

# A Novel Method of Cone Beam CT Projection Binning based on Image Registration

Seonyeong Park, *Student Member, IEEE*, Siyong Kim, Byongyong Yi, Geoffrey Hugo, H. Michael Gach, and Yuichi Motai, Jr., *Senior Member, IEEE*

**Abstract**—Accurate sorting of beam projections is important in four-dimensional Cone Beam Computed Tomography (4D CBCT) to improve the quality of the reconstructed 4D CBCT image by removing motion-induced artifacts. We propose Image Registration-based Projection Binning (IRPB), a novel markerless binning method for 4D CBCT projections, which combines Intensity-based Feature Point Detection (IFPD) and Trajectory Tracking using Random sample consensus (TTR). IRPB extracts breathing motion and phases by analyzing tissue feature point trajectories. We conducted experiments with two phantom and six patient datasets including both regular and irregular respiration. In experiments, we compared the performance of the proposed IRPB, Amsterdam Shroud method (AS), Fourier Transform-based method (FT), and Local Intensity Feature Tracking (LIFT) method. The results showed that the average absolute phase shift of IRPB was 3.74 projections and 0.48 projections less than that of FT and LIFT, respectively. AS lost the most breathing cycles in the respiration extraction for the five patient datasets, so we could not compare the average absolute phase shift between IRPB and AS. Based on the Peak Signal-to-Noise Ratio (PSNR) of the reconstructed 4D CBCT images, IRPB had 5.08 dB, 1.05 dB, and 2.90 dB larger PSNR than AS, FT, and LIFT, respectively. The average Structure SIMilarity Index (SSIM) of the 4D CBCT image reconstructed by IRPB, AS, and LIFT were 0.87, 0.74, 0.84, and 0.70, respectively. These results demonstrated that IRPB has superior performance to the other standard methods.

**Index Terms**—Cone beam computed tomography (CBCT), image registration, projection binning, respiratory signal.

## I. INTRODUCTION

TUMORS in moving organs require dynamic imaging techniques, such as four-dimensional Cone Beam Computed Tomography (4D CBCT) [1]-[4], because respiratory motion mainly causes inferior image quality, i.e., burry edges and inaccurate shapes, for the tumor in thoracic and upper abdominal regions [1], [5]-[9]. These dynamic techniques

This work was supported in part by Washington University in St. Louis Department of Radiation Oncology and National Institutes of Health National Cancer Institute under Grant R01 CA159471, in part by the National Science Foundation CAREER under Grant 1054333, in part by the National Center for Advancing Translational Sciences the CTSA Program under Grant UL1TR000058, in part by the Center for Clinical and Translational Research Endowment Fund within Virginia Commonwealth University (VCU) through the Presidential Research Incentive Program, and in part by VCU Graduate School Dissertation Assistantship.

S. Park and Y. Motai are with the Department of Electrical and Computer Engineering, Virginia Commonwealth University, Richmond, VA 23284 USA (e-mail: parks45@vcu.edu; ymotai@vcu.edu).

are able to reduce the estimation error of the target tumor's location by compensating for respiratory motion on medical images. 4D CBCT and other motion-compensated reconstruction methods have been developed to overcome motion-induced artifacts. Some of these methods are based on phase binning [2]-[4], [9]-[12] and the others use respiration models to deform either the reconstructed volume or the back-projection operation [13], [14]. The reconstructed image quality of all of these methods directly relies on the performance of respiratory signal extraction because all of these techniques require tagging raw data (e.g., radiographic projections) with a respiratory phase or state [3]-[11], [13].

Previous methods to acquire respiratory phases used one or more of the following approaches [11]: 1) additional equipment [15]-[18], 2) positional changes of a diaphragm [4], [9], [12], [19], and 3) projection [6], [20]-[22]. These three approaches are compared in Table I.

First, additional equipment-based methods employ either

TABLE I  
PREVIOUS BREATHING PHASE EXTRACTION METHODS

Method	Procedure	Limitation
Additional equipment [15]-[18]	Tracking external source, skin markers attached on the body, or internal source, radio-opaque markers implanted near the lesion	Correlation between external makers' motion and tumor's movement, invasive setup procedure, and expensive cost
Diaphragm [4], [9], [12], [19]	Observing changes of the diaphragm on images	Projection view restriction that the diaphragm should be always visible, correlation between target tumor's movement and the diaphragm's motion
Projection [6], [20]-[22]	Analyzing the intensity fluctuation, applying Principle Component Analysis/Fourier Transform, or reconstructing trajectories of tissue feature points in 3D	Restriction of the projection view and supplemental tasks depending on each patient: manual clustering of 4D CBCT projections as subsets and signal/image processing

S. Kim and G. Hugo are with the Department of Radiation Oncology, Virginia Commonwealth University, Richmond, VA 23298 USA (e-mail: siyong.kim@vcuhealth.org; gdhugo@vcu.edu).

B. Yi is with the Department of Radiation Oncology, University of Maryland School of Medicine, Baltimore, MD 21201 USA (e-mail: byi@umm.edu).

H. M. Gach is with the Departments of Radiation Oncology and Radiology, Washington University in St. Louis, St. Louis, MO 63110 USA (e-mail: gachhm@wustl.edu).

Copyright (c) 2017 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

external or internal tools such as skin markers, abdominal belts, spirometry, or radio-opaque markers [15]-[18]. To obtain breathing signals, the external source is attached to the patient's body and the internal source is implanted near the tumor. Then, locational changes of these sources, i.e., skin and radio-opaque markers, are tracked [15]-[18]. Measurement processes utilizing skin markers are relatively simple but a supplementary task is required to correlate the external marker and tumor motion [15]. The internal source method is invasive and expensive [6], [15]. Implantation of radio-opaque makers inside the patient's body is invasive, requires physician involvement, and can affect clinical workflow.

Second, the diaphragm position-based methods observe changes of the diaphragm on CBCT projections with the assumption that the changes are related to respiration [4], [9], [12], [19], [23]. Accordingly, the diaphragm region must be shown in all CBCT projections to estimate breathing motion. However, some commonly used CBCT systems have a limited longitudinal field of view [4], [9], [12], [19]. The method will have poor accuracy if the target tumor and the diaphragm have different respiratory motion patterns [6], [23].

Third, the projection-based methods include the Amsterdam Shroud method (AS) [20], the Amsterdam Shroud image-based Local Principal Component Analysis (LPCA) [22], the Fourier Transform-based method (FT) [21], and Local Intensity Feature Tracking (LIFT) [6]. AS extracts the respiratory signal from an Amsterdam Shroud image that consists of average intensity values for each horizontal row of the CBCT projection. LPCA uses a Principal Component Analysis (PCA)-based sliding window approach to obtain the breathing signal on the Amsterdam Shroud image. The FT finds the breathing signal by analyzing the CBCT projections based on Fourier transform theory. LIFT allocates feature points that are distributed uniformly on the 4D CBCT images, and breathing signals are extracted by recreating three-dimensional (3D) movements of each feature point from multiple images. A main advantage of the projection-based methods is that they require only CBCT projections to extract breathing phases, so they overcome the limitation of the equipment-based methods (i.e., expensive setup procedures). However, the existing projection-based methods are restricted by the projection views, as they need to set the Region Of Interest (ROI) in their process. Besides, they contain supplemental tasks, such as manual clustering of the 4D CBCT projections as subsets and signal/image processing, depending on each patient.

In this paper, we propose a novel CBCT projection binning method without the use of markers, called Image Registration-based Projection Binning (IRPB), which is a tissue feature-based approach. The proposed method selects tissue feature points near the edge areas only based on image intensity of adjacent pixels. The use of selective tissue feature points can enhance the computational efficiency of IRPB. Then, the trajectories of feature points are tracked by applying an image registration algorithm, the Random Sample Consensus (RANSAC) [24]. From the obtained trajectories, respiratory phases are calculated via periodicity analysis of transitional patterns of the feature point positions.

The contributions of this paper are twofold. First, IRPB is expected to provide better image guidance for radiation therapy. This is because IRPB provides higher quality reconstructed 4D CBCT images compared to the existing methods—AS, FT, and LIFT—by improving the accuracy of the assigned breathing phases. Second, the proposed method achieves automatic 4D CBCT projection binning, requiring no internal or external marker and no restriction of the projection views. We do not need to find the correlation between the marker's and tumor's motion, we do not need to include the diaphragm in every projection view, and we do not need any manual or arbitrary process to detect breathing signals.

## II. IMAGE REGISTRATION-BASED PROJECTION BINNING

IRPB consists of the following three steps: *A.* intensity-based feature point detection; *B.* feature point trajectory tracking; and *C.* respiratory signal extraction and projection binning. We describe these three steps in this section.

### A. Intensity-based Feature Point Detection

We adjust the image intensity values of the overall 4D CBCT projection in advance to utilize the full scale of intensity. We linearly scale the lowest 1% of image intensity values to the new value, i.e., '0,' and the highest 1% to '1,' thus, the contrast of the overall image increases. The image intensity adjustment can be formulated as  $(I_{cutoff} - l_{0.01}) / (h_{0.99} - l_{0.01})$ , where  $l_{0.01}$  presents the lowest 1% of image intensity values,  $h_{0.99}$  indicates the highest 1% of image intensity values, and  $I_{cutoff}$  is an image intensity array bounded by  $l_{0.01}$  and  $h_{0.99}$  (i.e.,  $l_{0.01} \leq I_{cutoff} \leq h_{0.99}$ ). This also allows all 4D CBCT projections to have the identical intensity range of the whole image by revising their contrast. Fig. 1 illustrates the procedure of the proposed Intensity-based Feature Point Detection (IFPD).

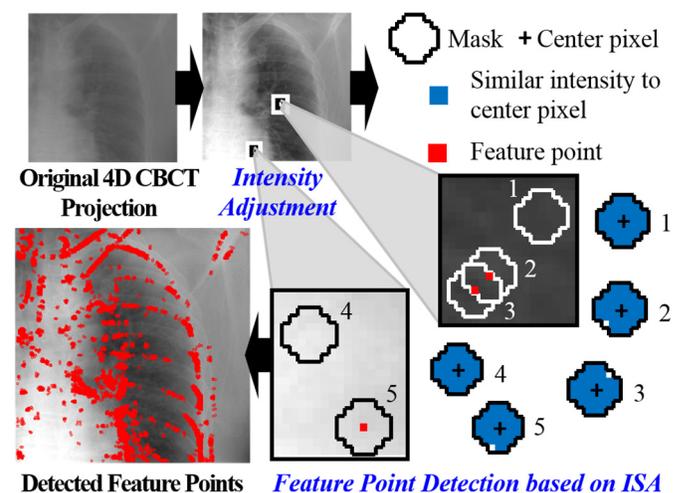


Fig. 1. Intensity-based Feature Point Detection (IFPD): A pixel circle is a mask covering adjacent pixels, a symbol '+' is a center pixel of the mask, a blue colored area presents pixels with similar intensities to the center pixel, and a red colored pixel is a selected feature point.

We assume that feature points related to breathing motion are densely located in areas near edges on the 4D CBCT projections, which was observed in the interim findings [6]. Therefore, we choose tissue feature points only near edges.

Selection of feature points is based on intensity variation within a circular pixel mask determined based on Bresenham circle [25]. In the mask, the number of pixels with intensities similar to the center pixel is calculated and denoted as Intensity Similarity of Adjacent pixels (ISA). A center pixel of the mask is determined to be a feature point depending on ISA.

In Fig. 1, a circle is a mask covering adjacent pixels, a symbol '+' is a center pixel of the mask, a blue colored area presents pixels with similar intensity to the center pixel, and a red colored pixel is a selected feature point. The value of ISA is computed as:

$$S'(\mathbf{a}_0) = \sum_a \{1 - H(|I(\mathbf{a}) - I(\mathbf{a}_0)| - \tau)\}, \quad (1)$$

where  $\mathbf{a}_0$  is the center pixel position,  $\mathbf{a}$  is the position of an adjacent pixel in the mask,  $I(\cdot)$  is an intensity function,  $H(\cdot)$  is a unit step function, and  $\tau$  is an intensity difference threshold. In (1), when all the pixels in a mask have the similar intensity based on  $\tau$ , the H term becomes '0' and ISA is calculated as the largest value, i.e., the total number of pixels in the mask  $\psi$ . This means that a center pixel is located far from the edge, so the center pixel is discarded in the feature point selection. Contrariwise, a center pixel near an edge has a smaller ISA value than  $\psi$ , and it is chosen as a feature point.

As shown in (1), the feature point selection depends on the intensity value  $I(\cdot)$  and parameters—the mask radius  $r$  and the intensity difference threshold  $\tau$ . It is possible to determine the feature points by tuning the parameters without an intensity adjustment. In this case, however, the parameters should be set according to each patient/projection view since it cannot be assured that every projection dataset has similar intensity contrast. If we adjust the image contrast of any projection dataset to a fixed range, then we do not need to tune the parameters each time. An algorithm to seek the optimal parameters can be another research topic. Thus, IFPD includes an intensity adjustment to avoid manual tasks for setting parameters for each operation.

Our goal is to select points that are likely to represent respiration with as high a ratio as possible. Not only do these targeted points include edges, but they are also located in the region near edges as assumed above. In the proposed IFPD, the mask radius  $r$  determines the number and the range of feature points. The larger  $r$  selects more feature points in a wider range. An appropriate choice for the mask radius  $r$  boosts the efficiency and performance of the binning method. The previous optical flow-based binning methods [26] require substantial computation for uniformly distributed points: tracking their locational changes over projections and extracting respiratory motion. However, the proposed method reduces the amount of computation for binning by selecting feature points locally according to  $r$ . In addition, a suitable choice of  $r$  increases the rate of the breathing motion-related feature points among all of the selected points, as well as contribute to improved accuracy performance of binning.

### B. Feature Point Trajectory Tracking

The trajectory chase of the detected feature points requires identifying corresponding feature points and tracking their

locational changes over sequential 4D CBCT projections. The problem is that an identical feature point may not exist in all 4D CBCT projections, because each 4D CBCT image is projected from different views due to the scanner rotation. To overcome this problem, we can use image registration methods to acquire feature point trajectories. The image registration aligns images with the same scene taken from different views or at different times, and has been widely used in medical imaging [3], [24], [27]–[29].

We apply one of the image registration methods, RANSAC [24], to track the locational changes of feature points in consecutive 4D CBCT images, and call this new method Trajectory Tracking using RANSAC (TTR). RANSAC has shown precise point-matching performance by iterating error correction and minimizing outliers, and many studies have proven its robustness [29]–[32]. To discover the periodic nature of breathing motion more effectively, we solely focus on Superior-Inferior (SI) coordinates of the feature points. The identified correspondence presents only the locational change of the feature point between two consecutive projections, but it does not clearly expose respiration-related regularity. As aforementioned, it is not always possible to find a trajectory connected from the first to the last projection because the identical feature point does not appear in every projection. The previous 4D CBCT projection binning method, LIFT [6], arbitrarily groups 701 4D CBCT projections into four portions, each with a different number of projections—Group1 and Group2 with 150 projections; Group3 with 250 projections; and Group4 with 251 projections—in order to track trajectories in each group belonging to the limited range of the shooting angles. Manually and arbitrarily selected, each group has an overlap of 50 projections [6].

To avoid manual and arbitrary grouping, we find the locational changes of feature points that are common to at least three consecutive projections. As shown in Fig. 2, we specifically find common corresponding feature points in the

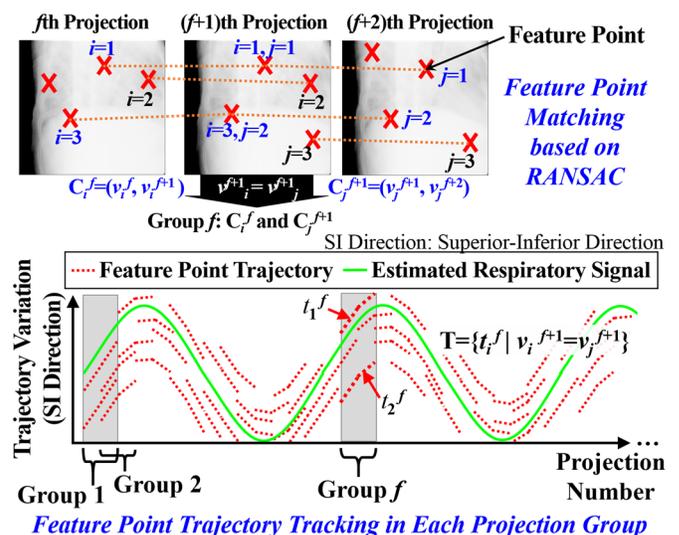


Fig. 2. Feature point trajectory tracking: corresponding feature point matching  $C_i^f$  by RANSAC [24], and tracking feature point trajectory  $t_i^f$  in three consecutive projections, where  $f$  is a projection number, and  $i$  is an index for trajectories in  $f$ th,  $(f+1)$ th, and  $(f+2)$ th projections.

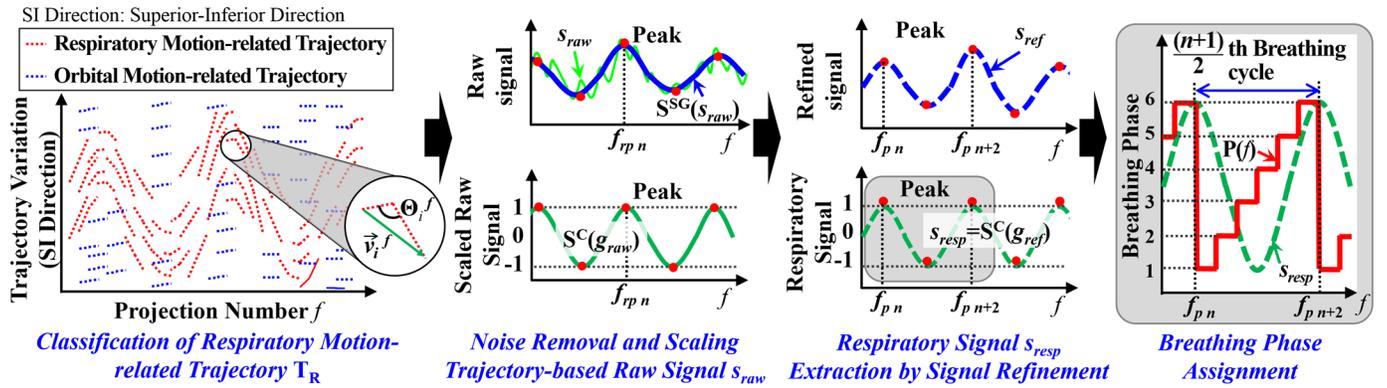


Fig. 3. Respiratory signal extraction and projection binning: classification of the respiratory motion-related trajectory  $\mathbf{T}_R$  using PCA [33], smoothing and scaling a discrete derivative raw signal  $s_{raw}^f$  obtained from  $\mathbf{T}_R$ , signal refinement for respiratory signal  $s_{resp}$  extraction, and breathing phase assignment according to  $s_{resp}$  by Eq. (2)-(6). Here,  $g_{ref}$  represents a signal scaled to  $[-1, 1]$  according to peak values of  $s_{ref}$  for all projections,  $f_{rp n}$  represents a projection number corresponding to a peak value of  $s_{raw}^f$  and  $S^C(g_{raw})$ , and  $f_{p n}$  represents a projection number corresponding to a peak value of  $s_{ref}$  and  $s_{resp}$ .

$f$ th and  $(f+1)$ th projections denoted as  $C^f=(v_i^f, v_i^{f+1})$  and those in the  $(f+1)$ th and  $(f+2)$ th projections denoted as  $C^{f+1}=(v_j^{f+1}, v_j^{f+2})$ . Afterwards, the matched feature point sets between adjacent  $C_i^f$  and  $C_j^{f+1}$  are determined as trajectories  $\mathbf{T}=\{t_i^f | v_i^{f+1}=v_j^{f+1}\}$ , where  $t_i^f$  is the  $i$ th trajectory in the  $f$ th projection, defined as  $t_i^f=(t_{i,1}^f, t_{i,2}^f, t_{i,3}^f)$ . Then, the breathing signal is extracted by using the three projection-long connected trajectories  $t_i^f$  of such featured points.

### C. Respiratory Signal Extraction and Projection Binning

Breathing signal extraction from the feature point trajectories influences the accuracy performance of 4D CBCT projection binning. In practice, not all tracked trajectories show a breath-like oscillating pattern, even if we consider the SI direction only. At some feature points that are not relevant to respiration, the trajectories occur due to movement of the x-ray source orbiting the patient during scanning, which is useless for extracting respiration patterns. For example, some of the selected feature points, such as the points near the apex of lungs, move toward SI direction over time. They may create non-periodic horizontal trajectories. To estimate a patient's respiration, we need to filter out the orbital motion-related trajectories. It is better that the overall feature point trajectories include enough breathing motion-related trajectories such that every projection has at least one respiratory motion-related trajectory. Furthermore, reliable classification of the orbital motion-related trajectories facilitates precise extraction of breathing signals.

To exclude the orbital motion-related trajectories, we analyzed patterns of the feature point trajectories by using PCA [33], one of the commonly used statistical data analysis tools. PCA transforms a dataset with a large number of interrelated variables to an uncorrelated set of variables. The criteria used for PCA, i.e., the input for PCA, consists of the angle  $\Theta_i^f$  and the trajectory vector  $\vec{v}_i^f$  for each trajectory  $t_i^f$ , that are assumed to be involved in respiration behaviors. As shown in Fig. 3,  $\Theta_i^f$  is the included angle among three points on each trajectory  $t_i^f$ , and the trajectory vector  $\vec{v}_i^f$  is determined by the first and the last (i.e., the third) points on each trajectory  $t_i^f$ . Then, a respiratory motion-related trajectory set,  $\mathbf{T}_R$  over three consecutive projections, is discriminated by PCA.

The three projection-long trajectories  $\mathbf{T}_R$  are pieces of the patient's breathing signal distributed over the 4D CBCT projections, so we need to analyze a common tendency with the periodicity among them. We cannot derive the respiration signal by simply averaging  $\mathbf{T}_R$  because trajectories are located at different positions, and each projection has a different number of trajectories. To overcome these limitations, we analyzed a discrete derivative raw signal  $s_{raw}^f$  i.e., the average of the discrete derivatives of the trajectories for a projection number variable:

$$s_{raw}^f = \frac{1}{n(\mathbf{T}_R^f)} \sum_{i=1}^{n(\mathbf{T}_R^f)} (t_i^f)', \quad (2)$$

where  $t_i^f$  is the discrete derivative of  $t_i^f$ , defined as  $t_i^f'=(t_{i,2}^f-t_{i,1}^f, t_{i,3}^f-t_{i,2}^f)$ .

Noise is reduced by smoothing  $s_{raw}^f$  with a Savitzky-Golay (SG) filter [34], and we represent the discrete derivative raw signal  $s_{raw}^f$  after applying the SG filter as  $S^{SG}(s_{raw})$  as shown in Fig. 3. In principle, the discrete derivative of the trajectory is similar to a sinusoidal function, which becomes '0' at both ends of a breathing cycle, i.e., end inhalation and end exhalation. These ends correspond to the peaks of the trajectory. The discrete derivative of the trajectory has its peaks and valleys corresponding to the mid phases of either inhalation or exhalation. Therefore, it is possible to distinguish each breathing cycle by searching either points with '0' value or local maxima/minima. In reality, however, a discrete derivative function may show irregular non-smooth behavior and further refinement may be required. For precise identification of the breathing cycle, we refine  $s_{raw}^f$  into the peak-prominent signal with clear curves, using a cubic spline interpolation [35]:

$$S^C(g_{raw}) = \sum_{i=1}^{n(f_p)-1} [a_i(f-f_{pi})^3 + b_i(f-f_{pi})^2 + c_i(f-f_{pi}) + d_i], \quad (3)$$

where  $g_{raw}$  represents a signal scaled to  $[-1, 1]$  according to peak values of  $S^{SG}(s_{raw})$  for all projections, and  $f_{pi}$  indicates the projection number corresponding to the  $i$ th peak value (i.e., '-1' or '1'). The peak values are the local maxima/minima of  $S^{SG}(s_{raw})$  that are either greater/smaller than their neighboring values. Their projection numbers are defined as  $f_{pi}$ . In (3),  $a_i$ ,  $b_i$ ,  $c_i$ , and  $d_i$  are the cubic polynomial coefficients:

$$\begin{aligned}
 a_i &= \frac{1}{(f_{p\ i+1} - f_{p\ i})^2} \left\{ 2 \frac{g_{raw}^{f_{p\ i}} - g_{raw}^{f_{p\ i+1}}}{f_{p\ i+1} - f_{p\ i}} + g_{raw}^{f_{p\ i}} + g_{raw}^{f_{p\ i+1}} \right\} \\
 b_i &= \frac{1}{f_{p\ i+1} - f_{p\ i}} \left\{ 3 \frac{g_{raw}^{f_{p\ i+1}} - g_{raw}^{f_{p\ i}}}{f_{p\ i+1} - f_{p\ i}} - 2g_{raw}^{f_{p\ i}} - g_{raw}^{f_{p\ i+1}} \right\} \\
 c_i &= g_{raw}^{f_{p\ i}} \quad \text{and} \quad d_i = g_{raw}^{f_{p\ i}}
 \end{aligned} \quad (4)$$

$S^C(g_{raw})$  from  $S^{SG}(s_{raw})$  is considered as an amplified variation of  $\mathbf{Tr}$ , so we can compute the refined signal  $s_{ref}$  of  $\mathbf{Tr}$  by the cumulative sum of  $S^C(g_{raw})$  as follows:

$$s_{ref}^f = \sum_{i=1}^f S_i^C(g_{raw}). \quad (5)$$

We repeat (3) for the refined signal  $s_{ref}$  to obtain a peak-prominent respiratory signal  $s_{resp}$ . With projection number corresponding to peak values  $f_p$ , the respiratory phase of 4D CBCT projection  $f$  is defined as:

$$P(f) = \left[ \frac{c(f - f_{p_n})}{f_{p_{n+2}} - f_{p_n}} \right], \quad f_{p_n} < f < f_{p_{n+2}}, \quad (6)$$

where  $c$  is the number of breathing phases, and  $f_{p_n}$  and  $f_{p_{n+2}}$  are projection numbers corresponding to the beginning and the end of the  $\{(n+1)/2\}$ th breathing cycle, respectively. Based on  $P(f)$ , 4D CBCT projections are sorted into each phase bin before their reconstruction.

### III. EXPERIMENTAL RESULTS

We used a PC with Inter Core i7-6790 CPU 3.60GHz CPU and 16GB RAM in the experiment. The descriptions of experimental datasets are given in the following subsection *A*. The experimental results of the proposed IFPD, TTR, and the breathing phase assignment are presented in subsection *B*, *C*, and *D*, respectively.

#### A. Experimental Datasets

##### 1) Phantom Datasets

4D extended Cardiac-Torso (XCAT) phantom datasets used for validation of the proposed method offer a virtual model of the human torso with cardiac and respiratory motions, used in many previous studies [36]. The size of the generated phantom model was  $256 \times 256 \times 201$  voxels. We produced a regular breathing phantom dataset with 5 s of respiration and an irregular phantom dataset with 5 s and 2.5 s long breathing durations. To get the 4D CBCT images of the phantom datasets, the Reconstruction ToolKit (RTK) was used [37]. The size of the created 4D CBCT projections are  $256 \times 256$  pixels. We set the detector panel size as  $256 \text{ mm} \times 256 \text{ mm}$ , and the distance between the source and the detector was 1500 mm. In addition, the distance between the source and the object was 1100 mm. The full scan was conducted with a single  $360^\circ$  gantry rotation, and 720 projections were created with  $0.5^\circ$  angle increments. Projections were sorted into six respiratory phase bins.

##### 2) Patient Datasets

4D CBCT datasets of six subjects were collected from Virginia Commonwealth University Massey Cancer Center under an institution review board approved study after patients provided informed consent and were enrolled on the study. The

TABLE II  
PATIENT DATASETS

Patient #	# of projections	Projection size (pixel)	Calibrated pixel size (mm/pixel)	Phase binning ground truth
1	701	$512 \times 512$	0.518	Diaphragm
2	2369	$1024 \times 768$	0.388	Marker
3	2413	$1024 \times 768$	0.388	Diaphragm and Marker
4	2386	$1024 \times 768$	0.388	Marker
5	2460	$1024 \times 768$	0.388	Diaphragm and Marker
6	2447	$1024 \times 768$	0.388	Diaphragm and Marker

Patient1 dataset was acquired on an Elekta (Elekta, Inc., Crawley, UK) Synergy XVI R4.0 system, and the remaining datasets on a Varian (Varian Medical Systems, Inc. Palo Alto, CA) OBI v1.3 system.

In the Patient1 dataset, the distance between the source and the origin was 1000 mm, the distance between the detector and the origin was 536 mm, and the panel size was  $400 \text{ mm} \times 400 \text{ mm}$  (calibrated to  $265.2 \text{ mm} \times 265.2 \text{ mm}$  at scan isocenter). As shown in Table II, the projection size was  $512 \times 512$  pixels, the number of projections was 701, and the calibrated projection pixel size (at scan isocenter) was 0.518 mm/pixel. The scan was acquired with 120 kVp and 1.6 mAs per projection. In Patient2 to Patient6 datasets, the distance between the detector and the origin was 500 mm, the projection size was  $1024 \times 768$  pixels, and the calibrated projection pixel size was 0.388 mm/pixel. The scan was acquired with 125 kVp and 4.0 mAs per projection. As shown in Table II, the numbers of projections in the Patient2 to Patient6 datasets were 2131, 2461, 2176, 2171, and 2198, respectively. Patient2 to Patient6 had fiducial markers implanted in or near the primary tumor that were segmented and extracted from each projection. The SI marker motion was converted into a respiratory signal.

##### 3) Ground Truth

Theoretically, no ground truth exists to evaluate the extracted breathing signal because different regions of the lungs such as upper lung, lower lung, and regions near the heart may not exhibit the same motion. We implemented diaphragm position-based and marker-based phase projection binning methods as surrogates for the ground truth for phase binning. Henceforth, we will use the term ‘‘Ground Truth Surrogate (GTS)’’ to refer to ‘‘phase binning ground truth’’ in this paper. For two phantom datasets and Patient1 dataset, the markers were not used, so their GTS used the diaphragm location-based method only. For Patient3, Patient5, and Patient6 datasets, the markers were used as well as their projections covered the diaphragm area. Accordingly, both diaphragm position-based and marker-based methods were GTS for Patient3, Patient5, and Patient6 datasets. GTS for Patient2 and Patient4 datasets used the marker-based method only because the markers were present but the diaphragm did not appear in the projections.

##### B. Intensity-based Feature Point Detection

We adjusted the image intensity values of the overall 4D CBCT projections as explained in Section II-A. To verify the effect of intensity adjustment in the image with outliers such as a high-density marker, we present the intensity adjustment

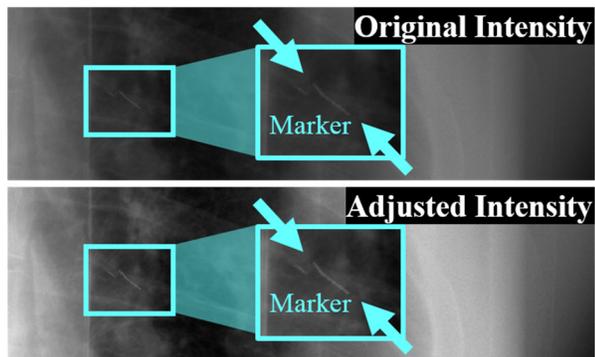


Fig. 4. Intensity adjustment result of the 340th projection of the Patient2 dataset.

results of the 340th projection of Patient2 dataset in Fig. 4. We randomly selected the Patient2 dataset from the patient datasets including markers in their projections. The area where the marker appeared was enlarged for better visibility in Fig. 4.

As shown in Fig. 4, the entire contrast of the projection got sharper so the high-density marker became brighter by the intensity adjustment. Average intensity values of the overall original projection and the overall intensity-adjusted projection were 0.30 and 0.44 with an intensity range of  $[0, 1]$ , respectively. Thus, the intensity adjustment increased the overall intensity of the 340th projection of the Patient2 dataset. Table III shows the average intensity before and after intensity adjustment. As shown in Table III, the average intensity values of the two phantom datasets that have the obvious contrast were decreased by the intensity adjustment. However, the average intensity values of six patient datasets with/without markers increased after the intensity adjustment.

TABLE III  
AVERAGE INTENSITY

Dataset	Original intensity		Adjusted intensity	
	AVG $\pm$ STD	Range	AVG $\pm$ STD	Range
Phantom1	0.61 $\pm$ 0.03	[0, 1]	0.55 $\pm$ 0.09	[0, 1]
Phantom2	0.61 $\pm$ 0.03	[0, 1]	0.55 $\pm$ 0.09	[0, 1]
Patient1	0.68 $\pm$ 0.09	[0, 0.96]	0.70 $\pm$ 0.09	[0, 1]
Patient2 (with markers)	0.32 $\pm$ 0.03	[0, 0.82]	0.44 $\pm$ 0.06	[0, 1]
Patient3 (with markers)	0.45 $\pm$ 0.04	[0.08, 0.94]	0.51 $\pm$ 0.03	[0, 1]
Patient4 (with markers)	0.43 $\pm$ 0.02	[0.15, 0.78]	0.51 $\pm$ 0.05	[0, 1]
Patient5 (with markers)	0.34 $\pm$ 0.03	[0.12, 1]	0.42 $\pm$ 0.02	[0, 1]
Patient6 (with markers)	0.41 $\pm$ 0.02	[0.14, 0.92]	0.55 $\pm$ 0.05	[0, 1]

The results of the proposed IFPD from Patient4 dataset are shown in Fig. 5. We empirically set the intensity difference threshold  $\tau$  as '0.07'. To demonstrate the roles of the mask radius  $r$  and the prior intensity adjustment, we provide the results according to the different conditions: We applied a different  $r$  of '1 to 5' for each column of Fig. 5, and we present results with intensity adjustment in Fig. 5(a)-(e) and results without intensity adjustment in Fig. 5(f)-(j). Here, we excluded the border of the entire image including a striking contrast from the search space because points selected over the image border degraded the performance as well as raised computation in tracking.

As shown in Fig. 5, IFPD selected points covering the edge areas where relatively large intensity variance appeared. A

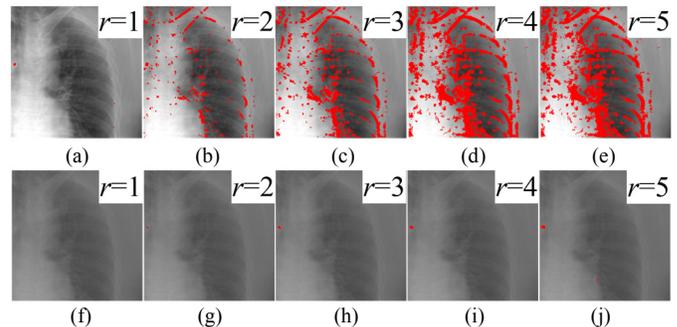


Fig. 5. Results of intensity-based feature point detection according to the mask radius  $r$ : the 500th projection of the Patient4 dataset (a)-(e) with and (f)-(j) without image intensity adjustment. Image intensity values of the projections in (a)-(e) were adjusted to saturate 1% of each projection as its low and high intensities.

smaller  $r$  selected less feature points as illustrated in Fig. 5, and every pixel was chosen as the feature point when  $r$  equaled 78. We obtained insufficient feature points without intensity adjustment as illustrated in Fig. 5(f)-(j). This is because the original 4D CBCT projections of Patient4 dataset did not have a stark contrast. Therefore, the experimental results showed that the extracted feature points depended on the radius  $r$  and the prior process for intensity adjustment. We conducted the rest of the experiments with the mask radius  $r$  of '1 to 5.' When  $r$  was '5,' the experiment could not continue due to insufficient computer memory, and the performance of IRPB was inferior when  $r$  was less than '4.' Therefore, we applied intensity adjustments and chose the mask radius  $r$  of '4' considering the computer memory limitations.

For the performance evaluation of the proposed IFPD, we compared it with Smallest Univalued Segment Assimilating Nucleus algorithm (SUSAN) [38] and Scale Invariant Feature Transform algorithm (SIFT) [39]. SUSAN is based on image intensity and yields the points satisfying the condition of the lowest univalued segment assimilating nucleus value [38]. SIFT is one of well-known algorithms in computer vision that detects local image features in applications of image stitching, video tracking, and various computer vision applications [39]. Table IV represents the number of feature point matches in IFPD, SUSAN, and SIFT. We compared the number of matches instead of feature points, because the chosen feature point with no correspondence in the adjacent projection is useless for trajectory tracking.

Table IV shows that the proposed IFPD had more matched feature points per projection than SUSAN and SIFT in the two phantom and the six patient datasets. Whereas SUSAN and SIFT had the maximum number of matches per projection of 1.79 and 14.89, respectively, the minimum number of the matches per projection in IFPD was 131.20. The experimental results showed that SUSAN and SIFT chose insufficient feature points with their matches per projection. This is because SUSAN and SIFT were designed for non-medical images that contain objects with definite borders. Without modification of SUSAN and SIFT, there is a limitation to use them for the relatively unclear CBCT projections. In the proposed method, we narrow down the initially selected feature points to the feature points with their matches in the adjacent projections then finally to the breathing motion-related points only. Thus,

TABLE IV  
 COMPARISON OF NUMBER OF FEATURE POINT MATCHES

Dataset	IFPD		SUSAN [38]		SIFT [39]	
	AVG $\pm$ STD	Range	AVG $\pm$ STD	Range	AVG $\pm$ STD	Range
Phantom1	279.21 $\pm$ 17 2.22	[41, 959]	0.02 $\pm$ 0.24	[0, 4]	10.00 $\pm$ 3.0 4	[4, 22]
Phantom2	266.46 $\pm$ 17 4.28	[45, 959]	0.01 $\pm$ 0.67	[0, 6]	9.85 $\pm$ 2.79	[4, 19]
Patient1	1956.27 $\pm$ 1659.18	[58, 8957]	1.28 $\pm$ 4.84	[0, 48]	7.15 $\pm$ 7.84	[0, 67]
Patient2	1731.08 $\pm$ 1534.79	[16, 9486]	0.31 $\pm$ 1.11	[0, 15]	8.16 $\pm$ 3.67	[0, 23]
Patient3	131.20 $\pm$ 134.38	[0, 1334]	0.25 $\pm$ 0.93	[0, 7]	3.24 $\pm$ 2.62	[0, 18]
Patient4	678.40 $\pm$ 780.36	[0, 4248]	0.07 $\pm$ 0.46	[0, 4]	0.81 $\pm$ 1.21	[0, 9]
Patient5	992.10 $\pm$ 57 5.41	[34, 4316]	1.79 $\pm$ 1.79	[0, 5]	14.89 $\pm$ 3.4 4	[5, 29]
Patient6	1616.43 $\pm$ 2 478.95	[39, 12663]	1.19 $\pm$ 2.29	[0, 29]	10.53 $\pm$ 5.0 3	[0, 32]

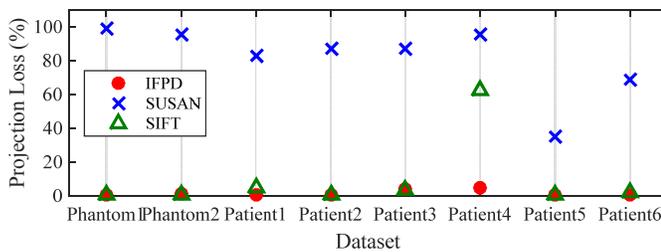


Fig. 6. Projection loss in terms of the trajectories for the two phantom and six patient datasets: the ratio of projections with no feature point matched to its correspondence in the next projection in the two phantom and six patient datasets.

the small number of feature point matches is likely to lead to the difficulty of tracking their fully connected trajectory over all of the projections.

A comparison of projection loss in terms of the trajectories between IFPD, SUSAN, and SIFT is shown in Fig. 6. Fig. 6 shows the ratio of projections with one feature point matched to its correspondence in the next projection. The average projection loss ratio of IFPD, SUSAN, and SIFT were 1.11%, 81.47%, and 8.92%, respectively. In other words, we cannot track an average of 81.47% of the trajectories with SUSAN.

If we can select one point that has its correspondences at every CBCT projection, and at the same time, its trajectory reveals the desired breathing pattern, there would be no relation between the number of the matches and the accuracy result of the respiration extraction. However, we could not acquire a single trajectory connected over the entire CBCT projections with our experimental datasets. Moreover, some of the matches have orbital patterns, not breathing patterns. The next step of IRPB, i.e., TTR, gets rid of the orbital motion-related trajectories formed by some of the matches we obtained. Consequentially, some of the matches of IFPD, SUSAN, and SIFT might be excluded and cause more projection losses. Therefore, we cannot expect to extract precise respiratory signal embracing all projections using SUSAN and SIFT.

### C. Trajectory Tracking using Random Sample Consensus

The first step for the proposed TTR is to find matches of corresponding points in consecutive projections. Not every selected feature point has correspondence in the prior and

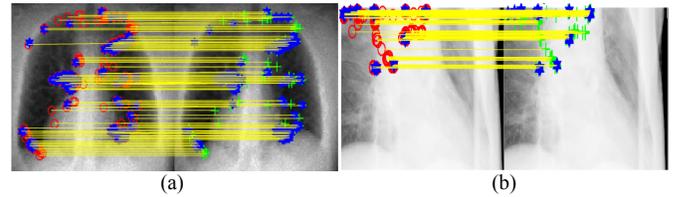


Fig. 7. Corresponding feature point matches: (a) between the 371st and the 372nd projections of the phantom1 dataset and (b) between the 212th and 213th projection of the Patient1 dataset.

posterior projections, because each projection was taken from the different rotation angle of the scanner and the image intensity is affected by the patient's movements. Fig. 7 presents matching results of the proposed TTR.

Fig. 7(a) illustrates corresponding feature point matches between the 371st and the 372nd projections of Phantom1 dataset, and Fig. 7(b) presents corresponding feature point matches between the 212th and the 213th projections of the Patient1 dataset. In Fig. 7, red and green markers are detected feature points by IFPD; yellow lines are feature point matches including inliers and outliers; and blue markers connected with yellow lines are inliers of the feature point matches in the two consecutive projections. As shown in Fig. 7, we discarded the outliers of the feature point matches from the trajectory candidates. Fig. 8(a) and Fig. 8(b) present three projection-long feature point trajectories  $T$  of Phantom1 and Patient3 datasets, respectively. As mentioned above, we considered the SI direction only for trajectory variation. Among detected trajectories in Fig. 8, some trajectories showed correlation with respiratory motion, but the others represented orbital motion.

In addition, Fig. 9 illustrates the number of the feature point trajectories  $T^f$  compared to their corresponding feature point matches  $C_i^f$ . In Fig. 9, a blue line is the number of the

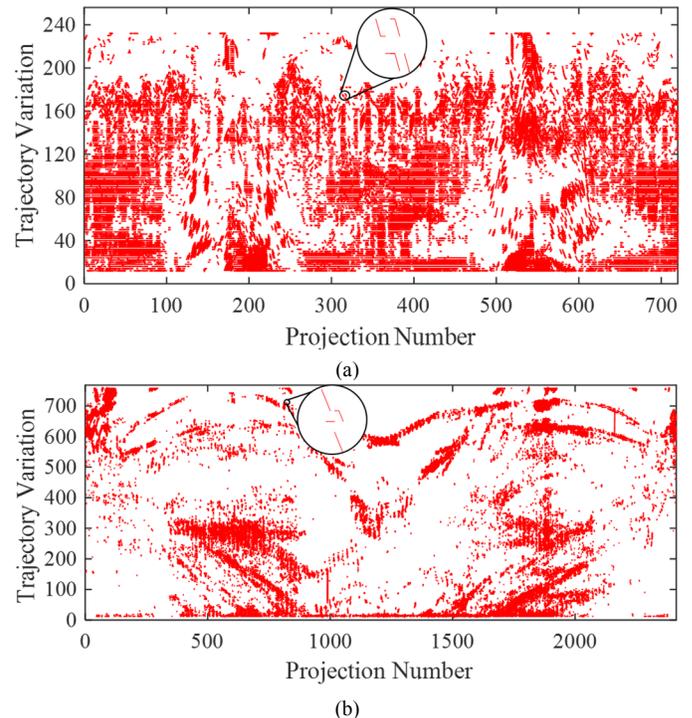


Fig. 8. Three projection-long feature point trajectory  $T$  of: (a) the phantom1 and (b) Patient3 datasets:  $T$  contains breathing motion-related  $T_R$  and orbital motion related trajectories, where the y-axis represents the SI direction.

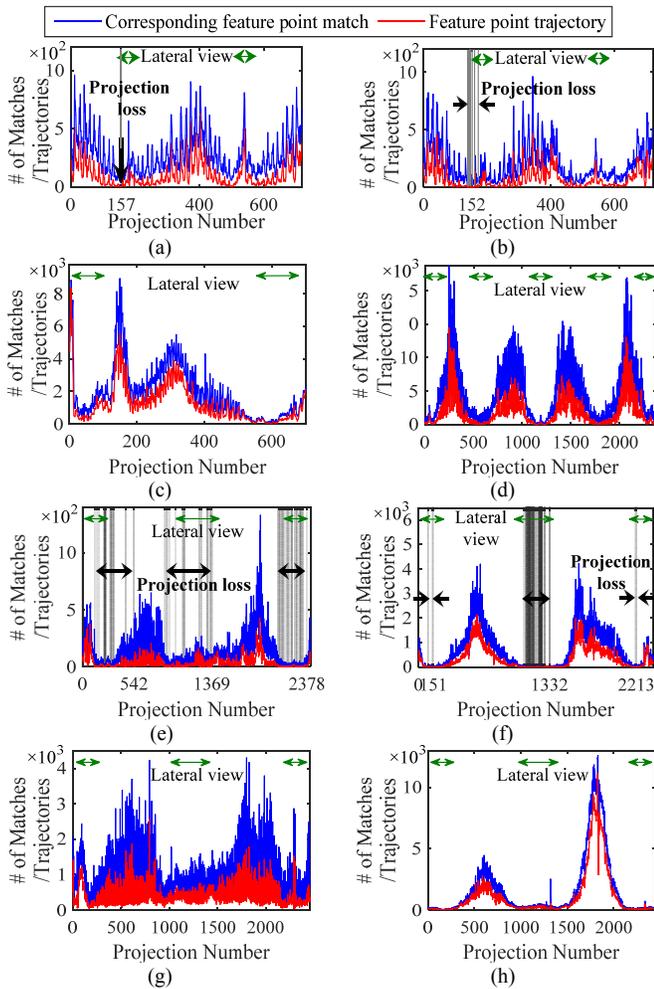


Fig. 9. Comparison of the number of corresponding feature point matches  $C_i^f$  and feature point trajectory  $T^f$  from: (a) phantom1, (b) phantom2, (c) Patient1, (d) Patient2, (e) Patient3, (f) Patient4, (g) Patient5, and (h) Patient6 datasets.

corresponding feature point matches  $C_i^f$ ; a red line is the number of the feature point trajectories  $T^f$ ; and gray vertical lines in Fig. 9(a), (b), (e), and (f) indicate where projection loss occurred. As shown in Fig. 9, the feature point trajectories of Phantom1 and Phantom2 datasets had ‘1’ (0.14%) and ‘8’ (1.11%) projection losses, respectively. The projection loss for the trajectories of Patient3 and Patient4 datasets were respectively ‘90’ (3.73%) and ‘106’ (4.45%). There was no projection loss in the Patient1, Patient2, Patient5, and Patient6 datasets. Nearby feature point trajectories can compensate for the intermittent losses unless projection losses are successive. Therefore, the projection losses in Phantom1, Phantom2, Patient3, and Patient4 datasets are negligible. Furthermore, the projections near the lateral views, e.g., x-ray gantry angle close to  $90^\circ$  and  $270^\circ$ , include less corresponding feature point matches and trajectories than other projections do, as shown in Fig. 9. In the lateral view, the overlapped lungs decrease the contrast on the image, and this results in less feature points because they are selected by IFPD based on their image intensity difference.

#### D. Respiratory Signal Extraction and Projection Binning

From the detected feature point trajectories, it is difficult to discover the patterns relevant to breathing variation. Therefore,

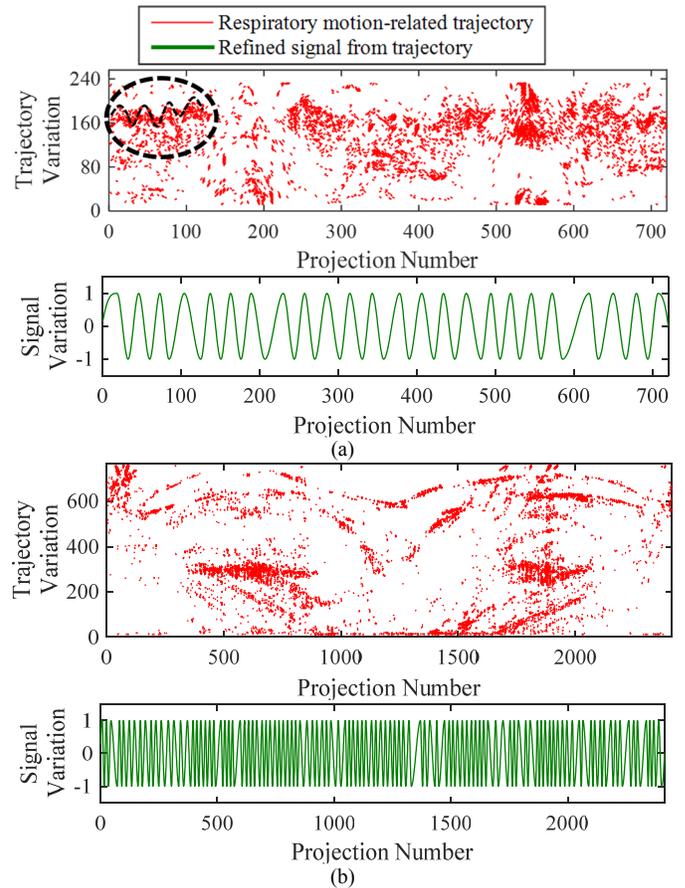


Fig. 10. Respiratory motion-related trajectories and respiratory signal of (a) the phantom1 and (b) Patient3 datasets.

we identified trajectories involved in respiration and the extracted respiratory pattern as described in Section II. Fig. 10(a) and (b) depict respiratory motion-related trajectories and the respiratory signal of Phantom1 and Patient3 datasets, respectively. In Fig. 10, red lines are the respiratory motion-related trajectories  $T_R$ , and the green line is the respiratory signal  $s_{resp}$  acquired by (3) and (5). Here, we already excluded orbital motion-related trajectories by applying PCA [33]. As explained above, the criterion for 4D CBCT binning is the projection number corresponding to the start and the end of the breathing cycle, not an amplitude of the respiratory signal. Unlike the feature point trajectories  $T$ , the respiratory motion-related trajectories  $T_R$  in Fig. 10 (a) exposed a periodic pattern marked with a black dashed circle.

Fig. 11 shows a comparison of the assigned breathing phases  $P(f)$  by IRPB, AS, and FT. Particularly, we utilized FT using the phase information of the CBCT projection intensities. In Fig. 11, a blue dotted line is the result of the diaphragm position-based method, a green dash-dot line is the result of the marker-based method, a red line is the result of IRPB, a purple line is the result of AS, and a magenta line is the result of FT. In addition, we implemented AS and FT using RTK [37] and MATLAB. For better visibility of phase assignment, we partially presented the breathing phases with a 250-projection window for two phantom datasets in Fig. 11(a) and (b) and a 150-projection window for six patient datasets in Fig. 11(c)-(h).

According to the results, IRPB, AS, and FT allocated

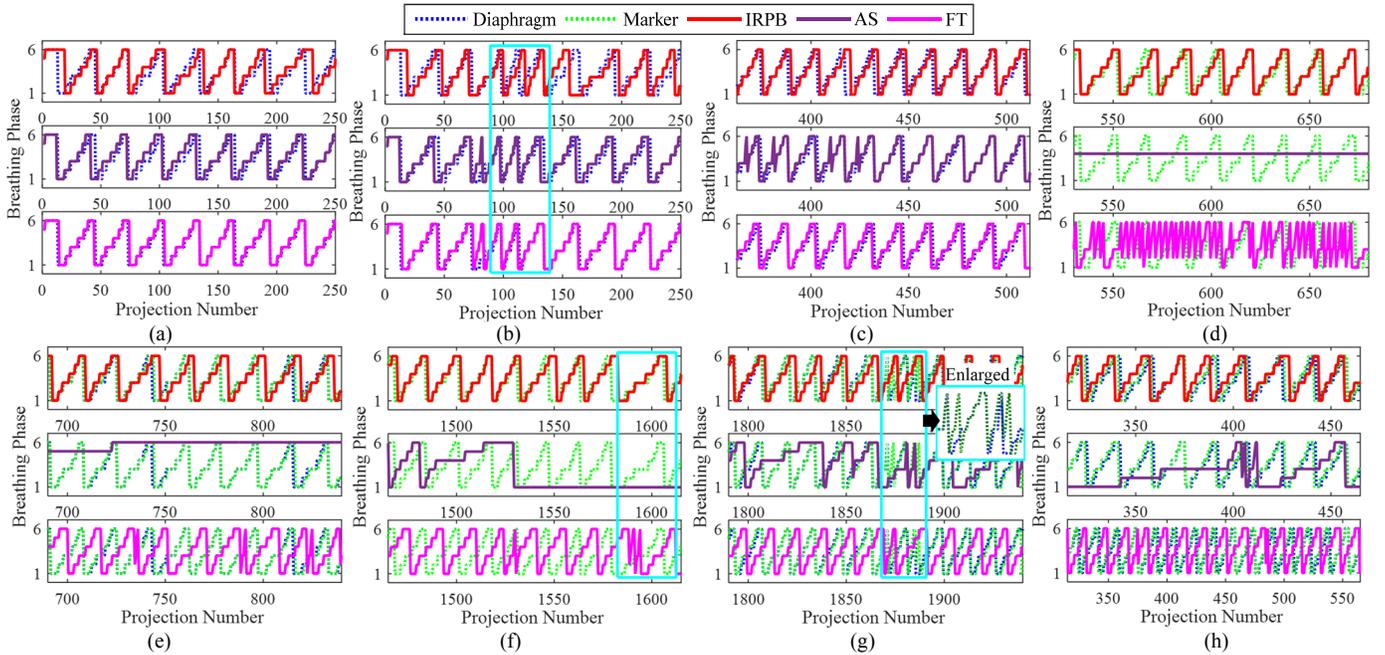


Fig. 11. Breathing phase comparison between IRPB, AS, and FT: (a) phantom1, (b) phantom2, (c) Patient1, (d) Patient2, (e) Patient3, (f) Patient4, (g) Patient5, and (h) Patient6 datasets.

breathing phases close to GTS—the diaphragm position-based and marker-based methods—as shown in Fig. 11(a)-(c). However, IRPB far outperformed AS and FT in breathing phase assignment for Patient2 to Patient6 datasets as shown in Fig. 11(c)-(h). The cyan rectangles in Fig. 11(b), (f) and (g) indicate noticeably irregular breathing. IRPB could set nearly precise breathing phases for both abnormally short (Fig. 11(b)) and long (Fig. 11(f)) breathing, but AS and FT could assign them for irregular breathing in Fig. 11(b) only. For the area marked as the cyan rectangle in Fig. 11(g), IRPB, AS, and FT could not set breathing phases accurately.

Interestingly, the length of irregular breath was estimated differently between the diaphragm position-based method and the markers-based method. For instance, in the enlarged cyan rectangle of Fig. 11(g), whereas the length of observed irregular breath was three cycles for the diaphragm position-based method, it was five cycles for the marker-based method. We believe this difference was because the fiducial markers were implanted near the heart in Patient5 and thus affected by cardiac motion.

Table V compares the average absolute phase shift produced by IRPB, AS, FT, and LIFT. The average absolute phase shift was calculated as:

$$\bar{P}(f, f_{GT}) = \sum_{i=1}^n |f^i - f_{GT}^i| / n_c, \quad P(f^i) = P(f_{GTS}^i) = 1,$$

where  $f^i$  and  $f_{GTS}^i$  are the projection numbers assigned as Phase1 in the  $i$ th breathing cycle by IRPB, AS, FT, and LIFT, and the diaphragm position-based/marker-based method, respectively, and  $n_c$  is the total number of breathing cycles.

IRPB, the diaphragm position-based method, and the marker-based methods extracted different numbers of respiratory cycles due to noticeably irregular breathing and the influence of other organs' movement as shown in Fig. 11(g). Therefore, we excluded those cycles in computation of the

TABLE V  
COMPARISON OF AVERAGE ABSOLUTE PHASE SHIFT

Method	Dataset	GTS	Phase shift (projections)
IRPB	Phantom1	Diaphragm	$3.54 \pm 2.78$
	Phantom2	Diaphragm	$3.56 \pm 2.60$
	Patient1	Diaphragm	$1.20 \pm 0.79$
	Patient2	Marker	$1.22 \pm 1.30$
	Patient3	Diaphragm	$2.67 \pm 1.41$
	Patient3	Marker	$2.44 \pm 1.24$
	Patient4	Marker	$1.56 \pm 1.24$
	Patient5	Diaphragm	$1.90 \pm 1.10$
	Patient5	Marker	$1.70 \pm 1.06$
	Patient6	Diaphragm	$2.33 \pm 1.00$
Patient6	Marker	$1.78 \pm 0.67$	
AS [20]	Phantom1	Diaphragm	$2.00 \pm 1.12$
	Phantom2	Diaphragm	$2.89 \pm 2.52$
	Patient1	Diaphragm	$0.30 \pm 0.48$
FT [21]	Phantom1	Diaphragm	$0.67 \pm 0.50$
	Phantom2	Diaphragm	$1.56 \pm 3.21$
	Patient1	Diaphragm	$0.70 \pm 0.48$
	Patient3	Diaphragm	$10.00 \pm 1.60$
	Patient3	Marker	$9.75 \pm 1.83$
	Patient4	Marker	$9.00 \pm 2.65$
	Patient5	Diaphragm	$5.90 \pm 0.32$
	Patient5	Marker	$5.70 \pm 0.67$
	Patient6	Diaphragm	$8.67 \pm 1.32$
Patient6	Marker	$8.11 \pm 1.17$	
LIFT [6]	Patient1	Marker	$1.68 \pm 1.09$ [6]

average absolute phase shift. AS and FT failed to assign proper breathing phases for most of the breathing cycles of the Patient2 to Patient6 datasets and Patient2 dataset, respectively, as shown in Fig. 11, so we computed the average absolute phase shift of two phantom and Patient1 datasets for AS and two phantom, Patient1, and Patient3 to Patient6 datasets for FT. In addition, we could compare IRPB and LIFT for the Patient1 dataset since it was also used in the experiment of [6].

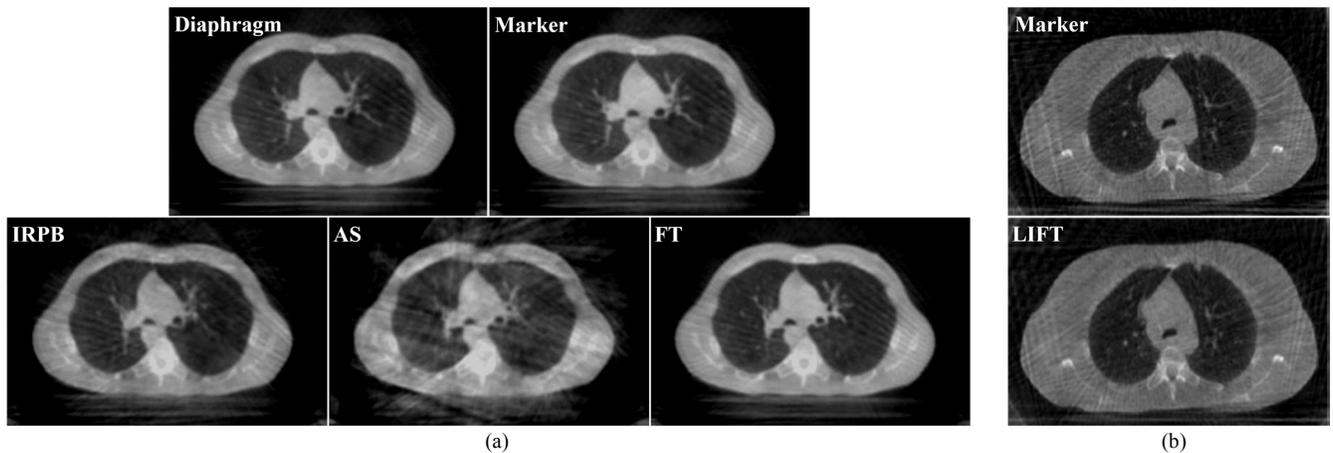


Fig. 12. Comparison of reconstructed 4D CBCT images: (a) the Patient5 dataset for Phase6 by the diaphragm position-based method, the marker-based method, IRPB, AS, and FT, and (b) the Patient2 dataset for Phase3 by the marker-based method and LIFT.

As shown in Table V, the average absolute phase shift of IRPB for the diaphragm position-based method was 2.41 projections in all datasets (2.53 projections except Patient2 dataset) and that for the marker-based method was 1.74 projections in all datasets (1.87 projections except Patient2 dataset). AS had the average absolute phase shift of 1.73 projections for the diaphragm position-based method in two phantom and Patient1 datasets. The average absolute phase shift of FT was 4.58 projections for the diaphragm position-based method and 8.14 projections for the marker-based method, in all datasets except Patient2 dataset. For Patient1, IRPB, AS, and FT showed slightly better performance with the average absolute phase shift of 1.20, 0.30, and 0.70 projections for the diaphragm position-based method, respectively, whereas LIFT had the average absolute phase shift of 1.68 projections for the diaphragm position-based method.

Fig. 12 shows the comparison of reconstructed 4D CBCT images: Patient5 dataset for Phase6 by the diaphragm position-based method, the marker-based method, IRPB, AS, and FT, and Patient2 dataset [6] for Phase3 by the marker-based method and LIFT. We used the Feldkamp, Davis, and Kress algorithm [40] for the reconstruction. The reconstructed image size was  $512 \times 512 \times 100$  pixels, and the spacing was 0.88 mm for left-right and anterior-posterior directions and 2 mm for SI direction. In addition, we set the intensity range of the reconstructed images to Hounsfield units. In Fig. 12, the bright streaking artifact, i.e., view-aliasing artifact, appears due to the insufficient number of projections used for 4D CBCT reconstruction compared to the total projections.

As illustrated in Fig. 12, the reconstructed 4D CBCT image by AS contained blurry boundaries near moving organs—the heart, vessels near the heart, and diaphragm—and significant streaking artifacts. Contrariwise, the reconstructed 4D CBCT image via IRPB, FT, and LIFT shows comparable image quality to that via the marker-based/diaphragm position-based method, manually implemented as GTS, as shown in Fig. 12. In addition, the reconstructed 4D CBCT image by FT seems more clear than the other images by IRPB, AS and LIFT, but the heart and vessel shapes on the reconstructed 4D CBCT image by FT are somewhat different from the images by GTS. As the result of the assigned breathing phases (Fig. 11(g)), FT nearly sorted

TABLE VI  
 PSNR AND SSIM OF RECONSTRUCTED 4D CBCT IMAGES BY IRPB, AS, FT, AND LIFT

Reconstructed 4D CBCT images		PSNR (dB)	SSIM
Fig. 12(a)	IRPB and Diaphragm	31.13	0.87
	AS [20] and Diaphragm	24.59	0.74
	FT [21] and Diaphragm	28.38	0.84
	IRPB and Marker	28.57	0.86
	AS [20] and Marker	24.96	0.74
	FT [21] and Marker	29.23	0.84
Fig. 12(b)	LIFT [6] and Marker	26.95	0.70

the projections in even cycles, so the view-aliasing artifact does not appear on the reconstructed 4D CBCT image via FT as much as the other images in Fig. 12. However, FT showed constant phase shifts, which led to the inaccurate heart and vessel shapes on the reconstructed 4D CBCT image.

In Table VI, we numerically evaluated the quality of the reconstructed 4D CBCT images of Fig. 12, by using Peak Signal-to-Noise Ratio (PSNR) and Structure Similarity Index (SSIM), commonly used as image quality metrics [41]. PSNR is computed as  $PSNR=20 \log_{10}(MAX/MSE)$ , where  $MAX$  is the maximum intensity value, and  $MSE$  is the mean square error between the reconstructed 4D CBCT images by IRPB, AS, FT, LIFT, and GTS—either the diaphragm position-based or marker-based method. SSIM is calculated as [41]:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)},$$

where  $x$  and  $y$  indicate the reconstructed 4D CBCT image by IRPB, AS, FT, LIFT, and GTS, respectively, and  $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_{xy}$  are means, variances, and covariance of  $x$  and  $y$ , respectively. In  $SSIM$ ,  $C_1$  and  $C_2$  denote stabilization constants calculated as  $C_1=(K_1 MAX)^2$  and  $C_2=(K_2 MAX)^2$ , where  $K_1 \ll 1$  and  $K_2 \ll 1$ . Here, we set  $K_1$  as 0.1 and  $K_2$  as 0.3 according to Image Processing Toolbox of MATLAB.

As shown in Table VI, IRPB had larger PSNR and SSIM than AS and FT when GTS was the diaphragm position-based method. When GTS was the marker-based method, FT showed the greatest PSNR and IRPB had the largest SSIM. In other words, IRPB was superior to AS and comparable to FT in terms of the image quality. LIFT had the smallest SSIM results among IRPB, AS, FT, and LIFT. Based on Table VI, we could

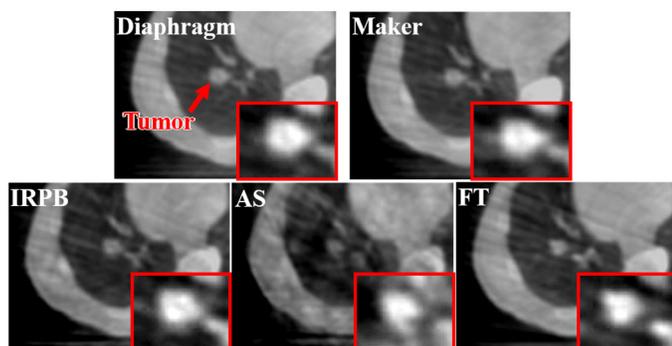


Fig. 13. Tumor on the reconstructed 4D CBCT image.

conclude that IRPB showed outstanding performance in the quality of the reconstructed 4D CBCT images, compared to AS, FT, and LIFT.

Fig. 13 compares reconstructed 4D CBCT images from the same image slice containing a tumor. The data was taken from the Patient5 dataset shown in Fig. 12(a). We magnified the area containing the tumor and increased the image contrast to compare the reconstructed tumor shapes, as shown in the insets within red boundaries in Fig. 13.

The tumor shape reconstructed using IRPB was similar with that generated by the marker-based/diaphragm position-based method. AS and FT had slightly distorted tumor shapes in the reconstructed 4D CBCT images resulting from inaccurate projection sorting. The tumor boundary was blurry for all five methods including GTS, as shown in Fig. 13. The blurry boundary is a disadvantage of the phase binning approach. Breathing motion-induced artifacts would be minimized in the reconstructed 4D CBCT images if the number of projections and breathing phases were sufficiently large. Thus, increasing the number of the phase bins may further decrease the motion-induced artifacts in the reconstructed 4D CBCT images.

In summary, the experimental results indicate that the proposed IFPD is robust for 4D CBCT projections that cannot capture the object clearly, unlike images used in many other simulations of computer vision. In our experiments, IFPD selected sufficient tissue feature points to track them over entire projections, but the existing intensity-based point selection approaches, SUSAN and SIFT, did not. Selection of the insufficient feature points was directly connected to the breathing phase shift. Moreover, IRPB resulted in outstanding image quality in comparison of the reconstructed 4D CBCT images by AS, FT, and LIFT.

We measured the computation time of the proposed method from the beginning of IFPD to the end of phase binning. The computation time for the proposed method was 4 hours and 24 minutes for Phantom1 dataset. We speculate on three possible solutions to reduce the computation time for IRPB down to minutes: First, we developed our programming code using Image Processing Toolbox of MATLAB for research purposes. Programming optimization on a faster platform than MATLAB such as OpenCV, C, or C++ is expected to shorten the computing time. Second, a machine with a GPU and parallel computing allows much faster IRPB processing than one with a CPU. The GPU memory specification also affects the computation time. Third, another solution is to apply IRPB to a

ROI of the projection, instead of the entire projection. The computation time would depend on the size of the ROI.

As mentioned above, the phase shift result of the marker-based method did not match the result of the diaphragm position-based method. This indicates that the diaphragm position-based method may not yield the precise tumor shape and clear tumor boundary on the reconstructed 4D CBCT images, depending on the tumor location. For example, the tumor near the diaphragm had similar motion with the diaphragm, and the tumor near the heart moved with both cardiac and respiratory motion. In the experiments, we extracted the breathing signal based on the dominant intensity features of the entire image of each 4D CBCT projection. Accordingly, IRPB has a potential limitation that its accuracy performance depends on the tumor location. Nevertheless, the need and risks of invasive implantation of the makers may be avoided by using IRPB. Setting the ROI to enclose the tumor in IRPB may result in a clearly delineated tumor in the reconstructed 4D CBCT image although other areas like the diaphragm might remain blurry.

#### IV. CONCLUSION

In this study, we introduced IRPB, the new marker-less 4D CBCT projection binning method, which combines IFPD's selection of tissue feature points using image intensity and TTR tracking the feature points by RANSAC algorithm. IRPB extracts the breathing motion and phases via periodicity analysis of feature point trajectories. In experiments, an average absolute phase shift of IRPB for GTS, i.e., the diaphragm location-based and the marker-based methods, was 2.05 projections and 6.27 projections less than that of FT, respectively. IRPB had an average absolute phase shift 0.48 projections less than LIFT under the same conditions. In comparison of average PSNR from the reconstructed 4D CBCT images, IRPB had 5.08 dB, 1.05 dB, and 2.90 dB greater PSNR average than AS, FT, and LIFT, respectively. An average SSIM of the 4D CBCT image reconstructed by IRPB was 0.13, 0.03, and 0.17 larger than that by AS, FT, and LIFT, respectively. These results regarding PSNR and SSIM demonstrate that IRPB achieved a higher quality of the reconstructed 4D CBCT images compared to the other methods we tested.

IRPB achieved automatic 4D CBCT projection binning without any manual/arbitrary process, additional equipment, restriction of the projection view, or supplemental signal/image processing depending on each patient. However, there are still remaining limitations such as computation time, memory requirements during processing, and accuracy performance. Many image registration methods introduced for computer vision have not been applied to respiratory signal extraction and 4D CBCT reconstruction methods yet. Therefore, our future work will involve modification, application, and evaluation of existing image registration algorithms for more precise respiratory signal extraction and 4D CBCT reconstruction.

#### REFERENCES

- [1] S. Kida, Y. Masutani, H. Yamashita, T. Imae, E. Matsuura, N. Saotome, K. Ohtomo, K. Nakagawa, A. Haga, "In-treatment 4D cone-beam CT with

- image-based respiratory phase recognition," *Radiol. Phys. Technol.*, vol. 5, no. 2, pp. 138-147, July 2012.
- [2] R. A. Sweeney, B. Seubert, S. Stark, V. Homann, G. Müller, M. Flentje, and M. Guckenberger, "Accuracy and inter-observer variability of 3D versus 4D cone-beam CT based image-guidance in SBRT for lung tumors," *Radiat. Oncol.*, vol. 7, pp. 81-88, June 2012.
- [3] C. P. V. Christoffersen, D. Hansen, P. Poulsen, and T. S. Sorensen, "Registration-based reconstruction of four-dimensional cone beam computed tomography," *IEEE Trans. Med. Imaging*, vol. 32, no. 11, pp. 2064-2077, Nov. 2013.
- [4] J. Sonke, L. Zijp, P. Remeijer, and M. van Herk, "Respiratory correlated cone beam CT," *Med. Phys.*, vol. 32, no. 4, pp. 1176-1186, Mar. 2005.
- [5] P. Keall, "4-dimensional computed tomography imaging and treatment planning," *Seminars in Radiat. Oncol.*, vol. 14, no. 1, pp. 81-90, Jan. 2004.
- [6] S. Dhou, Y. Motai, and G. D. Hugo, "Local intensity feature tracking and motion modeling for respiratory signal extraction in cone beam CT projections," *IEEE Trans. Biomed. Eng.*, Vol. 60, No. 2, pp. 332-342, Feb. 2013.
- [7] S. Dhou, G. D. Hugo, and A. Docef, "Motion-based Projection Generation for 4D-CT Reconstruction," *IEEE Int. Conf. Image Processing (ICIP)*, pp. 1698-1702, Oct. 2014.
- [8] 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, pp. 1513-1525, Oct. 2009.
- [9] M. Chen, K. Cao, Y. Zheng, and R. A. C. Siochi, "Motion-compensated mega-voltage cone beam CT using the deformation derived directly from 2D projection images," *IEEE Trans. Med. Imaging*, vol. 32, no. 8, pp. 1365-1375, Aug. 2013.
- [10] S. Leng, J. Zambelli, R. Tolakanahalli, B. Nett, P. Munro, J. Star-Lack, B. Paliwal, and G. Chen, "Streaking artifacts reduction in four-dimensional cone-beam computed tomography," *Med. Phys.*, vol. 35, no. 10, pp. 4649-4659, Oct. 2008.
- [11] C. Shieh, J. Kipritidis, R. T. O'Brien, Z. Kuncic, and P. J. Keall, "Image quality in thoracic 4D cone-beam CT: A sensitivity analysis of respiratory signal, binning method, reconstruction algorithm, and projection angular spacing," *Med. Phys.*, vol. 41, no. 4, Apr. 2014.
- [12] 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.
- [13] J. Hinkle, M. Szegedi, B. Wang, B. Salter, and S. Joshi, "4D CT image reconstruction with diffeomorphic motion model," *Med. Image Anal.*, vol. 16, no. 6, pp. 1307-1316, Aug. 2012.
- [14] M. Szegedi, J. Hinkle, P. Rassiah, V. Sarkar, B. Wang, S. Joshi, and B. Salter, "Four-dimensional tissue deformation reconstruction (4D TDR) validation using a real tissue phantom," *J. Applied Clinical Med. Phys.*, vol. 14, no. 1, pp. 115-132, Sept. 2013.
- [15] 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, no. 5, pp. 1587-1596, Dec. 2008.
- [16] 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. 3825-3833, Sept. 2006.
- [17] 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, pp. 1087-1098, Feb. 2008.
- [18] J. C. Park, S. H. Park, J. H. Kim, S. M. Yoon, S. S. Kim, J. S. Kim, Z. Liu, T. Watkins, and W. Y. Song, "Four-dimensional cone-beam computed tomography and digital tomosynthesis reconstructions using respiratory signals extracted from transcutaneously inserted metal markers for liver SBRT," *Med. Phys.*, vol. 38, no. 2, pp. 1028-1036, Feb. 2011.
- [19] S. Rit, M. vanHerck, L. Zijp, and J. J. Sonke, "Quantification of the variability of diaphragm motion and implications for treatment margin construction," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 82, no. 3, pp. 399-407, Mar. 2012.
- [20] L. Zijp, J. J. Sonke, and M. van Herk, "Extraction of the respiratory signal from sequential thorax cone-beam X-ray images," *Int. Conf. the Use of Computers in Radiation Therapy*, pp. 507-509, May 2004.
- [21] I. Vergalaso, J. Cai, and F. Yin, "A novel technique for markerless, self-sorted 4D-CBCT: Feasibility study," *Med. Phys.*, vol. 39, no. 3, pp. 1442-1451, Mar. 2012.
- [22] H. Yan, X. Wang, W. Yin, T. Pan, M. Ahmad, X. Mou, L. Cervino, X. Jia, and S. B. Jiang, "Extracting respiratory signals from thoracic cone beam CT projections," *Phys. Med. Biol.*, vol. 58, no. 5, pp. 1447-1464, Feb. 2013.
- [23] 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, June 2009.
- [24] M. A. Fischler and R. C. Bolles, "Random consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381-395, June 1981.
- [25] J. R. Van Aken, "An efficient ellipse-drawing algorithm," *IEEE Computer Graphics and Applications*, vol. 4, no. 9, pp. 24-35, Sept. 1984.
- [26] B. K. Horn and B. G. Schunk, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1-3, pp. 185-203, Aug. 1981.
- [27] D. Zhou, Y. Gao, L. Lu, H. Wang, Y. Li, and P. Wang, "Hybrid corner detection algorithm for brain magnetic resonance image registration," *2011 4th Int. Conf. Biomedical Engineering and Informatics (BMEI)*, vol. 1, pp. 308-313, Oct. 2011.
- [28] J. Ehrhardt, R. Werner, A. Schmidt-Richberg, and H. Handels, "Statistical modeling of 4D respiratory lung motion using diffeomorphic image registration," *IEEE Trans. Med. Imaging*, vol. 30, no. 2, pp. 251-265, Feb. 2011.
- [29] D. Fuiorea, V. Gui, D. Pescaru, and C. Toma, "Kernel based image registration versus MLESAC: a comparative study", *5th Int. Symp. Applied Computational Intelligence and Informatics*, pp. 255-260, May 2009.
- [30] P. H. S. Torr and D. W. Murray, "Outlier detection and motion segmentation," *Proc. SPIE 2059, Sensor Fusion VI*, pp. 432-443, Aug. 1993.
- [31] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 138-156, Apr. 2000.
- [32] S. Choi, T. Kim, and W. Yu, "Performance evaluation of RANSAC family," *Conf. Proc. of British Machine Vision*, pp. 1-12, 2009.
- [33] I. T. Jolliffe, "Properties of Sample Principal Components," in *Principal Component Analysis*, 2nd ed., New York, USA: Springer, 2002, pp. 29-61.
- [34] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627-1639, July 1964.
- [35] F. N. Fritsch and R. E. Carlson, "Monotone piecewise cubic interpolation," *SIAM J. Numerical Anal.*, vol. 17, no. 2, pp. 238-246, Apr. 1980.
- [36] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," *Med. Phys.*, vol. 37, no. 9, pp. 4902-4915, Sept. 2010.
- [37] S. Rit, M. Vila Oliva, S. Brousmiche, R. Labarbe, D. Sarrut, and G. C. Sharp, "The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)," *Journal of Physics: Conference Series*, vol. 489, no. 1, pp. 1-4, Mar. 2014.
- [38] S. M. Smith and J. M. Brady, "SUSAN — A new approach to low level image processing," *International Journal of Computer Vision*, vol. 23, no. 1, pp. 45-78, May 1997.
- [39] David G. Lowe, "Object recognition from local scale-invariant features," *the 7th IEEE Int. Conf. Proc. on Computer Vision*, vol. 2, pp. 1150-1157, Sept. 1999.
- [40] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Amer. A-Opt. Image Sci. Vis.*, vol. 1, no. 6, pp. 612-619, June 1984.
- [41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, Apr. 2004.