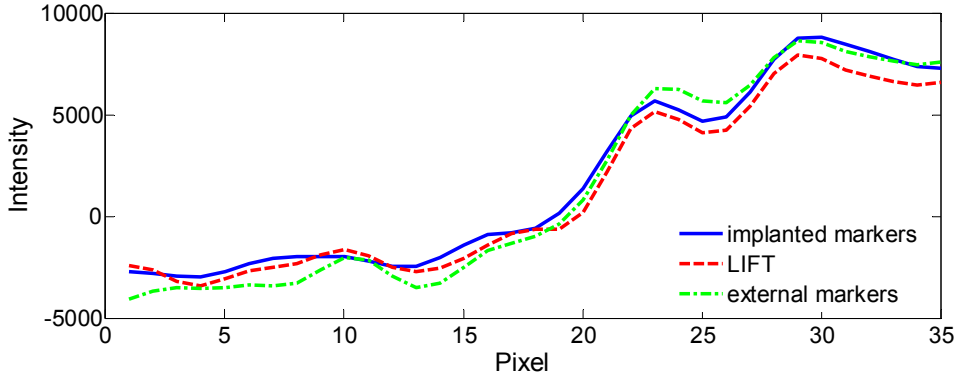
Fig. 32. ROI positions for edge profiles
Five different regions of interest (a)–(e) in three different views of the reconstructed images of patient 1. The ROIs are selected for the evaluation of spatial resolution.



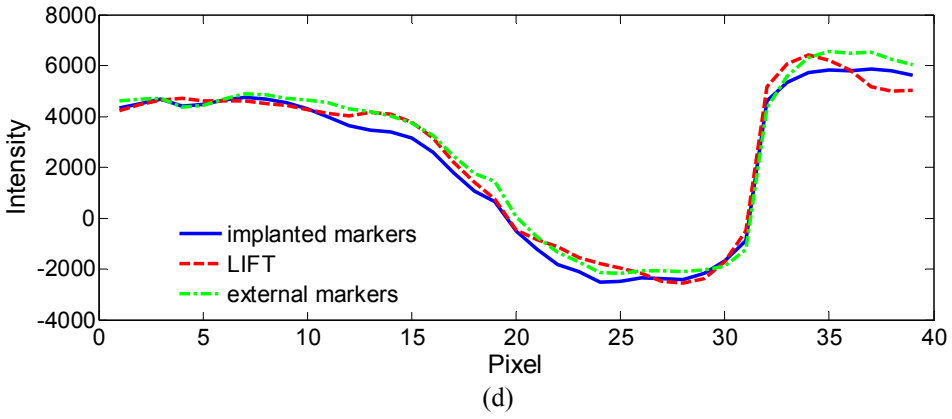
(a)



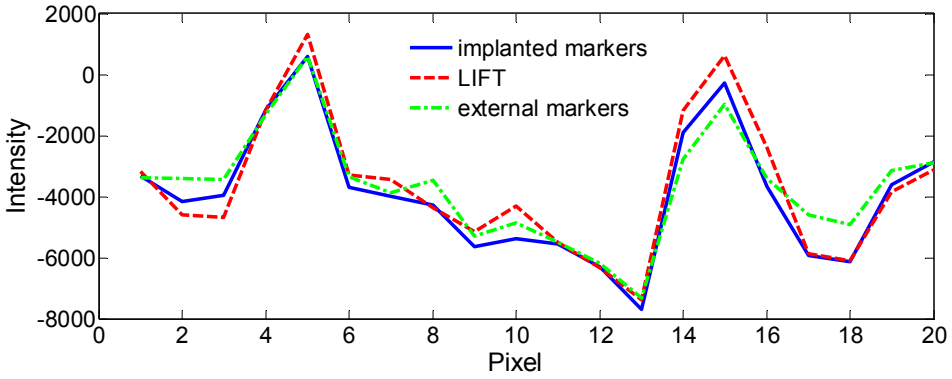
(b)



(c)



(d)



(e)

Fig. 33. Edge profiles in reconstructed images based on different phase sorting methods
Phase sorting methods used are implanted markers as ground truth and image-based sorting algorithms (LIFT, IFDR) for five different regions of interest (a)–(e) as depicted in Fig. 32. The respective edge profiles are given for internal markers (solid blue line), LIFT (dashed red line), and IFDR (dashed-dotted green line). The first profile (a) shows a part of the moving chest wall, the second profile (b) a moving pulmonary blood vessel, the third profile (c) a stationary edge next to the spine, the fourth profile (d) a part of a soft tissue, and finally the fifth profile (e) two moving pulmonary blood vessel again.

Histogram entropy and normalized mutual information (NMI)

Table 14 shows quantitative measurements for image quality. The entropy is used as an image quality measure to determine the impact of artifacts such as streak artifacts or blurring on the reconstruction result. 4D-CBCT results shown earlier have streaks and artifacts due to the small number of projections used in the reconstruction. Regarding the entropy, all compared methods have similar results. The second criterion used is the normalized mutual information. It shows the similarity between the reconstructed

volumes. As you can see, the NMI for the reconstructed volume using markers result with LIFT and External markers is very similar.

Table 14 Entropy and normalized mutual information measurements for different phase sorting techniques of Patient 4.
The volumes are cropped before calculating those results.

X	E(X)	NMI(X, markers)
Markers	6.626	0.9992
LIFT	6.719	0.4111
External markers	6.625	0.4418

The experimental results in this section showed that the proposed image-based phase sorting methods LIFT and IFDR, are performing well when used for 4D reconstruction comparing to the standard phase sorting methods.

4.4.2 Respiration-Based Generation and Reconstruction

The experimental results of the respiratory-based projection generation and the 4D-CBCT reconstruction are discussed in this section. The image generation and 4D-CBCT reconstruction was evaluated using the digital XCAT phantom [103] [104]. The phantom body represents an average human male in shape, proportion and composition. The physical pixel size used is 0.65 x 0.65 x 3.125 (mm). The phantom was used to generate 2D projections from the phantom each as an 512 x 512 array. The full scan was performed by a single 360° gantry rotation resulted in projecting 720 projections with 0.5° degrees of projection angle between each pair of adjacent projections. A total of 40 respiratory cycles were generated with each cycle consists of 18 projections. Those 2D projections are used to reconstruct a 4D-CBCT volume; the spatial resolution used was the same as the original resolution of the phantom.

Fig. 34 shows selected original 2D projections of the phantom dataset from one respiratory cycle (raw (a)) and the corresponding interpolated projections using the method discussed earlier in Section 4.3 (raw (b)). The auxiliary red dashed line is used to help the reader see the location of the anatomy in multiple images. As seen in row (a), the 2D projections are from different respiratory phases, and the position of the anatomy, such as the diaphragm, differs among projections. In raw (b), the original projections are re-generated using the proposed method to represent an identical respiratory phase (end of exhale). It is clearly shown that the generated projections in raw (b) have similar anatomy position and thus can be sorted to the same respiratory phase.

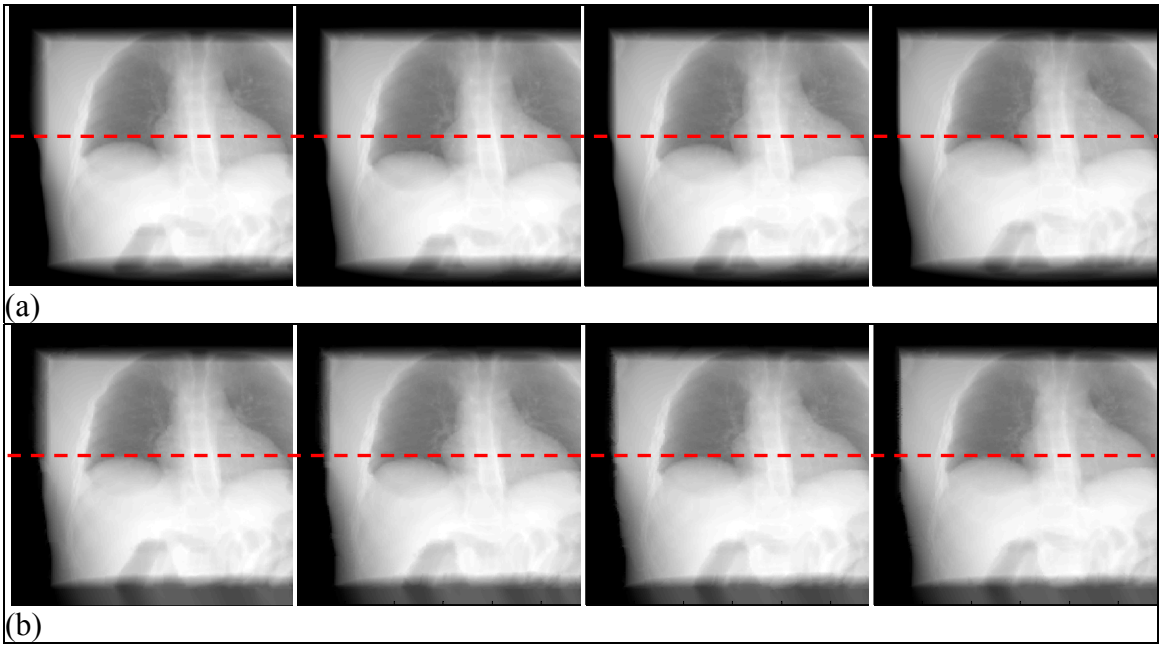


Fig. 34. Original and generated projections from the phantom dataset
Row (a) shows the original projections #(4,8,12,16). Row (b) shows the corresponding 'generated' projections using the interpolation method suggested. The red dashed line is an auxiliary line to show the difference in the diaphragm position. In row (a), the projections are in different phases, while in row (b), the generated projections are in the same phase.

Fig. 35 shows the reconstructed 4D-CBCT images for the phantom dataset using the FDK algorithm [42]. The figure shows the reconstructed images in three views: axial, coronal and sagittal views. The images reconstructed using all the original projection sorted in all phases (a), only the original projections in the end of exhale (b) and using all the generated projections that are sorted in end of exhale (c).

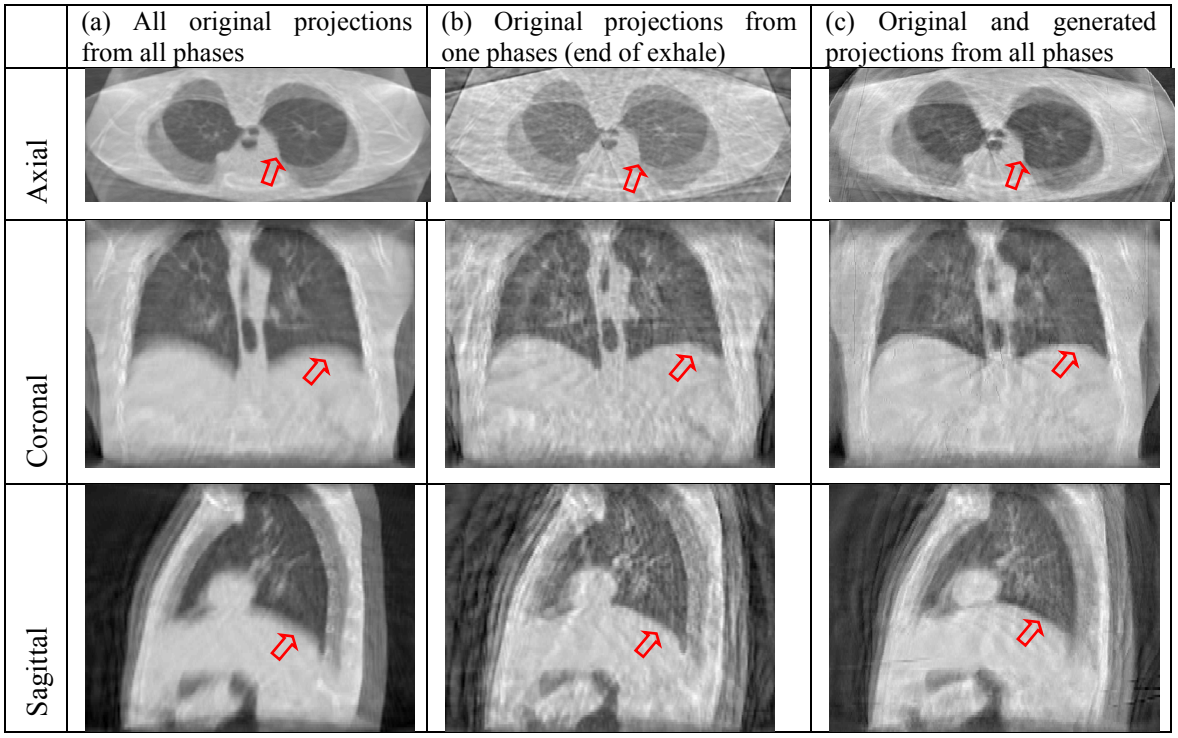


Fig. 35. Reconstructed 4DCBCT images of phantom dataset
Phantom projections are reconstructed using (a) all original projections in the dataset, (b) only original projections in the end of the exhalation phase and (c) all generated projections in the dataset. The reconstructed volumes are shown in three views: Axial, coronal and sagittal. The red arrows show the motion-affected regions of the which appear blurred in the reconstructed image

As you can see in Fig. 35, the reconstructed images in (a) are clear and do not contain noise comparing to the other images in (b) and (c). However, images in (a) are blurry due to the motion artifacts that result from reconstructing a volume using all projections in all phases. The red arrows in the images point to those motion affected

edges. The reconstructed images in (b) have almost sharp (non-blurry) edges comparing to (a). However, they contain a lot of noise and streaking artifacts due to the insufficient number of projections used in the reconstruction. In (c) the noise is less than (b) because of using the generated projections. Also, the motion affected edges appear to be sharp due to using generated projections in the same respiratory phase.

Several edge profiles were measured for axial, coronal and sagittal views of the phantom dataset in the ROIs shown in Fig. 36. The ROIs are placed in various motion-affected regions. The edge profiles for (a) show the change of CT values for a part of the moving anatomy (heart), for (b) a moving pulmonary vessel and for (c) a part of a moving anatomy. Fig. 37 compares these measured edge profiles to demonstrate that effect of using the generated projections instead of the original ones on the spatial resolution. The profiles (b) and (c) indicate that the reconstructed images using the generated projections have sharper edges comparing to the reconstruction images using all the original projections. Also, using only original projections sorted in one phase shows loss of edge information due to the streaks and noise. In particular, the edge profiles for Fig. 37 (b) and (c) demonstrate that the position of the moving blood vessels and diaphragm edge is sharper in the images reconstructed using the one phase reconstruction or generated projections as pointed at by the black arrows. The edge profile (a) shows a blurry edge in all the three reconstructed images. This indicates that the motion of the heart cannot be corrected using 4D reconstruction.

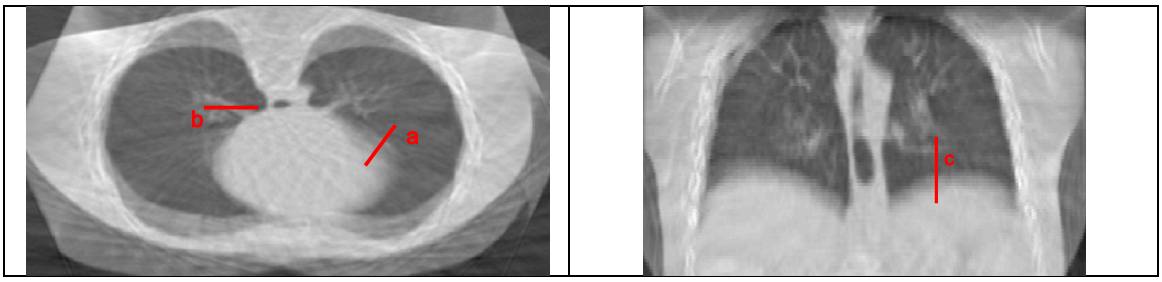


Fig. 36. ROI positions for edge profiles
Three different regions of interest (a)–(c) in two views of the reconstructed images of the phantom. The ROIs are selected for the evaluation of spatial resolution.

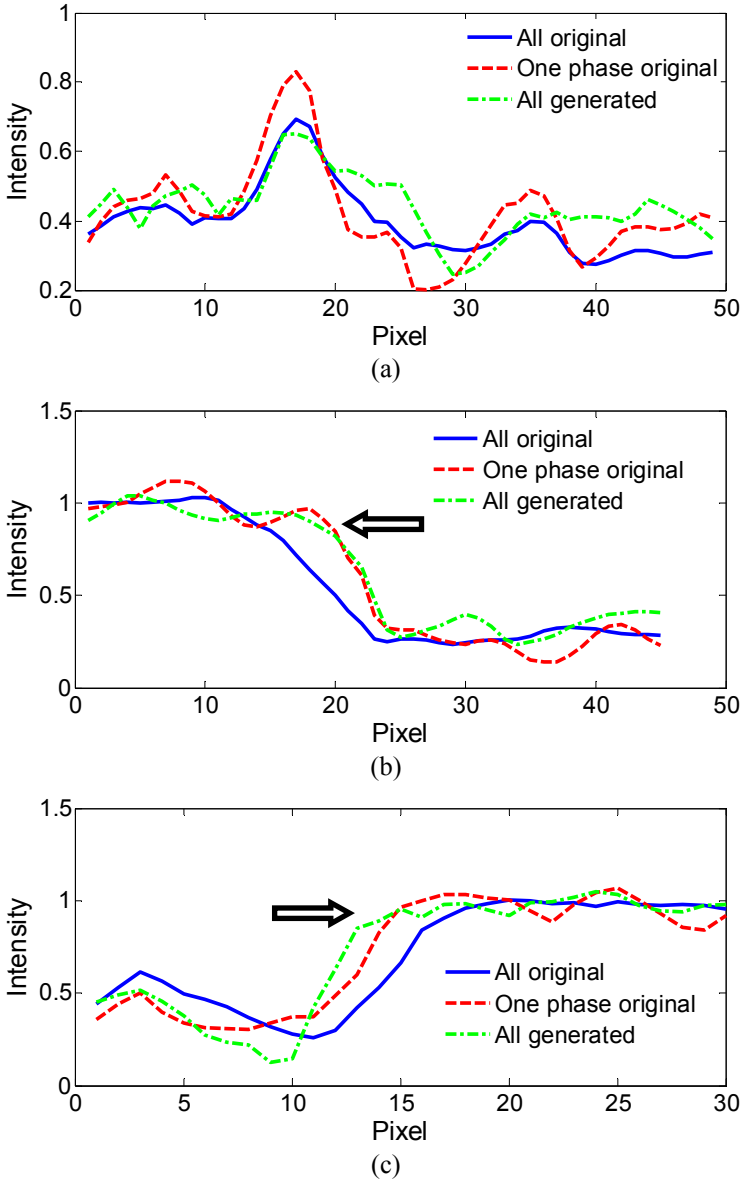
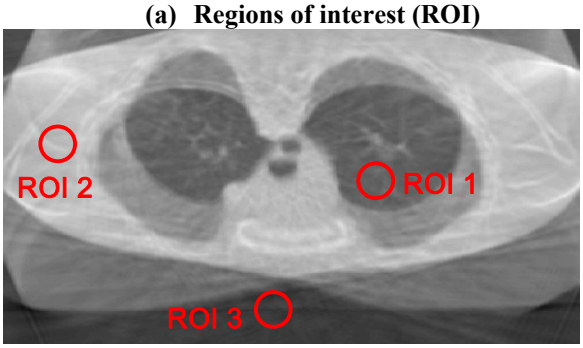


Fig. 37. Edge profiles in reconstructed images
Edge profiles in images reconstructed using all original projections in all phases, original projections in one phase (end of exhale) and all generated projections for three different regions of interest (a)–(c) as depicted

in Fig. 36. The respective edge profiles are given for all original projections (solid blue line), original projections in one phase (dashed red line), and all generated projections (dashed-dotted green line). The edge profiles for (a) show the change of CT values for a part of the moving anatomy, for (b) a moving pulmonary vessel and for (c) a part of a moving anatomy.

In Fig. 38, the noise level is evaluated in three ROIs. A lung tissue region (ROI 1), a soft tissue region (ROI 2), and an air region (ROI 3) are considered in Fig. 38 (a). The table in Fig. 38 (b) shows the corresponding standard deviations in a single reconstructed image for each ROI. The noise level of the reconstructed images using all generated projections is always less than the one using the reconstruction of one phase.

(a) Regions of interest (ROI)



(b) the standard deviation of the intensity in a ROI

	All original projection	Projections in one phase	All generated projections
ROI 1	0.048	0.086	0.061
ROI 2	0.022	0.051	0.032
ROI 3	0.038	0.137	0.053

Fig. 38. ROI positions and noise values
Noise was measured in reconstructed images using all original projections, original projections in one phase, and all generated projections.

Table 15 shows the entropy and the NMI measurements of reconstructed images using all original projections, original projections in one phase and all generated projections.

Table 15: Entropy and normalized mutual information measurements for reconstruction of phantom dataset using all original projections, one phase of original projections, and all generated projections.
The volumes are cropped and before calculating the entropy measurement (E).

X	E(X)	NMI(X, all original)
All original	6.803	0.9993
One phase original	6.769	0.2397
All generated projections	6.746	0.2364

Results in

Table 15 show that the entropy values are very similar in all reconstructed images. Also, the NMI measurements of the reconstructed images using all generated projections have similar values as the one reconstructed using only original projections in one phase.

This section showed the experimental results of the motion-based image generation method. This method used a phantom dataset to generate additional projections to be used in the 4D-CBCT reconstruction. The experimental results demonstrated the ability of the method in recovering the edges, adding no additional noise and visibly reducing streak artifacts. To determine the potential clinical impact of this study, clinical dataset should be used.

4.5 SUMMARY

This chapter analyzed the effects of using different phase sorting algorithms on the 4D-CBCT reconstruction. The phase sorting algorithms used are internal markers, external markers, and image-based phase sorting methods (LIFT, described in CHAPTER 2 and IFDR, described in CHAPTER 3). Also, a respiration-based image generation method is presented. This method is based on a motion analysis of image pixels through projections in a CBCT projection sequence. This method has been developed to generate new projections to be used in 4D-CBCT reconstruction to reduce the respiration streaking artifacts and improve the reconstruction quality. The experimental results showed that the proposed image-based phase sorting methods LIFT and IFDR, are performing well when used for 4D reconstruction comparing to the implanted markers result. The experimental results of the projection generation method demonstrated good recovery of edges, no additional noise and visibly reduced streak artifacts. To determine the potential clinical impact of this study, clinical datasets should be used.

As the study shows the feasibility of the proposed image-generation method, future work can be done to improve it. Motion detection can be improved by eliminating the point mis-correspondences and regularizing the optical flow motion. Interpolation can be improved from pixel-to-pixel interpolation to regional interpolation using motion-based image segmentation techniques. Also, advanced methods can be used to generate the intermediate projections. Machine learning and training techniques such as artificial neural networks (ANN) can be useful.

CHAPTER 5 CONCLUSIONS AND CONTRIBUTIONS

5.1 CONCLUSIONS

The following conclusions can be drawn from the results obtained in CHAPTER 2:

5.1.1 LOCAL INTENSITY FEATURE TRACKING AND MOTION MODELING FOR RESPIRATORY SIGNAL EXTRACTION IN CONE-BEAM CT PROJECTIONS

- LIFT is a novel method for respiratory motion extraction and breath phase sorting using CBCT projections. Feature points were extracted and tracked to form point trajectories. Trajectories with shapes similar to breathing curve were selected to be used in the 3D motion modeling module to recover the 3D motion of the lung. The 3D rotation around the Z-axis of the patient represented the respiratory motion in this study and the CBCT projections were then sorted according to the respiration signal.
- LIFT was able to extract the respiration signal in all projections of all datasets without the dependence on a particular anatomical structure (such as the diaphragm).
- The respiratory motion extracted using LIFT correlates with signals extracted using other standard methods with an average phase shift of 1.78 projections estimated between LIFT based signal and markers based signal, and of 1.68 projections between LIFT based signal and the diaphragm-based signal.
- The average breathing amplitude error of LIFT compared to the diaphragm-based method was 11.2% while it is 10.68% compared to the internal markers method.

- LIFT is able to extract regular and irregular breathing patterns existing in the datasets.

The following conclusions can be made from the results obtained in CHAPTER 3:

5.1.2 Intensity Flow Dimensionality Reduction for Respiratory Signal Extraction in Cone-Beam CT (CBCT) Projections

- IFDR is an image-based respiratory motion extraction method for breath phase sorting in cone beam CT images was developed and evaluated in three clinical subjects. The respiratory motion extracted was based on the intensity flow of patient's organ tissues existing in the CBCT scan images. Dimensionality reduction methods are applied to the dataset of the intensity flow displacement vectors to recover those motion patterns.
- Experimental results conducted showed that Non-linear dimensionality reduction method showed less phase shift and breathing amplitude error than the linear method.
- IFDR using the non-linear implementation has an average phase shift of 2.94 ± 1.26 projections with the diaphragm position-based signal and an average breathing amplitude error of 18.04 ± 7.62.
- Compared to the implanted-markers based signal, IFDR under the non-linear approach has an average phase shift of 2.72 ± 1.99 projections and an average breathing amplitude error of 14.28 ± 10.44. IFDR using the linear and non-linear approaches were able to extract the breath signal in all projections of the patients' dataset.

The following conclusions can be made from the results obtained in CHAPTER 4:

5.1.3 Projection Generation based on Respiration Motion for 4DCBCT Reconstruction

- Different respiration signals affect the 4D-CBCT reconstruction quality.
- Image-based phase sorting methods (LIFT, and IFDR) perform reasonably well compared to internal markers.
- A respiration-based interpolation method for image generation proposed to reduce the noise, streaking and motion artifacts in 4DCBCT projections. The proposed method is based on estimation the optical flow motion in the sequence of CBCT projections. It regenerates all the original projections from their current phase to one desired phase for reconstruction.
- The proposed method has been applied on phantom dataset and the results showed that the reconstructed images has good edge recovery, visibly reduced streak artifacts and no additional noise.
- Advanced motion detection and interpolation methods can be used in this work to improve the results.
- Clinical studies are required to prove the clinical impact and feasibility of this method.

REFERENCES

[1] T. Li, L. Xing, P. Munro, C. McGuinness, M. Chao, Y. Yang, B. Loo and A. Koong, "Four-dimensional cone-beam computed tomography using an on-board imager," *Med. Phys.,* vol. 33, no. 10, pp. pp. 3825-3833, Sep. 2006.

[2] S. S. Vedam, V. R. Kini, P. J. Keall, V. Ramakrishnan, H. Mostafavi and R. Mohan, "Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker," *Med. Phys.,* vol. 30, no. 4, pp. 505-513, Mar 2003.

[3] J. Sonke, L. Zijp, P. Remeijer and M. van Herk, "Respiratory correlated cone beam CT," *Med. Phys.,* vol. 32, no. 4, p. 1176–1186, 2005.

[4] S. Rit, D. Sarrut and C. Ginestet, "Respiratory signal extraction for 4D CT imaging of the thorax from conebeam CT projections," *Med. Image Comput. Comput. Assist. Interv.,* vol. 8, no. 1, pp. 556-563, 2005.

[5] L. Zijp, J. -J. Sonke and M. Herk, "Extraction of the respiratory signal from sequential thorax cone-beam x-ray images," in *The 14th International Conference on the Use of Computers in Radiation Therapy (ICCR)*, Seoul, Korea, 2004.

[6] T. E. Marchant, A. M. Amer and C. J. Moore, "Measurement of inter and intra fraction organ motion in radiotherapy using cone beam CT projection images," *Phys. Med. Biol.,* vol. 53, p. 1087–1098, Feb. 2008.

[7] P. R. Poulsen, B. Cho and P. J. Keall, "A method to estimate mean position, motion magnitude, motion correlation, and trajectory of a tumor from cone-beam CT projections for image-guided radiotherapy," *Int. J. Radiat. Oncol. Biol. Phys.,* vol. 72, p. 1587–1596, Dec. 2008.

[8] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. 7th Int. Joint Conf. Artificial Intelligence*, British Columbia, 1981.

[9] S. Dhou, Y. Motai and G. D. Hugo, "Local Intensity Feature Tracking for Respiratory Signal Extraction in Cone Beam CT images," *IEEE Trans. Biomed. Eng.,* vol. 60, no. 2, pp. 232-242, Feb. 2013.

[10] D. Jaffray, J. Siewerdsen, J. Wong and A. Martinez, "Flat-panel conebeam computed tomography for image-guided radiation therapy," *Int. J. Radiat. Oncol. Biol. Phys.,* vol. 53, pp. 1337-1349, 2002.

[11] J. Pouliot, A. Bani-Hashemi, J. Chen, M. Svatos, F. Ghelmansarai, M. Mitschke, M. Aubin, P. Xia, O. Morin, K. Bucci, M. Roach, P. Hernandez, Z. Zheng, D. Hristov and L. Verhey, "Low-dose megavoltage cone-beam CT for radiation herapy," *Int. J. Radiat. Oncol. Biol. Phys.,* vol. 61, p. 552–560, 2005.

[12] M. Oldham, D. Létourneau, L. Watt, G. Hugo, D. Yan, D. Lockman, L. Kim, C. P., A. Martinez and J. Wong, "Cone-beam-CT guided radiation therapy: A model for on-line application," *Radiother. Oncol.,* vol. 75, p. 271–278, 2005.

[13] E. C. Ford, G. S. Mageras, E. Yorke and C. C. Ling, "Respiration-correlated spiral CT: A method of measuring respiratory-induced anatomic motion for radiation treatment planning," *Med. Phys.,* vol. 30, p. 88–97, 2003.

[14] D. A. Low, M. Nystrom, E. Kalinin, P. Parikh, J. F. Dempsey, J. D. Bradley, S. Mutic, S. H. Wahab, T. Islam, G. Christensen, D. G. Politte and B. R. Whiting, "A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing," *Med. Phys.,* vol. 30, p. 1254–1263, 2003.

[15] T. Pan, T. Y. Lee, E. Rietzel and G. T. Chen, "4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT," *Med. Phys. ,* vol. 31, p. 333–340, 2004.

[16] S. S. Vedam, P. J. Keall, V. R. Kini, H. Mostafavi, H. P. Shukla and R. Mohan, "Acquiring a four-dimensional computed tomography dataset using an external respiratory signal," *Phys. Med. Biol.,* vol. 48, p. 45–62, 2003.

[17] C. Ozhasoglu and M. J. Murphy, "Issues in respiratory motion compensation during external-beam radiotherapy," *Int. J. Radiat. Oncol. Biol. Phys. ,* vol. 52, no. 5, pp. 1389-1399, Apr. 2002.

[18] R. I. Berbeco, S. Nishioka, H. Shirato, G. T. Chen and S. B. Jiang, "Residual motion of lung tumours in gated radiotherapy with external respiratory surrogates," *Phys. Med. Biol.,* vol. 50, no. 16, pp. 3655-3667, Aug. 2005.

[19] H. H. Liu, N. Koch, G. Starkschall, M. Jacobson, K. Forster, Z. Liao, R. Komaki and C. W. Stevens, "Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: Part II-margin reduction of internal target volume," *Int. J. Radiat. Oncol. Biol. Phys.,* vol. 60, no. 5, pp. 1473-1483, Dec. 2004.

[20] L. I. Cervino, A. K. Chao, A. Sandhu and S. B. Jiang, "The diaphragm as an anatomic surrogate for lung tumor motion," *Phys. Med. Biol.,* vol. 54, no. 11, pp. 3529-3541, Jun. 2009.

[21] D. Gierga, G. Sharp, J. Brewer, M. Betke, C. Willett and G. T. Chen, "Correlation between external and internal markers for abdominal tumors: implications for respiratory gating," *Int. J. Radiat. Oncol. Biol. Phys.,* vol. 57, no. 2 Suppl., pp. S186-S187, Oct. 2003.

[22] J. D. Hoisak, K. E. Sixel, R. Tirona, P. C. Cheung and J. P. Pignol, "Correlation of lung tumor motion with external surrogate indicators of respiration," *Int. J. Radiat. Oncol. Biol. Phys.,* vol. 60, no. 4, pp. 1298-1306, Nov. 2004.

[23] D. Ionascu, S. B. Jiang, S. Nishioka, H. Shirato and R. I. Berbeco, "Internal-external correlation investigations of respiratory induced motion of lung tumors," *Med. Phys,* vol. 34, no. 10, pp. 3893-3903, Oct. 2007.

[24] S. Korreman, H. Mostafavi, Q. T. Le and A. Boyer, "Comparison of respiratory surrogates for gated lung radiotherapy without internal fiducials," *Acta. Oncol.,* vol. 45, no. 7, pp. 935-942, Sep. 2006.

[25] Y. Tsunashima, T. Sakae, Y. Shioyama, K. Kagei, T. Terunuma, A. Nohtomi and Y. Akine, "Correlation between the respiratory waveform measured using a respiratory sensor and 3D tumor motion in gated radiotherapy," *Int. J. Radiat. Oncol. Biol. Phys,* vol. 60, no. 3, pp. 951-958, Nov. 2004.

[26] H. Yan, F. F. Yin, G. P. Zhu, M. Ajlouni and J. H. Kim, "The correlation evaluation of a tumor tracking system using multiple external markers," *Med. Phys.,* vol. 33, no. 11, pp. 4073-4084, Oct. 2006.

[27] F. Bergner, T. Berkus, M. Oelhafen, P. Kunz, T. Pan and M. Kachelrieß, "Autoadaptive phase-correlated (AAPC) reconstruction for 4D CBCT," *Med. Phys.,* vol. 36, no. 12, pp. 5695-5706, Dec. 2009.

[28] R. A. C. Siochi, "Deriving motion from megavoltage localization cone beam computed tomography scans," *Phys. Med. Biol.,* vol. 54, no. 13, p. 4195–4212, Jun. 2009.

[29] A. Kavanagh, P. M. Evans, V. N. Hansen and S. Webb, "Obtaining breathing patterns from any sequential thoracic x-ray image set," *Phys. Med. Biol.,* vol. 54, no. 16, pp. 4879-4888, Aug. 2009.

[30] I. Vergalasova, J. Cai and F. -F. Yin, "A novel technique for markerless, self-sorted 4D-CBCT: Feasibility study," *Med. Phys.,* vol. 39, no. 3, pp. 1442-1451, Mar. 2012.

[31] B. K. Horn and B. G. Schunk, "Determining optical flow," *Artificial intelligence,* vol. 17, pp. 185-203, Aug. 1981.

[32] C. Wachinger, M. Yigitsoy, E. -J. Rijkhorst and N. Navab, "Manifold learning for image-based breathing gating in Ultrasound and MRI," *Medical Image Analysis,* vol. 16, no. 4, pp. 806-818, May 2012.

[33] J. Chang, J. Sillanpaa, C. C. Ling, E. Seppi, E. Yorke, G. Mageras and H. Amols, "Integrating respiratory gating into a megavoltage cone-beam CT system," *Med Phys.,* vol. 33, no. 7, pp. 2354-2361, 2006.

[34] L. Dietrich, S. Jetter, T. Tucking, S. Nill and U. Oelfke, "Linac-integrated 4D cone beam CT: first experimental results," *Phys. Med. Biol.,* vol. 51, pp. 2939-2952, May 2006.

[35] P. Keall, "4-dimensional computed tomography imaging and treatment planning," *Semin Radiat Oncol.,* vol. 14, no. 1, pp. 81-90, 2004.

[36] T. Pan, "Comparison of helical and cine acquisitions for 4D-CT imaging with multislice CT," *Med. Phys.,* vol. 32, no. 2, pp. 627-634, 2005.

[37] E. Rietzel, T. Pan and G. T. Chen, "Four-dimensional computed tomography: image formation and clinical protocol," *Med. Phys.,* vol. 32, no. 4, pp. 874-889, 2005 .

[38] Z. Qi and G. H. Chen, "Performance studies of four-dimensional cone beam computed tomography," *Phys Med Biol.,* vol. 56, no. 20, pp. 6709-6721, 2011.

[39] D. Yan, F. Vicini, J. Wong and A. Martinez, "Adaptive radiation therapy," *Phys. Med. Biol.,* vol. 42, p. 123–132, 1997.

[40] P. Keall, G. Mageras, J. Balter, R. Emery, K. Forster, S. Jiang, J. Kapatoes, D. Low, M. Murphy, B. Murray, C. Ramsey, M. van Herk, S. Vedam, J. Wong and E. Yorke, "The management of respiratory motion in radiation oncology report of AAPM Task Group 76," *Med. Phys. ,* vol. 33, p. 3874–3900, 2006.

[41] S. Kriminski, M. Mitschke, S. Sorensen, N. Wink, P. Chow, S. Tenn and T. Solberg, "Respiratory correlated cone-beam computed tomography on an isocentric C-arm," *Phys. Med. Biol.,* vol. 50, p. 5263–5280, 2005.

[42] L. A. Feldkamp, L. C. Davis and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Am. A-Opt. Image Sci. Vis.,* vol. 1, pp. 612-619,, 1984.

[43] T. Li, A. Koong and L. Xing, "Enhanced 4D cone-beam CT with inter-phase motion model," *Med. Phys.,* vol. 34, no. 9, pp. 3688-3695, Sep. 2007.

[44] J. Lu, T. M. Guerrero, P. Munro, A. Jeung, P. -C. M. Chi, P. Balter, X. R. Zhu, R. Mohan and T. Pan, "Four-dimensional cone beam CT with adaptive gantry rotation and adaptive data sampling," *Med. Phys.,* vol. 34, no. 9, pp. 3520-3529, Sep. 2007.

[45] T. Li and L. Xing, "Optimizing 4D cone-beam CT acquisition protocol for external beam radiotherapy," *Int. J. Radiation Oncology Biol. Phys.,* vol. 67, no. 4, p. 1211–1219, 2007.

[46] T. Li, L. Xing, P. Munro, C. McGuinness, M. Chao, Y. Yang, L. B. and A. Koong, "Four-dimensional cone-beam computed tomography using an on-board imager," *Med. Phys.,* vol. 33, no. 10, pp. 3825-3833, 2006.

[47] S. Leng, J. Zambelli, R. Tolakanahalli, B. Nett, P. Munro, J. Star-Lack, B. Paliwal and G. H. Chen, "Streaking artifacts reduction in four-dimensional cone-beam computed tomography.," *Med Phys.,* vol. 35, no. 10, pp. 4649-59, 2008.

[48] S. Leng, J. Tang, J. Zambelli, B. Nett, R. Tolakanahalli and G. H. Chen, "High temporal resolution and streak-free four-dimensional cone-beam computed tomography," *Phys Med Biol. ,* vol. 53, no. 20, pp. 5653-5673, 2008.

[49] Q. Zhang, Y. Hu, F. Liu, K. Goodman, K. E. Rosenzweig and G. Mageras, "Correction of motion artifacts in cone-beam CT using a patient-specific respiratory motion model," *Med Phys.,* vol. 37, no. 6, pp. 2901-2909, 2010 .

[50] E. Y. Sidky, C. -M. Kao and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *J. X-Ray Sci. Technol. ,* vol. 14, p. 119–139 , 2006.

[51] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.,* vol. 53, no. 17, p. 4777–4807, 2008.

[52] S. Leng, J. Tang, J. Zambelli, B. Nett, T. R. and G.-H. Chen, "High temporal resolution and streak-free four-dimensional cone-beam computed tomography," *Phys. Med. Biol.,* vol. 53, no. 20, p. 5653–5673, 2008.

[53] G. -H. Chen, J. Tang and S. Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med. Phys.,* vol. 35, no. 2, pp. 660-663, Feb. 2008.

[54] G. -H. Chen, J. Tang and S. Leng, "Prior Image Constrained Compressed Sensing (PICCS)," San Jose, CA, 2008.

[55] E. Y. Sidky, Y. Duchin, X. Pan and C. Ullberg, "A constrained, total-variation minimization algorithm for low-intensity x-ray CT," *Med. Phys.,* vol. 38, no. 7, pp. S117-S125, Jul. 2011.

[56] T. Li, E. Schreibmann, Y. Yang and L. Xing, "Motion correction for improved target localization with on-board cone-beam computed tomography," *Phys. Med. Biol.,* vol. 51, no. 2, p. 253–267, 2006.

[57] R. Zeng, J. Fessler and J. Balter, "Estimating 3-D respiratory motion from orbiting views by tomographic image registration," *IEEE Trans. Med. Imaging,* vol. 26, p.

153–163, 2007.

[58] S. Rit, J. W. H. Wolthaus, M. van Herk and J. -J. Sonke, "On-the-fly motion-compensated cone-beam CT using an a priori model of the respiratory motion," *Med. Phys.,* vol. 36, no. 6, pp. 2283-2296, Jun. 2009.

[59] S. Rit, D. Sarrut and L. Desbat, "Comparison of analytic and algebraic methods for motion-compensated cone-beam ct reconstruction of the thorax," *IEEE Trans. Med. Imaging,* vol. 28, no. 10, p. 1513–1525, 2009.

[60] R. Li, X. Jia, J. H. Lewis, X. Gu, M. Folkerts, C. Men and S. B. Jianga, "Real-time volumetric image reconstruction and 3D tumor localization based on a single x-ray projection image for lung cancer radiotherapy," *Med. Phys.,* vol. 37, no. 6, pp. 2822-2826, Jun. 2010.

[61] R. Li, J. H. Lewis, X. Jia, X. Gu, M. Folkerts, C. Men, W. Song and S. Jiang, "3D tumor localization through real-time volumetric x-ray imaging for lung cancer radiotherapy.," *Med. Phys.,* vol. 38, p. 2783–2794, 2011.

[62] M. Ahmad, P. Balter and T. Pan, "Four-dimensional volume-of-interest reconstruction for cone-beam computed tomography-guided radiation therapy," *Med. Phys.,* vol. 38, p. 5646–5656, 2011.

[63] F. Bergner, T. Berkus, M. Oelhafen, P. Kunz, T. Pan, R. Grimmer, L. Ritschl and M. Kachelrieß, "An investigation of 4D cone-beam CT algorithms for slowly rotating scanners," *Med. Phys.,* vol. 37, no. 9, pp. 5044-5053, Sep. 2010.

[64] X. Jia, Y. Lou, B. Dong and Z. J. S. Tian, "4D computed tomography reconstruction from few-projection data via temporal non-local regularization," in *MICCAI'10 Proceedings of the 13th international conference on Medical image computing and computer-assisted intervention: Part I*, Berlin, 2010.

[65] Z. Tian, X. Jia, B. Dong, Y. Lou and S. B. Jiang, "Low-dose 4DCT reconstruction via temporal nonlocal means," *Med. Phys.,* vol. 38, no. 3, pp. 1359-1365, 2011.

[66] X. Jia, Z. Tian, Y. Lou, J. -J. Sonke and S. B. Jiang, "Four-dimensional cone beam CT reconstruction and enhancement using a temporal nonlocal means method," *Med. Phys.,* vol. 39 , pp. 5592-5602, 2012.

[67] M. Bertram, G. Rose, D. Schafer, J. Wiegert and T. Aach, "Directional Interpolation of Sparsely Sampled Cone-Beam CT Sinogram Data," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Arlington, VA, 2004.

[68] M. Bertram, J. Wiegert, D. Schafer, T. Aach and G. Rose, "Directional View Interpolation for Compensation of Sparse Angular Sampling in Cone-Beam CT," *IEEE Trans. Med. Imag.,* vol. 28, no. 7, pp. 1011-1022, Jul. 2009.

[69] G. H. Weiss, A. J. Talbert and R. A. Brooks, "The use of phantom views to reduce CT streaks due to insufficient angular sampling," *Phys. Med. Biol.,* vol. 27, no. 9, p. 1151–1162, 1982.

[70] R. R. Galigekere, K. Wiesent and D. W. Holdsworth, "Techniques to alleviate the effects of viewaliasing artifacts in computed tomography," *Med. Phys.,* vol. 26, no. 6, p. 896–904 , 1999.

[71] T. M. Lehmann, C. Gönner and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Trans. Med. Imag.,* vol. 18, no. 11, p. 1049–

1075, Nov. 1999.

[72] T. M. Lehmann, C. Gonner and K. Spitzer, "Addendum: B-spline interpolation in medical image processing," *IEEE Trans. Med. Imag.,* vol. 20, no. 7, p. 660–665, Jul. 2001.

[73] H. Kostler, M. Prummer, U. Rude and J. Hornegger, "Adaptive variational sinogram interpolation of sparsely sampled CT data," in *The 18th International Conference on Pattern Recognition (ICPR'06) 2006*, Hong Kong, 2006.

[74] P. V. Sankar and L. A. Ferrari, "Simple Algorithms and Architectures for B-spline Interpolation," *IEEE Trans. Pattern Anal. Mach. Intell,* vol. 10, no. 2, pp. 271-276, 1988.

[75] A. K. Jain, Fundamentals of Digital Image Processing, Englewood Cliffs, NJ : Prentice-Hall, 1989.

[76] P. Thévenaz, T. Blu and M. Unser, "Interpolation revisited," *IEEE Trans. Med. Imag.,* vol. 19, no. 7, p. 739–758, Jul. 2000.

[77] G. J. Grevera and J. K. Udupa, "An objective comparison of 3-D image interpolation methods," *IEEE Trans. Med. Imag.,* vol. 17, no. 4, p. 642–652, Aug. 1998.

[78] A. Goshtasby, D. A. Turner and L. V. Ackerman, "Matching of tomographic slices for interpolation," *IEEE Trans. Med. Imag.,* vol. 11, no. 4, p. 507–516, Dec. 1992.

[79] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Med. Imag.,* vol. 10, no. 10, p. 1521–1527, Oct. 2001.

[80] J. Hladuvka and E. Gröller, "Direction-driven shape-based interpolation of volume data," in *in Proc. Vis., Model., Visualizat.*, Stuttgart, 2001.

[81] C. Penßel, M. Kachelrieß, M. Knaup and W. A. Kalender, "Azimuthal interpolation and noise reduction," *IEEE Nucl. Sci. Symp. Conf. Rec.,* vol. 4, p. 1946–1949, 2005.

[82] A. Garde, L. Sörnmo, R. Jané and B. F. Giraldo, "Correntropy-Based spectral characterization of respiratory patterns in patients with Chronic Heart Failure," *IEEE Trans. Biomed. Eng.,* vol. 57, no. 8, pp. 1964-1972, Aug. 2010.

[83] C. M. Ionescu, I. Muntean, J. A. Tenreiro-Machado, R. De Keyser and M. Abrudean, "A theoretical study on modeling the respiratory tract with ladder networks by means of intrinsic fractal geometry," *IEEE Trans. Biomed. Eng.,* vol. 57, no. 2, pp. 246-253, Feb. 2010.

[84] A. Yadollahi and Z. M. Moussavi, "A robust method for estimating respiratory flow using tracheal sounds entropy," *IEEE Trans. Biomed. Eng.,,* vol. 53, no. 4, pp. 662-668, Apr. 2006.

[85] M. Y. Wang, C. R. Maurer Jr., J. M. Fitzpatrick and R. J. Maciunas, "An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head," *IEEE Trans. Biomed. Eng.,* vol. 43, no. 6, pp. 627-637, Jun. 1996.

[86] L. Pan, J. L. Prince, J. A. C. Lima and N. F. Osman, "Fast tracking of cardiac motion Using 3D-HARP," *IEEE Trans. Biomed. Eng.,* vol. 52, no. 8, pp. 1425-1435, 2005.

[87] A. S. Naini, T.-Y. Lee, R. V. Patel and A. Samani, "Estimation of lung's air volume and its variations throughout respiratory CT image sequences," *IEEE Trans. Biomed. Eng.,* vol. 58, no. 1, pp. 152-158, Jan. 2011.

[88] C. Tomasi and T. Kanade, "Detection and Tracking of Point Features," Carnegie Mellon Uni.,, Pittsburgh, PA,, 1991.

[89] J. Shi and C. Tomasi, "Good features to track," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recogn.*, New York, 1994.

[90] R. R. Sokal and P. H. A. Sneath, Principles of Numerical Taxonomy, San Francisco, CA: Freeman, 1963.

[91] R. Dubes and A. K. Jain, "Validity Studies in Clustering Methodologies," *Pattern Recognition,* vol. 11 , no. 4, pp. 235-254, 1979.

[92] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," *Int. J. Computer Vision,* vol. 9, no. 2, pp. 137-154, Nov. 1992.

[93] A. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures," *Analytical Chemistry,* vol. 36, p. 1627–1639, 1964.

[94] J. Barron, D. Fleet and S. Beauchemin, "Performance of optical flow techniques," *Int. J. Comp. Vis.,* vol. 12, no. 1, p. 43–77, 1994.

[95] J. -Y. Bouguet, "Pyramidal Implementation of the Lucas-Kanade Feature Tracker,," OpenCV Documentation, Microprocessor Research Lab, Intel Corporation, 2000.

[96] "Online help of hierarchal clustering- Release R2012a," MATLAB, 2012. [Online]. Available: http://www.mathworks.com. [Accessed Nov. 2012].

[97] R. I. Berbeco, S. Nishioka, H. Shirato and S. B. Jiang, "Residual motion of lung tumors in end-of-inhale respiratory gated radiotherapy based on external surrogates," *Med. Phys.,* vol. 33 , no. 11, pp. 4149-4156, 2006.

[98] D. Sun, S. Roth and M. J. Black, "Secrets of optical flow estimation and their principles," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2010.

[99] A. K. Jain and R. C. Dubes, Algorithms for Clustering Data, Upper Saddle River, NJ: Prentice Hall Inc., 1988.

[100] J. W. Sammon, "A nonlinear mapping for data structure analysis," *IEEE Trans. Computers,* Vols. C-18, no. 5, pp. 401-409, 1969.

[101] R. Gordon, R. Bender and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography," *J. Theor. Biol.,* vol. 29, p. 471–481 , 1970.

[102] M. Brehm, P. Paysan, M. Oelhafen, P. Kunz and M. Kachelrieß, "Self-adapting cyclic registration for motion-compensated cone-beam CT in image-guided radiation therapy," *Med. Phys.,* vol. 39, no. 12, pp. 7603-7618, 2012.

[103] J. Pluim, J. Maintz and M. Viergever, "Mutual-information–based registration of medical images: A survey," *IEEE Trans. Med. Imaging ,* vol. 22, no. 8, pp. 986-1004, 2003.

[104] W. P. Segars, D. S. Lalush and T. M. W., "A realistic spline-based dynamic heart phantom," *IEEE Trans. Nucl. Sci.,* vol. 46, pp. 503-509, 1999.

[105] W. P. Segars, M. M., B. T.J., F. E. C. and T. B. M., "Realistic CT simulation using the 4D XCAT phantom," *Med Phys,* vol. 35, no. 8, pp. 3800-3808, Aug 2008 .

[106] W. R. Sung, E. K. Kang and J. S. Choi, "Adaptive motion estimation technique for motion compensated interframe interpolation," *IEEE trans. Consumer Electronics,* vol. 45, no. 3, pp. 753-761, Aug. 1999.

[107] R. Lu, P. Marziliano and C. -H. Thng, "Comparison of scene-based interpolation methods applied to CT abdominal images," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, 2004.