A probability-based multi-cycle sorting method for 4D-MRI: A simulation study

Xiao Liang, Fang-Fang Yin, Yilin Liu, and Jing Cai^{a)}

Medical Physics Graduate Program, Duke University, Durham, North Carolina 27705

and Department of Radiation Oncology, Duke University Medical Center, Durham, North Carolina 27710

(Received 15 June 2016; revised 14 October 2016; accepted for publication 17 October 2016; published 8 November 2016)

Purpose: To develop a novel probability-based sorting method capable of generating multiple breathing cycles of 4D-MRI images and to evaluate performance of this new method by comparing with conventional phase-based methods in terms of image quality and tumor motion measurement.

Methods: Based on previous findings that breathing motion probability density function (PDF) of a single breathing cycle is dramatically different from true stabilized PDF that resulted from many breathing cycles, it is expected that a probability-based sorting method capable of generating multiple breathing cycles of 4D images may capture breathing variation information missing from conventional single-cycle sorting methods. The overall idea is to identify a few main breathing cycles (and their corresponding weightings) that can best represent the main breathing patterns of the patient and then reconstruct a set of 4D images for each of the identified main breathing cycles. This method is implemented in three steps: (1) The breathing signal is decomposed into individual breathing cycles, characterized by amplitude, and period; (2) individual breathing cycles are grouped based on amplitude and period to determine the main breathing cycles. If a group contains more than 10% of all breathing cycles in a breathing signal, it is determined as a main breathing pattern group and is represented by the average of individual breathing cycles in the group; (3) for each main breathing cycle, a set of 4D images is reconstructed using a result-driven sorting method adapted from our previous study. The probability-based sorting method was first tested on 26 patients' breathing signals to evaluate its feasibility of improving target motion PDF. The new method was subsequently tested for a sequential image acquisition scheme on the 4D digital extended cardiac torso (XCAT) phantom. Performance of the probability-based and conventional sorting methods was evaluated in terms of target volume precision and accuracy as measured by the 4D images, and also the accuracy of average intensity projection (AIP) of 4D images.

Results: Probability-based sorting showed improved similarity of breathing motion PDF from 4D images to reference PDF compared to single cycle sorting, indicated by the significant increase in Dice similarity coefficient (DSC) (probability-based sorting, $DSC = 0.89 \pm 0.03$, and single cycle sorting, $DSC = 0.83 \pm 0.05$, p -value < 0.001). Based on the simulation study on XCAT, the probability-based method outperforms the conventional phase-based methods in qualitative evaluation on motion artifacts and quantitative evaluation on tumor volume precision and accuracy and accuracy of AIP of the 4D images.

Conclusions: In this paper the authors demonstrated the feasibility of a novel probability-based multicycle 4D image sorting method. The authors' preliminary results showed that the new method can improve the accuracy of tumor motion PDF and the AIP of 4D images, presenting potential advantages over the conventional phase-based sorting method for radiation therapy motion management.

© 2016 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4966705>]

Key words: 4D-MRI, probability-based sorting, XCAT

1. INTRODUCTION

Respiratory motion is known to cause artifacts in medical imaging, characterized by blurring or discontinuity of moving tissues in the images.^{1–8} For fast imaging techniques such as CT, moving tumor/tissues may be extended or compressed in the directions of breathing motion, causing potential error in target volume delineation and dose calculation in radiation therapy. For slow imaging techniques such as CBCT,⁹ SPECT (Ref. 10), and PET,¹¹ the resulting images reflect the averaged intensity of the anatomy throughout the breathing cycle.

4D imaging techniques have been developed to account for breathing motion for various imaging modalities.^{1,7,12–15} 4D imaging reduces or even eliminates the breathing-related motion artifacts by sorting image data into separate respiratory bins, providing a clear shot of anatomy at each phase of the breathing cycle. However, it has been shown that the quality of 4D imaging is affected by breathing irregularity.^{8,9,16} Clements *et al.*,⁹ studied the effect of irregular breathing patterns in 4D-CT on a Perspex thorax phantom imaged with the same acquisition settings used for stereotactic lung patients. It has been found that when irregular motion was

introduced 4D-CT volumes were reduced by up to 20% and internal target volume (ITV) discontinuity was present in 4D-CT maximum intensity projection (MIP) images. In a study conducted by Cai *et al.*,¹⁷ real time sagittal dynamic magnetic resonance imaging was resorted using a 4D-CT acquisition method based on phantom or patient images to simulate the imaging rebinning processes. The accuracy of average intensity projection (AIP) of 4D images was quantitatively evaluated by comparing AIPs generated from resorted dynamic MRI images and the full set of dynamic MRI images. It has been found that the AIP similarity shows a linear negative correlation with the respiratory variability normalized by tumor sizes. The impact of irregular motions on 4D-PET image reconstruction has also been investigated by Tian *et al.*¹⁶ by evaluating the shape of reconstructed point source activity distributions. It has been found that irregular motions with amplitude fluctuation and baseline drift partly influence the target shape, causing elongation of targets in the motion direction in 4D-PET images. To cope with breathing irregularity, breathing guidance techniques such as audio instructions and a video guiding interface have been developed to facilitate regular and predictable breathing motion. Although breathing guidance techniques have been reported to effectively improve breathing regularity, applicability of these techniques is restricted to patients capable of following guidance instructions.^{18–20}

Instead of eliminating breathing irregularity, previous studies suggest possibility of reducing deleterious effects of breathing irregularity by incorporating breathing variation into the 4D reconstruction process. Previous studies on breathing motion variation suggest that when breathing signals become longer, the probability density function (PDF) of breathing motion displacement will approach stability.^{21,22} Based on this finding, although breathing motion varies in both amplitude and period from cycle to cycle, each patient may have a relatively stable PDF. In addition, a previous study has shown that the breathing motion PDF of a single breathing cycle is dramatically different from the true stabilized PDF that resulted from many breathing cycles.¹⁷ However, conventional 4D sorting methods generate only one cycle, incapable of revealing breathing variation information. The lack of breathing variation information in 4D images may result in inaccurate target delineation and erroneous image intensity distribution, potentially causing inaccuracy in patient setup and dose calculation. For some imaging modalities such as CT and MRI using cine acquisition, the incompleteness of breathing variation information for each slice is inherent because image data at each slice location are acquired continuously in a short period of time, typically slightly over a breathing cycle, thus experiencing little to no breathing variation. For some other modalities such as MRI using sequential acquisition, it is possible to extract breathing variation information from raw data because image data at each slice location are acquired over relatively long intervals, typically several breathing cycles. Based on this observation, we developed a novel, probability-based sorting method capable of generating multiple breathing cycles of 4D-MRI images to overcome the aforementioned limitations

of single-cycle sorting methods. The feasibility of the proposed method was tested using patients' breathing signals and on digital human phantoms, and compared with conventional phase-based sorting methods in terms of image quality and tumor motion measurement.

2. METHODS AND MATERIALS

2.A. The probability-based multi-cycle 4D sorting method

Figure 1 shows the general workflow of the probability-based multi-cycle 4D sorting method. The overall idea is to identify a few main breathing cycles (and their corresponding weightings) that can best represent the main breathing patterns of the patient and then reconstruct a 4D image set for each of the identified main breathing cycles. This method requires image data to be randomly acquired in relation to breathing variations, i.e., different parts of the anatomy can experience the different breathing variations throughout the entire image acquisition process. As a result, this method cannot be applied to some 4D imaging techniques whose image acquisition scheme does not meet this criterion. For example, it is inapplicable to current cine/helical 4D-CT techniques, in which each slice is imaged for only a short period of time (slightly greater than one breathing cycle) and therefore does not experience the breathing variations throughout the image

FIG. 1. Workflow of the probability-based multi-cycle 4D sorting method.

acquisition. It is expected that this method not only results in multi-cycle 4D images that represent the main breathing patterns of the patient but also reduces breathing variation artifacts in each 4D image set. In addition, it also improves the accuracy of tumor motion PDF and AIP of the 4D images. The method is carried out in the following steps:

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

$$A_i = \frac{d_{\text{peak},i} + d_{\text{peak},i+1}}{2} - d_{\text{valley},i}, \quad (1)$$

$$T_i = t_{\text{peak},i+1} - t_{\text{peak},i}, \quad (2)$$

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

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

(3) For each main breathing cycle, a set of 4D images is reconstructed using a result-driven sorting method adapted from our previous study.²³ As a brief introduction, the result-driven sorting method has been developed for the purpose of effectively utilizing redundant images when multiple images are assigned to a bin. First, the average respiratory trajectory is calculated from the entire breathing signal as the targeted motion result. With the average respiratory trajectory, it is possible to identify the amplitude of a bin, defined as the amplitude of the average trajectory at the center of a bin. In each bin, the image with the amplitude closest to the amplitude of this bin is selected to generate the 4D-MRI dataset. In this study, an average respiratory cycle is calculated for each main breathing cycle. The full EOE to end-of-inhalation (EOI) range of the averaged motion trajectory is equally divided into

ten amplitude bins with five on the descending slope and five on the ascending slope. The total number of bins to be filled is the number of amplitude bins, 10, multiplied by the number of main breathing cycles. Each frame of acquired images is assigned to one or more of these bins based on the breathing amplitude when the frame was acquired. A frame is allowed to be assigned to bins in different main breathing cycles but will not be assigned to bins on the opposite slope at this stage. After assigning all the frames, there may be bins with either more than one frame or no frame. For empty bins, attempts to fill the bin will be made in the following order, copying from the bin in the opposite slope, interpolating from adjacent slices at the same phase, interpolating from the same slice at adjacent phases. For bins filled with multiple frames, the frame with the closest amplitude to the center of the bin is selected. The sorting is completed when every bin has been filled with one and only one frame. It should be noted that the probability-based multi-cycle 4D sorting method can be applied to any imaging modality, providing that the image data (images, or image raw data such as projections in CT and k -space data in MRI) are randomly sampled in relation to breathing variations. It can also be applied to any type of respiratory surrogate, such as real-time position management (RPM), diaphragm, tumor, internal markers, as long as it provides consistent breathing signal from the same source.

2.B. Breathing profile study

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

$$\text{PDF}_{\text{prob}}(x) = \frac{\sum_j p_j(x) T_{\text{main},j} f_{\text{main},j}}{\sum_j T_{\text{main},j} f_{\text{main},j}}, \quad (3)$$

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

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

where t is the time of a data point between the i th and the $i+1$ th peaks. The single-cycle motion trajectory was converted to a breathing motion PDF for the phase sorting method, $\text{PDF}_{\text{single}}(x)$. The original RPM breathing signal was also converted to a breathing motion PDF as reference, $\text{PDF}_{\text{ref}}(x)$.

To quantitatively evaluate the similarity of PDFs from 4D sorting methods with the reference PDF, Dice similarity coefficient (DSC) was calculated for each of the 26 RPM breathing signals between $\text{PDF}_{\text{single}}(x)$ and $\text{PDF}_{\text{ref}}(x)$, and $\text{PDF}_{\text{prob}}(x)$ and $\text{PDF}_{\text{ref}}(x)$, respectively, as follows:

$$\text{DSC}_{\text{single}} = \frac{2|S_{\text{single}} \cap S_{\text{ref}}|}{|S_{\text{single}}| + |S_{\text{ref}}|}, \quad (5)$$

$$\text{DSC}_{\text{prob}} = \frac{2|S_{\text{prob}} \cap S_{\text{ref}}|}{|S_{\text{prob}}| + |S_{\text{ref}}|}, \quad (6)$$

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

2.C. Digital phantom study

Feasibility of the probability-based multi-cycle 4D image sorting method was tested for a sequential image acquisition scheme using the 4D digital extended cardiac torso (XCAT) phantom.²⁴ The sequential image acquisition scheme, which was previously demonstrated for T_2 -weighted 4D-MRI sorting,²³ meets the requirement of probability-based multi-cycle 4D sorting since the image data are sampled randomly in relation to breathing variation.

To evaluate the probability-based multi-cycle 4D sorting method, the conventional phase-based single-cycle 4D sorting was also simulated using the same corresponding XCAT phantom for comparison, as outlined in Fig. 2. Furthermore, the phase-based sorting method was simulated for two different image acquisition schemes: cine acquisition and sequential acquisition, respectively. The former is the current clinical standard method used in 4D-CT, and the latter is the same as used in the probability-based method. Both image acquisition methods have been described in detail in the literature,^{13,23,25,26} and will only be briefed here. In cine

acquisition, image data are continuously acquired at a slice location for at least a breathing cycle before data at the next location are acquired. The data needed for 4D sorting are ready when all slice locations are imaged. In sequential acquisition, a slice location is imaged only once before the next slice location is imaged. The acquisition process will repeat multiple times for the entire volume until a certain percentage of all required data is obtained. 3D volumes were also acquired at different displacements along the breathing curves to serve as reference.

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

Depending on the source of breathing signal for XCAT, the simulation was performed in two types of scenarios: (1) the 4D XCAT phantom was set to move with an artificially generated breathing signal consisting of two breathing patterns that are repeated alternatively; motion amplitudes and periods are 14.5 mm and 2.77 s and 30.0 mm and 2.25 s, respectively. Simulation using the artificial breathing signal was done to demonstrate and validate the proposed methodology in the simplest scenario and (2) the 4D XCAT phantom was set to move with the patient's RPM breathing signal ($n = 6$). For all cases, the peak-to-peak motion amplitude for the diaphragm was set to 30 mm. The XCAT images were generated in the transverse plane using the following parameters: in-plane resolution: 1 mm; matrix size: 512×512 ; slice thickness: 5 mm; number of slices: 50. The simulated volume covers the upper abdominal region from the bottom of the liver at

Fig. 2. The study design for the comparison of the probability-based sorting method with the conventional phase sorting methods using a digital phantom.

the EOI phase, to the highest point of the diaphragm at the EOE phase. To simulate the clinical images for patients with liver cancer, a spherical target of 3 cm in diameter was added in the liver and set to move with the liver during respiration. The XCAT images were generated with similar appearance to T_2 -weighted MRI images by using the activity mode of the XCAT phantom, in which intensities of the organs and tissues were assigned using values extracted from FRFSE T_2 -weighted MRI images.

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

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

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

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

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

the probability-based method, the AIP was calculated as a weighted average of the phases of the multiple cycles. Mathematically, it was calculated as

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

where $I_{j,k}$ is the 3D volume at the k th amplitude bin of the j th main breathing pattern, $\mu_{j,k}$ is the ratio of the duration of the k th amplitude bin of the j th main breathing pattern to the period of the j th main breathing pattern, $T_{main,j}$. The reference AIP, AIP_{ref} , used as the ground-truth for evaluation, was determined as the average of all image data obtained throughout the entire image acquisition period.

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

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

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

where n_V is the number of voxels in the volume, n_P is the number of voxels on the profile, $AIP(i,j,k)$ is the image

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

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

3. RESULTS

3.A. Breathing profile study

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

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

3.B. Digital phantom study

Figure 4 shows the sorting results using the phase-based and the probability-based methods for the two-pattern artificial breathing signal. The artificial breathing signal [Fig. 4(a)], generated as described in Sec. 2.C, represents a simplified scenario of breathing variation. The two main breathing patterns comprising the artificial signal are shown in Fig. 4(b). In 4D images reconstructed using the conventional phase-based methods [Fig. 4(c), row 1, and 2], major motion artifacts can be observed in and around the target and the diaphragm regions, especially in phase 4, 5, and 6, where tissues are duplicate and/or missing. In contrast, in the two cycles of 4D images reconstructed with the probability-based sorting method [Fig. 4(c), row 3, and 4], visible motion artifacts are minimal.

Figure 5 shows the sorting results using the phase-based and the probability-based methods for an example patient's

TABLE I. DSC for 26 patient breathing signals, quantifying similarity between $PDF_{\text{single}}/PDF_{\text{prob}}$ and PDF_{ref} , and the difference in DSC for PDF_{single} and PDF_{prob} .

Case#	DSC		ΔDSC
	Single	Prob.	
1	0.82	0.84	0.02
2	0.83	0.88	0.05
3	0.90	0.91	0.01
4	0.83	0.90	0.07
5	0.88	0.92	0.04
6	0.93	0.90	-0.03
7	0.80	0.88	0.08
8	0.80	0.87	0.07
9	0.83	0.89	0.07
10	0.81	0.89	0.08
11	0.78	0.88	0.10
12	0.87	0.92	0.06
13	0.77	0.87	0.10
14	0.77	0.91	0.14
15	0.87	0.83	-0.04
16	0.82	0.86	0.04
17	0.88	0.96	0.08
18	0.90	0.93	0.04
19	0.86	0.92	0.06
20	0.81	0.93	0.12
21	0.82	0.90	0.08
22	0.78	0.91	0.12
23	0.79	0.90	0.11
24	0.80	0.85	0.05
25	0.85	0.89	0.04
26	0.76	0.86	0.10
Mean $\pm$ S.D.	0.83 ± 0.05	0.89 ± 0.03	$P < 0.001$

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

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

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

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

Figure 6 demonstrated AIPs of the 4D images at a region including the liver and the diaphragm and corresponding results on the accuracy of AIP. Results for the artificial breathing profile are shown in the left [Figs. 6(a)–6(c)] and results for an example patient breathing signal in the right [Figs. 6(d) and 6(e)]. In the AIPs from the conventional phase-sorting methods, motion artifacts are visible in the diaphragm region, represented by discontinuity and deformation of image intensity distribution. Also, the smeared targets in AIPs from the conventional methods exhibit noticeable difference from the targets in reference AIPs in terms of intensity distribution. In contrast, the AIPs from the probability-based method show minimal motion artifacts and similar intensity distribution in the target region as compared to the reference AIPs [Figs. 6(a) and 6(d)]. Intensity difference maps of AIPs are shown in Figs. 6(b)–6(e) for better visualizing the AIP accuracy for the three methods compared against the reference AIP. Since the liver in the XCAT phantom is assigned with a uniform intensity, major differences of AIPs are found at the diaphragm and the target regions for all three methods. However, the intensity differences are smaller for the probability-based method. Intensity differences along the blue line were plotted, showing that the intensity difference profile for the probability-based method has smaller maximum difference than profiles for the conventional methods for both cases [Figs. 6(c) and 6(f)]. AIP accuracy was also quantitatively measured by $dAIP_V$ and $dAIP_P$, which reflect AIP accuracy in the entire volume and along the profile, respectively, as summarized in Table III. Based on significant

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

4. DISCUSSION

In this study, a novel probability-based multi-cycle 4D sorting method was described and evaluated for its feasibility and potential clinical efficacy. The probability-based method was first applied on 26 patients' RPM breathing curves to extract the main breathing patterns. The PDF from the probability-based method has been shown to be more accurate than the PDF from conventional single-cycle methods in terms of similarity to the reference PDF from the original breathing signals. This result also indicates that the new method can be potentially used for generating more accurate AIPs. This potential was confirmed in the digital phantom simulation

FIG. 5. (a) An example patient breathing signal acquired during CT simulation. (b) The three main breathing patterns identified with the probability-driven sorting method. (c) Ten-phase 4D images generated using the phase sorting method for cine acquisition. (d) Ten-phase 4D images generated using the phase sorting method for sequential acquisition. (e) Ten-phase 4D images, cycle 1 (top), cycle 2 (middle), and cycle 3 (bottom), generated using the probability-driven sorting method.

study, in which the XCAT 4D images were generated using an artificially generated breathing curve and six patients' RPM breathing curves. The ability of the probability-based method to generate 4D datasets with better motion PDF and thus more accurate AIPs can be theoretically deduced. Suppose a patient breathing signal $x(t)$ continuously recorded over $t = 0 \rightarrow T$ can be decomposed into monotonic segments, $x_k(t)$. For each segment, one and only one inverse function, $t_k(x)$, exists. Therefore, $\text{PDF}_{\text{ref}}(x)dx$, i.e., the probability of motion displacement staying between x and $x + dx$, can be expressed as

$$\text{PDF}_{\text{ref}}(x)dx = \frac{\sum_k |dt_k|}{T} = \frac{\sum_k |t'_k(x)|}{T} dx, \quad (12)$$

$$\text{PDF}_{\text{prob}}(x) = \frac{\sum_j p_j(x) T_{\text{main},j} f_{\text{main},j}}{\sum_j T_{\text{main},j} f_{\text{main},j}}. \quad (13)$$

The equation for calculating PDF_{prob} is introduced in Sec. 2.B and is shown above for reference. In the extreme case when each individual breathing cycle is grouped as a main breathing cycle, $f_{\text{main},j} = 1, \forall j$, as $p_j(x)$ becomes the

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

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

FIG. 6. The top row shows the AIPs of the liver region generated using (from left to right) the phase sorting for cine acquisition, the phase sorting for sequential acquisition, the probability-driven sorting, and the reference, based on (a) the artificial breathing signal and (d) a patient breathing signal. The middle row shows the intensity difference maps of AIPs for (from left to right) the phase sorting for cine acquisition, the phase sorting for sequential acquisition, and the probability-driven sorting based on (b) the artificial breathing signal and (e) a patient breathing signal. The bottom row shows the profile of intensity difference along the blue lines for the phase sorting for cine acquisition (dashed-dotted), the phase sorting for sequential acquisition (dashed), and the probability-driven sorting (solid).

PDF for cycle j , $p_j(x)T_{\text{main},j}dx$ represents the total time of displacement staying between x and $x+dx$ during cycle j . Thus we have

$$\sum_j p_j(x)T_{\text{main},j}f_{\text{main},j}dx = \sum_j \sum_i |t_{j,i}'(x)|dx = \sum_k |t_k'(x)|dx, \quad (14)$$

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

$$\text{PDF}_{\text{prob}}(x) \rightarrow \text{PDF}_{\text{ref}}(x).$$

As for conventional single-cycle sorting methods, $\text{PDF}_{\text{single}}(x)$ may improve as the number of phases increases but is not guaranteed to converge to $\text{PDF}_{\text{ref}}(x)$. The above demonstration is intended to facilitate understanding of the results in Fig. 3 rather than to serve as a rigid proof of the properties of $\text{PDF}_{\text{prob}}(x)$ as the number of main breathing cycle increases.

Currently, in free-breathing treatment, on-board CBCT, effectively an AIP over multiple breathing cycles, is registered with the AIP of planning single-cycle 4D images for patient alignment. However, as introduced in Sec. 1, it has been shown that PDFs of single breathing cycles are considerably different from PDFs of multiple breathing cycles,¹⁷ which may affect appearance of AIPs of targets and, in turn, patient setup

TABLE III. Quantitative results of AIP difference from the reference AIP for the two conventional sorting methods and the probability-based sorting method.

	dAIP _V			dAIP _P		
	Cine	Seq.	Prob.	Cine	Seq.	Prob.
Artificial	1.46	1.27	0.65	3.89	3.93	1.88
pt #1	1.52	1.30	0.92	6.85	4.36	2.16
pt #2	1.36	1.42	0.81	3.69	4.05	2.01
pt #3	2.23	1.62	1.36	6.62	4.51	3.81
pt #4	1.66	1.25	0.95	5.03	3.42	2.98
pt #5	1.29	1.33	0.93	3.20	3.14	2.53
pt #6	1.60	1.34	0.89	5.87	5.23	2.65
mean ± SD	1.59 ± 0.31	1.36 ± 0.13	0.93 ± 0.22	5.02 ± 1.47	4.09 ± 0.70	2.57 ± 0.67
p-value	0.002	0.011	/	0.002	0.002	/

TABLE IV. CoM shifts of the target from the reference AIP to the 4D AIPs.

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

accuracy. For the purpose of more accurate target alignment, it is reasonable to register on-board CBCT with AIPs from the probability-based sorting instead of AIPs from conventional single-cycle sorting methods in free-breathing treatment, as this study shows that the probability-based sorting is able to provide more accurate PDFs than single-cycle sorting methods. Results from evaluation of accuracy of the target CoMs in the AIPs suggest that the probability-based sorting may provide AIPs with more accurate and stable target CoMs than the sequential acquisition phase-based sorting method. As there is large variation of the target CoM shifts between studied cases for the cine acquisition phase-based sorting, more inclusive studies are warranted to confirm the advantage of the probability-based sorting over the cine acquisition phase-based sorting. More accurate PDFs derived from the probability-based sorting method may also benefit PDF-based treatment planning and dose calculation and extend the applicability of PDF-based treatment planning for treating moving targets using the free-breathing technique.^{17,21,22} Currently, dose calculation for free-breathing treatment is performed on AIP images. For a moving target, the dose to AIP is calculated, which can be different from the actual dose to target. Sheng *et al.* calculated actual doses to target by convoluting the dose distribution with the target PDF and also proposed PDF-based treatment planning based on PDF-based dose calculation for the purpose of reducing dose to tissues surrounding the target. It was found in their study that if a highly reproducible PDF is used in optimization, the gross tumor volume (GTV) dose coverage can be maintained and the mean lung dose reduced by 9.6%. However, they showed that the dosimetric error increased nearly exponentially as the PDF error increased.¹⁷ Based on their study, as a rough estimate, a 6% increase (from 83% to 89% as demonstrated in this study) in tumor motion PDF reproducibility could lead to a 33% decrease in dosimetric error in GTV coverage.

In this study, the proposed probability-based sorting method was only tested in a 4D-MRI sequential acquisition scenario. It is expected, however, that the new method can be potentially applied for other 4D imaging techniques. A direct extension of our current work is to apply the method for 4D-MRI reconstruction based on retrospective *k*-space sorting. In *k*-space sorting, the *k*-space can be divided into

many small segments, each of which is associated with a time point and assigned to a respiratory phase. Regardless of the MRI data acquisition scheme (radial, spiral, etc.), the assignment of respiratory phases to the *k*-space segments is a random process, which meets the criteria for probability-based sorting. In addition, we expect the probability-based sorting method can also be applied to other imaging modalities. One possible application is for 4D-PET reconstruction because of the randomness of decay events relative to breathing variation. PET events detected within a short period of time can be used to reconstruct a volume with low signal-to-noise Ratio (SNR), which is associated with a time point on the breathing signal. These volumes will be subsequently resorted into bins for multiple main breathing cycles. Due to the low SNR of each volume, each bin must have a sufficient number of volumes for generating multi-cycle 4D images with acceptable image quality. Despite the potentially broad applicability, it is notable that this method may not work for 4D-CT due to the fact that 4D-CT is limited to cine acquisition mode. As a result, raw data for each slice only cover a small portion of breathing curves, which often shows similar breathing pattern without much variation. Extending CT acquisition time may not provide sufficient coverage for breathing variation and increased exposure to ionizing radiation is not desirable.

It has been shown in this study that the probability-based sorting method, capable of generating 4D images with multiple main breathing patterns, can improve 4D image quality. The improvement can be further enhanced by optimizing the algorithm. For example, in the grouping process, the groups were delimited by amplitude and period of individual breathing cycles as shown in Fig. 3(b). This approach, however, may separate individual cycles with close amplitudes and/or periods into two groups. This result is because of grouping individual cycles based on preset grids. One possible remedy to this problem is replacing the current grouping method with data-based grouping, or clustering method such as *k*-means clustering method with optional optimization of the number of groups. The number of groups may be reduced while keeping the PDF similarity as high as current results. Another modification to the current method that can potentially improve the outcome is taking the location of individual breathing cycles into consideration. Baseline shift has been observed in some patient's breathing curves. In the current method, if two individual breathing cycles share the same amplitude and period but have different locations, they will end up in the same group. For a patient with moderate-to-none baseline shift, the reconstructed images will not be affected. In some patients, however, considerable baseline shift has been observed. In this situation, the final motion range may be underestimated if positions of individual breathing cycles are not taken into consideration. To extend our method to patients with large baseline shifts, each individual breathing cycle will be characterized with three attributes, amplitude, period, and position. The probability-based sorting method identifies the main breathing patterns based on the analysis of breathing curves captured by external surrogates. In our digital phantom study, the breathing curves were directly used

to guide the motion of the phantom so the internal-external breathing motion is perfectly correlated. For patient imaging, although previous studies have pointed out the internal-external breathing motion correlation is usually high, it is still desired to test our method on real patient imaging to evaluate how the method performs with the presence of uncertainty in internal-external breathing motion correlation.

5. CONCLUSION

In this paper we demonstrated the feasibility of a novel probability-based multi-cycle 4D image sorting method. Our preliminary results showed that the new method can improve the image quality of 4D images, the accuracy of tumor motion PDF, and the AIP of 4D images, presenting potential advantages over the conventional phase-based sorting method for radiation therapy motion management.

ACKNOWLEDGMENT

This work is partly supported by funding from NIH (Grant No. 1R21CA165384).

^{a)} Author to whom correspondence should be addressed. Electronic mail: jing.cai@duke.edu; Telephone: 919-684-1089; Fax: 919-660-2180.

¹ A. F. Abdelnour, S. A. Nehmeh, T. Pan, J. L. Humm, P. Vernon, H. Schoder, K. E. Rosenzweig, G. S. Mageras, E. Yorke, S. M. Larson, and Y. E. Erdi, "Phase and amplitude binning for 4D-CT imaging," *Phys. Med. Biol.* **52**, 3515–3529 (2007).

² J. Ehrhardt, R. Werner, D. Saring, T. Frenzel, W. Lu, D. Low, and H. Handels, "An optical flow based method for improved reconstruction of 4D CT data sets acquired during free breathing," *Med. Phys.* **34**, 711–721 (2007).

³ M. J. Fitzpatrick, G. Starkschall, J. A. Antolak, J. Fu, H. Shukla, P. J. Keall, P. Klahr, and R. Mohan, "Displacement-based binning of time-dependent computed tomography image data sets," *Med. Phys.* **33**, 235–246 (2006).

⁴ W. Lu, P. J. Parikh, J. P. Hubenschmidt, J. D. Bradley, and D. A. Low, "A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4D CT," *Med. Phys.* **33**, 2964–2974 (2006).

⁵ Y. D. Mutaf, J. A. Antolak, and D. H. Brinkmann, "The impact of temporal inaccuracies on 4DCT image quality," *Med. Phys.* **34**, 1615–1622 (2007).

⁶ T. Pan, X. Sun, and D. Luo, "Improvement of the cine-CT based 4D-CT imaging," *Med. Phys.* **34**, 4499–4503 (2007).

⁷ E. Rietzel and G. T. Chen, "Improving retrospective sorting of 4D computed tomography data," *Med. Phys.* **33**, 377–379 (2006).

⁸ T. Yamamoto, U. Langner, B. W. Loo, Jr., J. Shen, and P. J. Keall, "Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients," *Int. J. Radiat. Oncol., Biol., Phys.* **72**, 1250–1258 (2008).

⁹ N. Clements, T. Kron, R. Franich, L. Dunn, P. Roxby, Y. Aarons, B. Chesson, S. Siva, D. Duplan, and D. Ball, "The effect of irregular breathing patterns on

internal target volumes in four-dimensional CT and cone-beam CT images in the context of stereotactic lung radiotherapy," *Med. Phys.* **40**, 021904 (10pp.) (2013).

¹⁰ I. Polycarpou, I. Chrysanthou-Baustert, O. Demetriadou, Y. Parpottas, C. Panagidis, P. K. Marsden, and L. Livieratos, "Impact of respiratory motion correction on SPECT myocardial perfusion imaging using a mechanically moving phantom assembly with variable cardiac defects," *J. Nucl. Cardiol.* **1**–10 (2015).

¹¹ W. Sureshbabu and O. Mawlawi, "PET/CT imaging artifacts," *J. Nucl. Med. Technol.* **33**, 156–161 (2005), quiz 163–154.

¹² M. von Siebenthal, G. Szekely, U. Gamper, P. Boesiger, A. Lomax, and P. Cattin, "4D MR imaging of respiratory organ motion and its variability," *Phys. Med. Biol.* **52**, 1547–1564 (2007).

¹³ J. Cai, Z. Chang, Z. Wang, W. Paul Segars, and F. F. Yin, "Four-dimensional magnetic resonance imaging (4D-MRI) using image-based respiratory surrogate: A feasibility study," *Med. Phys.* **38**, 6384–6394 (2011).

¹⁴ T. Li, B. Thorndyke, E. Schreiber, Y. Yang, and L. Xing, "Model-based image reconstruction for four-dimensional PET," *Med. Phys.* **33**, 1288–1298 (2006).

¹⁵ J. J. Sonke, L. Zijp, P. Remeijer, and M. van Herk, "Respiratory correlated cone beam CT," *Med. Phys.* **32**, 1176–1186 (2005).

¹⁶ Y. Tian, K. Stutzer, W. Enghardt, M. Priegnitz, S. Helmbrecht, C. Bert, and F. Fiedler, "Experimental investigation of irregular motion impact on 4D PET-based particle therapy monitoring," *Phys. Med. Biol.* **61**, N20–N34 (2016).

¹⁷ K. Sheng, J. Cai, J. Brookeman, J. Molloy, J. Christopher, and P. Read, "A computer simulated phantom study of tomotherapy dose optimization based on probability density functions (PDF) and potential errors caused by low reproducibility of PDF," *Med. Phys.* **33**, 3321–3326 (2006).

¹⁸ S. Pollock, R. Keall, and P. Keall, "Breathing guidance in radiation oncology and radiology: A systematic review of patient and healthy volunteer studies," *Med. Phys.* **42**, 5490–5509 (2015).

¹⁹ T. Neicu, R. Berbeco, J. Wolfgang, and S. B. Jiang, "Synchronized moving aperture radiation therapy (SMART): Improvement of breathing pattern reproducibility using respiratory coaching," *Phys. Med. Biol.* **51**, 617–636 (2006).

²⁰ P. H. Cossmann, "Video-coaching as biofeedback tool to improve gated treatments: Possibilities and limitations," *Z. Med. Phys.* **22**, 224–230 (2012).

²¹ J. Cai, P. W. Read, T. A. Altes, J. A. Molloy, J. R. Brookeman, and K. Sheng, "Evaluation of the reproducibility of lung motion probability distribution function (PDF) using dynamic MRI," *Phys. Med. Biol.* **52**, 365–373 (2007).

²² J. Cai, P. W. Read, J. M. Larnier, D. R. Jones, S. H. Benedict, and K. Sheng, "Reproducibility of interfraction lung motion probability distribution function using dynamic MRI: Statistical analysis," *Int. J. Radiat. Oncol., Biol., Phys.* **72**, 1228–1235 (2008).

²³ Y. Liu, F. F. Yin, B. G. Czito, M. R. Bashir, and J. Cai, "T2-weighted four dimensional magnetic resonance imaging with result-driven phase sorting," *Med. Phys.* **42**, 4460–4471 (2015).

²⁴ W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. Tsui, "4D XCAT phantom for multimodality imaging research," *Med. Phys.* **37**, 4902–4915 (2010).

²⁵ 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.* **31**, 333–340 (2004).

²⁶ 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.* **48**, 45–62 (2003).