Data-driven respiratory motion compensation for four-dimensional cone-beam computed tomography (4D-CBCT) using groupwise deformable registration

Matthew J. Riblett

Department of Radiation Oncology, Virginia Commonwealth University, Richmond, VA 23298, USA

Gary E. Christensen

Department of Electrical and Computer Engineering and Department of Radiation Oncology, University of Iowa, Iowa City, IA 52242, USA

Elisabeth Weiss

Department of Radiation Oncology, Virginia Commonwealth University, Richmond, VA 23298, USA

Geoffrey D. Hugo^{a)}

Department of Radiation Oncology, Washington University, St. Louis, MO 63110, USA

(Received 30 November 2017; revised 30 April 2018; accepted for publication 6 June 2018; published 18 September 2018)

Purpose: To demonstrate the feasibility of using a purely data-driven, *a posteriori* respiratory motion modeling and reconstruction compensation method to improve 4D-CBCT image quality under clinically relevant image acquisition conditions.

Methods: Evaluated workflows that utilized a combination of groupwise deformable image registration and motion-compensated image reconstruction algorithms. Groupwise registration is an approach that simultaneously registers all temporal frames of a 4D image to a common reference instead of one at a time so as to minimize the influence of any individual time point on the global smoothness or accuracy of the resulting deformation model. Four-dimensional cone-beam CT (4D-CBCT) Feldkamp–Davis–Kress (FDK) reconstructions were registered to either iteratively computed mean respiratory phase (mean-frame) or preselected respiratory phase (fixed-frame) reference images to model respiratory motion. The resulting 4D transformations were used to deform projection data during the FDK backprojection operation to create motion-compensated reconstructions. Tissue interface sharpness (TIS) was defined as the slope of a sigmoid curve fit to a mobile tissue boundary and was used to evaluate image quality in regions susceptible to motion artifacts. Image quality improvement was assessed for 19 clinical cases by evaluating mitigation of view aliasing artifacts, TIS, image noise reduction, and contrast for implanted fiducial markers.

Results: Average (standard deviation) diaphragm TIS recovery relative to initial 4D-CBCT reconstructions was observed to be 87% (46%) using fixed-frame registration alone; 87% (47%) using fixed frame with motion-compensated reconstruction; 101% (68%) using mean-frame registration alone; and 99% (65%) using mean frame with motion-compensated reconstruction. Noise was reduced in sampled soft tissue ROIs by 58% for both fixed-frame registration and registration with motion compensation and by 57% and 58% on average for the corresponding mean-frame methods, respectively. Average improvement in local CNR was observed to be respectively 93% and 98% for fixed-frame registration and registration with motion compensation methods and 116% and 111% for the corresponding mean-frame methods.

Conclusion: Data-driven groupwise registration and motion-compensated reconstruction offer a feasible means of improving the quality of 4D-CBCT images acquired under clinical conditions. The addition of motion compensation reconstruction after groupwise registration visibly reduced the impact of view aliasing artifacts for the clinical image datasets studied. © 2018 American Association of Physicists in Medicine [https://doi.org/10.1002/mp.13133]

Key words: cone-beam computed tomography, groupwise image registration, image reconstruction, lung cancer, motion compensation

1. INTRODUCTION

To ensure correct patient setup and periodically review therapeutic progress in radiation therapy, it is common to acquire images of target anatomy *in situ* using Cone-Beam CT

(CBCT) imaging.^{1,2} Current CBCT imaging modalities acquire a complete single-revolution dataset over a substantially greater duration (60–240 s) than the period of the typical respiratory cycle (4–8 s) of a patient. Due to respiratory motion, the position of anatomical features can vary

FIG. 1. Examples of respiratory motion related artifacts in CBCT reconstruction of the lung. The left column shows cross-sections of a FDK 3D-CBCT reconstruction. Notice that the lung–diaphragm boundary (arrows) appears blurred. Likewise, only large blood vessels are visible and they also appear blurred. The middle column shows the same data reconstructed using respiratory phase-correlated FDK 4D-CBCT. Notice that there is relatively little blurring due to breathing (arrows) using this approach. However, considerable streaking (view aliasing) artifacts are visible due to reduced angular projection sampling. The right column shows the same data reconstructed using motion-compensated 4D-CBCT. Notice that this method has less motion blurring (arrows) and streaking artifacts compared to the former two methods. [Color figure can be viewed at wileyonlinelibrary.com]

dramatically between adjacent projections. When reconstructing 3D volumes, anatomical regions and structures influenced by respiratory motion can appear blurred or distorted in the resulting 3D volume (see left-most panel, Fig. 1). The resulting degraded image quality can result in significant target localization discrepancies^{3,4} due to loss of contrast and lack of clear anatomical boundaries.

One potential solution to this problem is to use respiratory correlated CBCT (4D-CBCT).⁵ In this approach, images are retrospectively reconstructed from phase-binned subsets of projections to effectively arrest the motion in a series of rendered phases of the respiratory cycle. Such techniques have been adopted clinically because it has been shown that 4D-CBCT imaging can recover respiratory motion similar to that recovered by 4D-CT.⁶ However, in 4D-CBCT, reduction of the available projections per phase by respiratory phase-binning results in significant angular undersampling of the imaged volume, resulting in streak or view aliasing artifacts that contribute to an increase in image noise (see center panel, Fig. 1).

To address the limitations of 4D-CBCT, a range of improved acquisition and reconstruction methods have been investigated. One such approach, gated CBCT,⁷ avoids view aliasing challenges upfront by monitoring patient respiration and prompting the acquisition of uniformly spaced projection of data around the patient only at predefined points in the respiratory cycle. This prospective approach effectively arrests respiratory motion in reconstructions and eliminates reconstruction errors caused by limited projection data, but carries the implicit penalties of increased acquisition time and limited insight into patient motion.

Alternatively, motion compensation (MoCo) reconstruction methods, which estimate patient motion observed during CBCT acquisition and subsequently adapt all available projection data to account for any observed deformations between respiratory phases^{8–14} and motion-correction

methods, which apply similar motion-based adjustments to reconstructed CBCT volumes, have been developed. MoCo can be implemented either before or after completion of CBCT reconstruction. Motion correction^{10,15} and earlier MoCo reconstruction strategies^{8,9,16} utilize patient-specific deformation models estimated *a priori* from 4D-CT to adjust projection or reconstructed image data to account for motion likely experienced during CBCT acquisition. However, the clinical utility of such methods may be limited by the accuracy of the *a priori* model, which may not remain valid over the course of treatment.¹⁷ Conversely, Brehm et al.,^{11–13} Zhang et al.,¹⁶ and Wang and Gu¹⁴ each proposed strategies that generalize the MoCo approach without utilizing an *a priori* motion model. The MoCo technique proposed by Brehm^{11–13} uses a simulated, artifact-free prior image, generated by segmenting an initial reconstructed volume, to directly determine the impact of reconstruction artifacts on the image. Zhang¹⁶ demonstrated the ability to generate a realistic motion model directly from six-phase thoracic CBCT reconstructions when regularized by principle components analysis (PCA). Similarly, Wang and Gu proposed the simultaneous motion estimation and image reconstruction (SMEIR) method¹⁴ that seeks to iteratively update a motion model to minimize the difference between synthetic projection data forward projected through an iterative reconstruction and raw CBCT projections.

In the Brehm et al., Zhang et al., and Wang and Gu approaches, estimation of the motion model is performed in a phase-by-phase manner, which does not fully exploit the temporally redundant nature of data in 4D-CBCT image: regions of anatomy that are impacted by motion or reconstruction artifacts in one temporal frame may be unaffected in one or more adjacent temporal frames. In this paper, we present a new MoCo approach that uses groupwise registration to build the underlying motion model. Our approach incorporates this temporal redundancy of the data into the registration scheme

to mitigate the possible manifestation of phase-specific registration bias, which may otherwise occur in conventional phase-to-phase registrations, and ensure that reasonable motion modeling is achieved consistently across all respiratory phases. Details of the groupwise approach are discussed later in the Methods section. Groupwise registration has been previously demonstrated to produce reasonable models of subject motion when applied to 4D-CT, dynamic magnetic resonance (MR), and ultrasound (US) images.^{18,19}

2. MATERIALS AND METHODS

2.A. Free-breathing CBCT projection dataset

CBCT projection datasets used in this study were from a larger dataset of 4D-CBCT scans from 20 subjects undergoing concurrent radiochemotherapy for locally advanced nonsmall cell lung cancer.^{20,21} This dataset, collected under an institutional review board approved protocol, is described in more detail by Hugo *et al.*²² To demonstrate proof-of-concept in this study, one projection set was selected for each subject from those acquired during the first 2 weeks of treatment for which a reasonable respiratory signal could be extracted. Each of the 4D-CBCT datasets was acquired on a commercial CBCT scanner (On-Board Imager™, Varian Medical Systems, Inc., Palo Alto, CA). Approximately 2150–2750 projections per subject were acquired over a period of 8–10 min in half-fan mode with half bow-tie filter using 125 kVp, 20 mA, and 20 ms in a single 360° slow gantry arc. Audio-visual biofeedback was used during acquisition in all subjects. A purely data-driven method of signal extraction, discussed later, which extracts signal from the raw projection data, was utilized for initial projection phase-binning prior to 4D reconstruction.

2.B. Full vs clinically representative dataset

The acquisition technique used in this study is not representative of standard clinical techniques. Notably, the number of projections and scan duration far exceed the more typical 600–800 projections and 1–2-min scan duration of a routine clinical CBCT scan on the same hardware. While advantageous for the reconstruction of 4D volumes, the correspondingly fine angular resolution of the projection data greatly reduced the degree of image noise and view aliasing artifact otherwise expected in a clinical image. To evaluate the efficacy of the proposed methods under clinical conditions, a clinically representative, reduced projection dataset was produced for each subject by systematically sampling a portion of the corresponding full projection dataset.

Figure 2 illustrates two techniques that were investigated to create a downsampled set of projections from the full dataset. The first approach, called Projection Striding (PS), samples the full projection dataset with a uniform stride and then performs respiratory phase-binning of the sampled projections prior to reconstruction. Unfortunately, this approach can develop reconstruction failures, as shown in Fig. 2. Failure

occurs when the stride sampling synchronizes with a given phase of the respiratory cycle resulting in an uneven angular distribution of projections across phase bins. The second approach, called Phase-Binned Striding (PBS), performs nonuniform stride sampling of the full projection dataset such that each phase bin has approximately the same number of projections and these projections form an approximate uniform angular sampling of the circle. PBS mitigates the likelihood of developing the respiratory cycle synchronization-related reconstruction failures observed during application of the PS method. This is because the subsampled dataset has approximately even angular sampling and a relative distribution of projections in each phase bin that is consistent with that of the full dataset. For this study, the PBS method was used with a step size, or downsampling factor, of 4 to reduce the number of projections for each subject to approximately 540–690.

2.C. Respiratory signal acquisition

A respiratory signal is required to facilitate the initial phase-binning of the projection datasets and, later, the application of phase-specific motion compensation. The Amsterdam Shroud (AS) method²³ was used to generate shroud images and respiratory signals for the full projection dataset of each patient using the algorithm implementation available in the *Reconstruction Toolkit* (RTK).²⁴ Routine difficulties were experienced in signal extraction from the raw shroud image within projection spans that included (a) anatomical regions dominated by large, temporally invariant features of sharp intensity gradients (i.e., vertebrae) and (b) when mobile tissue structures moved out of the projection field of view (FOV) due to the half-fan CBCT acquisition mode utilized. Following Zijp *et al.*,²³ we applied a temporal derivative filter to the shroud image to highlight periodic motion exhibited by the diaphragm and other high-intensity structures, thereby correcting these issues. Out of the 20 subjects in the original database, a respiratory signal could not be reliably extracted from the projection set of one subject (102) using the shroud-based method, so this subject was excluded. To generate signals for the clinically representative datasets, the full projection set respiratory signal was sampled along with the projection data, so the respiratory signal remained consistent with the projections.

2.D. Projection sorting and initial FDK reconstruction

After acquiring a respiratory signal from the shroud image, projections were sorted into ten phase bins according to their respective position in the respiratory cycle. Initial 4D patient image reconstruction was accomplished for both the full and clinically representative datasets using the RTK implementation of the Feldkamp–Davis–Kress (FDK) reconstruction algorithm.²⁵ The phase-binned datasets were reconstructed individually and stacked temporally to create the initial 4D-CBCT image used for image registration and motion modeling.

FIG. 2. Illustration of workflow for generating clinically representative projection CBCT datasets from a highly sampled CBCT dataset using the Projection Striding (PS) and Phase-Binned Striding (PBS) techniques. In both techniques, the objective is to downsample and extract a collection of projections belonging to a specific phase from the remaining projections. At each step, the set of projections is reduced (depicted as a darkening of the projection) according to either a uniform projection stride sampling or phase-binning criteria. The upper-right panel shows that the PS method may have significant view aliasing artifacts in regions of reconstructed anatomy that have been severely undersampled. The PBS technique does not exhibit an undersampling problem since it has a more even angular and phase-distributed sampling of the projection space and results a more realistic representation of a clinical reconstruction. [Color figure can be viewed at wileyonlinelibrary.com]

2.E. Groupwise registration and motion modeling

Groupwise 4D (3D + t) registration was used to create a representative respiratory motion model instead of using a conventional, frame-to-frame registration approach. Groupwise registration simultaneously registers more than two images in the spatial domain, S , as opposed to pairwise registration, which co-registers a pair of images. For spatiotemporal data, groupwise registration simultaneously registers N temporal frames of a 4D image $I(\mathbf{x}, t)$, $t \in T = \{1, 2, \dots, N\}$, to a common reference frame using a time-varying transform $\mathbf{h}_\mu(\mathbf{x}, t)$. In this work, the transformation $\mathbf{h}_\mu(\mathbf{x}, t)$ was parameterized using a linear combination of B-splines with weights μ and the 4D images registered were constructed from a stack of $N = 10$ temporal image frames each corresponding to a phase-binned reconstruction from a different point in the respiratory cycle. The common reference frame may be an average coordinate system that does not correspond to one of the original 4D image,¹⁸ or the coordinate system of one of the temporal frames.¹⁹ Both approaches were used and compared.

$$C_{IF}(\mu) = \frac{1}{|S||T|} \sum_{\mathbf{x} \in S} \sum_{t \in T} (I(\mathbf{h}_\mu(\mathbf{x}, t)) - \bar{I}_\mu(\mathbf{x}))^2 \quad (1)$$

$$\bar{I}_\mu(\mathbf{x}) = \frac{1}{|T|} \sum_{t \in T} I(\mathbf{h}_\mu(\mathbf{x}, t)) \quad (2)$$

$$C_{FF}(\mu) = \frac{1}{|S||T|} \sum_{\mathbf{x} \in S} \sum_{t \in T} (I(\mathbf{h}_\mu(\mathbf{x}, t)) - I_{FF}(\mathbf{x}))^2 \quad (3)$$

$$I_{FF}(\mathbf{x}) = I(\mathbf{x}, t_{FF}) \quad (4)$$

Two hierarchical, groupwise image registration methods were assessed. The first evaluated method, a modified version of the groupwise method proposed by Metz et al.¹⁸ detailed in

Eq. (1), was used to compute the cost of a transform, $\mathbf{h}_\mu(\mathbf{x}, t)$, mapping the frames of the 4D-CBCT image to an iteratively computed mean reference frame, $\bar{I}_\mu(\mathbf{x})$, seen in Eq. (2), by minimizing the variance of voxel intensities along the temporal dimension of the 4D image. This mean reference frame (MF) method initially included a bending energy penalty term²⁶ in the cost function to restrict the flexibility of the transform. However, after preliminary testing, it was found that the bending energy penalty did not have much effect on the registrations evaluated, and it was excluded from all final registrations.

The second method assessed in this study registered all phases of the 4D image set to a target image phase $I_{FF}(\mathbf{x})$, as seen in Eq. (3). Instead of defining the deformation between the initial image and some iteratively generated mean image, this second fixed-reference frame (FF) approach directly computes the deformation to a reference, $I_{FF}(\mathbf{x})$, seen in Eq. (4), for a chosen time point, t_{FF} , in the respiratory cycle. For this study, the end of inhalation frame of the 4D image was selected as the fixed frame because it generally possessed the least sorting-related artifact relative to the other frames. While the cost functions of the two approaches are otherwise very similar, in implementation, the FF method is less computationally expensive and offers modest speed improvements over the MF approach.

All registrations were implemented in elastix v4.8.^{27,28} Both methods were applied using a hierarchical registration scheme consisting of two resolution levels. For both methods, the first resolution used a $2 \times$ downsampled image and computed cubic B-spline transform parameters on a 32 mm grid spacing. A full-resolution image and 16 mm grid spacing were used in the second resolution of both methods. Anatomical masks were used to confine the registration to the subject body in the image and minimize the influence of view aliasing and reconstruction artifacts in nonsubject regions.

For both techniques, the result of the registration is a spatiotemporal transformation represented by a set of b-spline transform parameters (i.e., control points). A dense

deformation vector field (DVF) was created using elastix, which was subsequently used as the motion model for motion-compensated reconstruction.

2.F. Motion-compensated reconstruction

To improve image quality using knowledge of subject motion, MoCo reconstruction was employed. The motion-compensated algorithm utilized in this study, originally described by Rit *et al.*,⁹ backprojects the projection data into the subject space to produce a tomographic image. However, unlike the FDK approach which backprojects along straight lines of x-ray acquisition, the algorithm used in this study backprojects along curved trajectories that compensate for motion. In this study, the motion model computed in the preceding groupwise registration facilitates this warping in both the spatial and temporal domains. The RTK implementation of this motion-compensated FDK algorithm was used to perform the MoCo reconstruction. Rather than using the phase-binned projection subsets produced for the initial FDK reconstruction, all available projection data were utilized in the reconstruction of the MoCo image volumes.

2.G. Workflow experiments

To evaluate the impact of each component of the MoCo workflow on the quality of the reconstructed images, two experiments were performed. First, the impact of type of groupwise registration (MF vs FF) on image quality was assessed. While both the MF and FF methods discussed have the demonstrable potential to produce a realistic model of subject deformation, they vary in computational expense and complexity. In the second experiment, to evaluate the tangible contribution of MoCo reconstruction toward improving image quality, the results of a workflow with and without the MoCo reconstruction component were compared. The workflow without MoCo is termed the “registration-only” result, as it consists of using the motion model from groupwise registration to deform each frame of the 4D image to the reference, and then averaging the registered 4D image along the temporal dimension to produce an improved 3D image at the reference image position.

Additionally, evaluation of method performance was carried out for the two proposed motion compensation methods using both the full and the downsampled projection datasets to demonstrate proof-of-concept using clinically representative images. Analysis of image quality improvement was completed for both the registration-only results and the MoCo reconstruction results. In total four methods (MF vs FF, $\pm$ MoCo) were evaluated for each subject in each of the two datasets (full and downsampled projection sets).

2.H. Image quality assessments

Image quality performance of the proposed methods was evaluated by visual inspection and three quantitative measures: (a) tissue interface sharpness (TIS), a measure of the sharpness of high-contrast tissue interfaces subject to motion, (b) image noise in homogeneous regions of interest, and (c)

local contrast-to-noise ratio (CNR) measured at fiducial markers. The details of these assessments are provided in the following subsections.

2.H.1. Tissue interface sharpness

Motion blurring in reconstructed CBCT images is most visible in the direction of motion, which should be improved by 4D and MoCo reconstruction methods. To characterize the degree of motion blur in the superior-inferior direction, tissue interface sharpness (TIS) is defined as the magnitude of the intensity gradient across interfaces approximately parallel to the axial imaging plane. Notably, high-contrast tissue interfaces that meet this definition include the lung-diaphragm interface and medial wall of the bronchus intermedius.

To compute the TIS semiautomatically, we identified three locations of interest, field of view (FOV) permitting, at both the apex of the lung–diaphragm and the bronchus intermedius interfaces and sampled image intensities in rays traversing the selected interfaces. To reduce the influence of noise in the TIS assessments, a set of rays, $\vec{r}$, spaced one voxel width apart were cast along the superior-inferior (Z) axis at each identified location: forming an $N \times N$ voxel sampling region in the axial (X-Y) plane centered about the voxel (x_0, y_0) as defined in Eq. (5). In this study, a sampling region width of $N = 5$ was used to produce a total of 25 sampled rays per TIS measurement. Following intensity normalization of the set of rays to a range of $[0,1]$, a sigmoid curve function, $S_i(z)$, as defined in Eq. (6), was fit to each ray intensity profile, $r_i(z)$. Shown in Eq. (7), for each ray index, the coefficient b_i , which characterizes the peak slope of the fit sigmoid curve, was recorded. The remaining parameter, c_i , corresponds to the offset of the interface boundary along the intensity profile dimension as determined by the sigmoid inflection point. Per Eq. (8), averaging the individual ray measures produces the TIS measure for a given interface location. A diagram depicting the generalized method is shown in the left panel of Fig. 3.

$$\begin{aligned} \vec{r}_i(z) &= \vec{r}_i(x_i, y_i, z) \\ &= \vec{r}_i \left(\left(x_0 + \left(\frac{(i-1) - (i-1) \bmod N}{N} \right) - \left(\frac{N-1}{2} \right) \right), \left(y_0 + ((i-1) \bmod N) - \left(\frac{N-1}{2} \right) \right), z \right), \text{ where } \{N \in \mathbb{Z} | N \bmod 2 = 1\} \\ &\quad \text{and } \{i \in \mathbb{Z} | [1, N^2]\} \end{aligned} \quad (5)$$

$$S_i(z) = \frac{1}{1 + e^{-b_i \cdot (z - c_i)}} \quad (6)$$

$$b_i = \frac{-\ln \left(\frac{1 - S_i(z)}{S_i(z)} \right)}{(z - c_i)} \quad (7)$$

FIG. 3. Illustrations of image quality assessments. Left: Tissue Interface Sharpness (TIS) is computed using a sigmoid curve fit to intensity profiles across high-contrast tissue interfaces. Right: ROI-based sampling of image noise in air, soft tissue, and the aorta permits direct characterization of visual image quality. [Color figure can be viewed at wileyonlinelibrary.com]

$$TIS = \frac{1}{N^2} \sum_{i=1}^{N^2} b_i \quad (8)$$

It is important to note that while the initial 4D-CBCT image suffers from noise and reconstruction artifact degradation, the interfaces evaluated in this study are expected to be crisp in the sense that little motion blurring exists in the initial 4D reconstruction and the corresponding gradient of image intensities at these locations are expected to be high. The introduction of any motion blurring which may be sourced to errors in respiratory motion modeling or model application in the MoCo reconstructions would therefore be expected to reduce the measured intensity gradients at these interfaces. Therefore, comparisons of method performance were completed by evaluating the recovery of TIS in the MoCo images as a relative ratio to that of the initial 4D-CBCT reconstruction.

2.H.2. ROI sampling of image noise

A second measure of visual image quality is established by the observation of statistical image noise in key regions of interest (ROIs). Key sites for evaluation were identified: (a) the aorta, for describing quality of the central region of the reconstructed image; (b) the soft tissue, for describing the variability of image intensities in a reasonably homogeneous region toward the periphery of the FOV; and (c) the air space around the subject, for capturing the effects of the streaking and view aliasing in a homogeneous region of the image. This method is also depicted in Fig. 3.

2.H.3. Local contrast-to-noise ratio for fiducial markers

The third quantitative measure of image quality utilized was an assessment of the local contrast-to-noise ratio (CNR) in the region immediately surrounding a fiducial marker. The purpose

of this evaluation was to quantify the visibility of fine, high-contrast structures, like blood vessels or fiducial markers. The local CNR assessment method for fiducial markers, originally proposed by Shieh *et al.*,²⁹ was adopted for use in this study. Of the subjects included in this study, six had between 1 and 3 visible, implanted markers for which an average measure of local CNR was computed.

3. RESULTS

3.A. Quantitative performance

Performance of the proposed methods was characterized for each study subject using the image quality assessment measures described in the preceding section. Notably, TIS recovery was only evaluated at the diaphragm interface for the 10 subjects where the structure was visible within the field of view. TIS recovery at the bronchus intermedius, and all three noise reduction measures were evaluated for each of the 19 subjects. In Table I, the average method performance in the full projection subject dataset is highlighted. Figure 4 demonstrates the variability in image quality improvement for each subject in the full projection dataset under the evaluated methods by presenting the average of TIS recovery metrics for the three evaluated interfaces as a relative percentage of those from the initial 4D-CBCT reconstruction. Similarly, Table II and Fig. 5 present comparable results for the down-sampled projection subject dataset. Local CNR was assessed for six subjects containing a total of 11 fiducial markers. Figure 6 depicts the range of fiducial marker CNR improvement relative to that of the initial FDK reconstruction for each of the evaluated methods across both the full and downsampled projection datasets.

3.B. Visual observations of image enhancement

The researchers observed that the proposed methods demonstrated the ability to remove view aliasing artifacts and clarify the visibility of small structures (i.e., vessel trees) for

TABLE I. Summary of method performance averaged over the 19 subjects of the full projection dataset as quantified by the image quality measures of TIS at the diaphragm and bronchus intermedius and noise of three ROIs (air, aorta, and soft tissue). Percentage metrics presented represent the relative percentage of TIS recovered across the three evaluated interfaces relative to the initial 4D-CBCT reconstruction.

	Reference frame and method									
	Average of 4D-CBCT		Fixed frame (ff); registration-only		Fixed frame (FF); MoCo recon.		Mean frame (MF); registration-only		Mean frame (MF); MoCo recon.	
	Std.		Std.		Std.		Std.		Std.	
	Mean (%)	Dev. (%)	Mean (%)	Dev. (%)	Mean (%)	Dev. (%)	Mean (%)	Dev. (%)	Mean (%)	Std. Dev. (%)
TIS recovery										
Diaphragm	95.7	(39.0)	98.5	(20.0)	98.4	(19.9)	122	(43.8)	121	(43.4)
Bronchus intermedius	82.5	(20.7)	95.4	(9.24)	94.9	(10.3)	94.3	(20.4)	93.8	(21.8)
ROI noise reduction										
Ambient air	64.4	(9.95)	64.1	(8.86)	7.6	(11.2)	62.7	(9.98)	67.6	(11.4)
Aorta	59.5	(6.82)	57.0	(5.86)	57.3	(5.44)	59.0	(5.55)	58.2	(5.55)
Soft tissue	52.3	(11.7)	42.7	(8.41)	43.5	(9.42)	40.5	(8.64)	41.7	(8.95)

FIG. 4. Variability in individual subject TIS recovery across each of the assessed methods. Results are shown for the analysis of the full projection subject dataset. [Color figure can be viewed at wileyonlinelibrary.com]

the dataset utilized. Figures 7 and 8 offer representative examples of the types of improvement achievable through the proposed MoCo methods for reader inspection by highlighting one of the evaluated subjects using the full projection dataset and the downsampled datasets, respectively. The left-most column in both figures contains the preliminary reconstructions that were used as the basis for comparison for the results of each proposed method. It is important to note that the reference frame (initial FDK) for the FF and MF methods differ; the end of inhalation image frame is used as the reference for the FF method while the mean respiration frame is used for the MF method. The end of inhalation frame shown for the FF method highlights the considerable amount of noise and view aliasing typically observed in 4D-CBCT acquisitions while the mean reference frame pictured for the

MF method illustrates the potential for motion blurring in both 3D- and 4D-CBCT reconstructions.

Looking at the registration-only results, the FF method shows a visible improvement in image noise and a correspondingly improved visibility of finer vessel structures in the upper lobes of the lungs as seen in the coronal view. Alternatively, the MF method results reveal a reduction in motion blurring along the diaphragm and larger vessel interfaces. The addition of MoCo reconstruction further enhances the results of the registration-only approaches by minimizing the incidence of the view aliasing artifacts as seen in the axial views. The most visible examples of these improvements can be seen in views of the clinically representative downsampled dataset results, shown in Fig. 8. Similar image enhancement effects were also observed for the other 18 subjects in the study.

TABLE II. Summary of method performance averaged over the 19 subjects of the downsampled projection dataset as quantified by the image quality measures of TIS at the diaphragm and bronchus intermedius and noise of three ROIs (air, aorta, and soft tissue). Percentage metrics presented represent the relative percentage of TIS recovered across the three evaluated interfaces relative to the initial 4D-CBCT reconstruction.

	Reference frame and method									
	Average of 4D-CBCT		Fixed frame (FF); registration-only		Fixed frame (FF); MoCo recon.		Mean frame (MF); registration-only		Mean frame (MF); MoCo recon.	
	Std.				Std.		Std.		Std.	
	Mean (%)	Dev. (%)	Mean (%)	Std. Dev. (%)	Mean (%)	Dev. (%)	Mean (%)	Dev. (%)	Mean (%)	Dev. (%)
TIS recovery										
Diaphragm	78.6	(57.1)	86.5	(46.4)	87.2	(47.4)	101	(68.1)	99.3	(65.3)
Bronchus intermedius	77.7	(27.2)	87.5	(22.9)	85.4	(21.5)	88.4	(27.7)	85.2	(26.7)
ROI noise reduction										
Ambient air	65.4	(4.29)	65.5	(4.82)	69.7	(4.78)	65.4	(3.58)	69.5	(4.62)
Aorta	64.7	(2.78)	66.8	(2.94)	65.8	(3.84)	66.7	(3.83)	65.5	(2.82)
Soft tissue	59.8	(6.28)	58.1	(6.19)	58.0	(5.97)	57.3	(7.39)	57.9	(6.26)

FIG. 5. Variability in individual subject TIS recovery across each of the assessed methods. Results are shown for the analysis of the downsampled projection subject dataset. [Color figure can be viewed at wileyonlinelibrary.com]

4. DISCUSSION

In this study, we presented two methods of performing motion modeling using groupwise deformable image registration and adopted the use of MoCo reconstruction to enhance the quality of reconstructed 4D-CBCT images. While this study does not represent the first groupwise registration nor the first MoCo reconstruction of 4D-CBCT images, to the best of our knowledge, it is the first time that both have been combined to produce a purely data-driven, *a posteriori* reconstruction technique aimed at improving image quality. The purely data-driven methods proposed produce representative models of subject respiration directly from day-of-treatment CBCT acquisitions and circumvent the need for either an *a priori*

motion model or an external respiratory signal. Applying this motion model during MoCo reconstruction facilitates the use of all available CBCT projection data to produce a visibly clearer reconstruction impacted by fewer view aliasing artifacts.

Two resulting images were produced for each of the two proposed methods, characterizing the outcome of applying the groupwise registration transform to the initial FDK reconstruction as well as that of the MoCo reconstruction. For the FF method, the registration-only images show improvements in image noise in bulk tissue and reductions in the effects of view aliasing. The corresponding results from the MoCo reconstruction suggest additional gains in mitigating streaking (view aliasing) effects as well as slight enhancement of fine structures as indicated by the CNR assessments

FIG. 6. Plot of relative improvement of fine structure contrast, as measured by local CNR of fiducial markers, compared to that of the initial FDK reconstruction. Outliers, shown as points, are improvement observations that fall more than 1.5 times the interquartile range above the upper quartile and below the lower quartile. [Color figure can be viewed at wileyonlinelibrary.com]

performed. The MF method registration-only approach demonstrated considerable reduction in motion blurring compared to the initial mean-frame reference. This is particularly apparent in subjects where the diaphragm is visible. The MoCo reconstruction results for the MF method reveal similar motion blurring mitigation as well as reductions in view aliasing impact. While the differences between the registration-only and MoCo reconstruction results are at best subtle when using the full projection dataset, as seen in Fig. 7, they become more apparent in the results for the downsampled dataset, seen in Fig. 8.

To consider the quantitative performance of the proposed methods, evaluation of TIS recovery and noise reduction in three ROIs was made. Looking at TIS recovery, which serves as a surrogate measure for the arresting of subject respiratory motion, Figs. 4 and 5 show that each proposed method was generally able to recover a substantial amount of the sharpness of the initial tissue interface. For the full and downsampled cases, the MF method was routinely able to recover more of the interface sharpness than the FF method at both the diaphragm and the bronchus intermedius interfaces and on average enhanced the contrast of the assessed diaphragm interfaces beyond that of the reference image. It is suggested that the iteratively computed reference frame used by the MF method may be the source of this discrepancy in performance, providing a lower noise reference point for the group-wise registration. The FF method, on the other hand, is limited by the quality of the fixed-reference frame selected from the initial 4D-CBCT image. It is for this reason that the end of inhalation reference was chosen, as image frames reconstructed at this reference point in the respiratory cycle can benefit from an increased number of projections being acquired at the momentary pause in subject motion and greater initial reconstruction quality.

FIG. 7. Coronal and axial slices of initial FDK image, the registration-only results, and MoCo reconstruction results for PT-105 using the full projection dataset.

The characterization of noise in three ROIs was used to provide insight into the visual quality of the images. Evaluation of ROIs in the aorta, soft tissue, and air media were made to provide quality measures in well-sampled, homogeneous, and view-aliased regions, respectively. Table I shows that for the cases which employed the full projection dataset noise reductions achieved for each of the proposed methods fell slightly short of that realized by averaging the temporal frames of the initial 4D-CBCT images. However, in the downsampled cases, it appears that proposed methods achieve approximately the same amount of noise reduction as averaging. This is suspected to be a consequence of the

FIG. 8. Coronal and axial slices of initial FDK image, the registration-only results, and MoCo reconstruction results for PT-105 using the downsampled projection dataset.

improved initial image quality in the full dataset cases, where temporal averaging may appear to reduce image noise in the ROIs, but contribute to reduction in the visibility of small structures and tissue texture.

It is clear from this variability in TIS and ROI noise performance that a method which benefits one subject may not benefit another; in other words, there is no apparent one-size-fits-all method for motion compensation. It is, however, worth noting that Figs. 4 and 5 demonstrate that for approximately half of subjects evaluated, each of the proposed methods appear to achieve similar image quality outcomes. In the course of this study, evaluation of subject factors that might contribute to the success or failure of a particular method for a particular subject was considered, but was ultimately deemed to be outside the scope of this work and, instead, focused on demonstrating the feasibility of the proposed methods. Analysis of underlying subject factors, acquisition conditions, and their impact on the performance of the proposed methods may be undertaken as part of future work.

To contextualize the proposed method and presented results, comparisons can be drawn to related motion-compensated reconstruction methods previously described in the literature. The MoCo methods presented by Brehm *et al.*^{11,12} initially demonstrated the feasibility of computing a respiratory motion model directly from the registration of CBCT datasets (both simulated and measured) instead of prior CT images for the purposes of performing a motion-compensated reconstruction. These studies utilized a cyclically constrained, demons-based registration algorithm to model the deformations between adjacent 3D-CBCT frames and subsequently composed these deformations to produce a full 4D motion model. The proposed methods in this work instead utilize variations of a groupwise registration algorithm to directly compute a 4D motion model from 4D-CBCT images thereby overcoming the need to compose successive DVFs and the

potential error associated with this process. The SMEIR method presented by Wang and Gu¹⁴ notably demonstrated an ability to improve image quality by simultaneously estimating respiratory motion during an iterative reconstruction. While effective, it was demonstrated that the SMEIR method had the potential to oversmooth certain anatomical features (i.e., tumors or vessels) resulting from the use of a prior image reconstructed using Total Variation leading to errors in local motion estimates which in turn affect reconstructed image quality. Furthermore, the researchers indicated that there is potential for the method to error in cases of irregular patient respiration where the number of projections utilized to reconstruct any given frame of the 4D image varies considerably from the next. The proposed method does not incorporate any explicit image smoothing in its initial FDK reconstruction but still accomplishes a smooth deformation via a B-spline representation of the registration transformation. Additionally, the methods presented in this study have been shown to robustly produce improved images in cases of varying projection counts per reconstructed frame, as is the case in the clinical dataset used in this study. Finally, a number of alternative methods of improving image quality, including the works of Shieh *et al.*³⁰ with the AAIR method and Chen *et al.*³¹ with the PICCS method, have shown promising results using iterative reconstruction techniques which capitalize on data in prior reconstructed images instead of motion compensation. The approach we evaluated here could be integrated with these alternative techniques, by using one of these iterative reconstruction methods as the initial reconstruction input into the groupwise registration algorithm.

There are several known issues and limitations associated with the data and methods proposed in this study. First, in addition to the issues identified by Hugo *et al.*²² for the dataset utilized in this study, the clinically representative downsampled dataset was produced artificially rather than through a standard clinical acquisition. Second, and on a related note, this retrospective dataset did not afford any sort of standard, such as gated CBCT acquisitions, to allow for a comprehensive characterization of the proposed methods abilities to accurately reconstruct subject anatomy. Evaluating the reconstruction accuracy of the proposed methods along these lines is the subject of future planned studies. Finally, while the methods proposed offer an avenue for improving the quality of 4D-CBCT images taken under clinical conditions, the implemented workflow remains far from optimized for a clinical environment. While any tomographic reconstruction is computationally demanding, the proposed techniques carry the added expenses of generating motion models and performing MoCo reconstruction. To illustrate this point, it is worth noting that in their nonoptimized state, the workflows presented can take between 6 and 10 h to complete the MoCo reconstruction of a single subject. However, both groupwise registration and MoCo reconstruction are highly parallelizable, e.g., through implementation on graphics processing units. The results of this study can aid in selection of a workflow, which we then plan to optimize for eventual clinical use.

5. CONCLUSIONS

The overarching objectives of this study were achieved through the demonstration of the feasibility of using a purely data-driven, *a posteriori* respiratory motion modeling and reconstruction compensation method to improve 4D-CBCT image quality. We demonstrated that groupwise registration methods can be successfully applied to 4D-CBCT to produce models of subject motion under clinically realistic acquisition conditions. We also demonstrated that it is possible to improve the quality of 4D-CBCT image frames by performing groupwise registration of the initial 4D image and averaging the resulting aligned frames in order to reduce image noise and motion blurring. Applying the registration-derived motion model in a MoCo reconstruction, it was shown that 4D-CBCT image quality can be enhanced with visible reduction in view aliasing artifact and mitigation of motion blurring. While response varied subject-to-subject, the proposed methods, on average, each demonstrated a similar ability to improve CBCT image quality, under both clinically superior and clinically relevant conditions, as measured by the defined TIS and ROI noise metrics.

ACKNOWLEDGMENTS

This work was supported in part by a research grant from the National Cancer Institute of the National Institutes of Health under award number R01CA166119. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CONFLICTS OF INTEREST

We disclose the following potential conflicts of interest in the manuscript: Virginia Commonwealth University has a research agreement with Philips Medical Systems and a licensing agreement with Varian Medical Systems. EW receives royalties from UpToDate. GEC has received research funding from Roger Koch.

^{a)}Author to whom correspondence should be addressed. Electronic mail: gdhugo@wustl.edu.

REFERENCES

- Jaffray DA, Siewerdsen JH, Wong JW, Martinez AA. Flat-panel cone-beam computed tomography for image-guided radiation therapy. *Int J Radiat Oncol Biol Phys.* 2002;53:1337–1349.
- Létourneau D, Wong JW, Oldham M, et al. Cone-beam-CT guided radiation therapy: technical implementation. *Radiother Oncol.* 2005;75:279–286.
- Hugo GD, Liang J, Campbell J, Yan D. On-line target position localization in the presence of respiration: a comparison of two methods. *Int J Radiat Oncol Biol Phys.* 2007;69:1634–1641.
- Rit S, Nijkamp J, van Herk M, Sonke J-J. Comparative study of respiratory motion correction techniques in cone-beam computed tomography. *Radiother Oncol.* 2011;100:356–359.
- Sonke J-J, Zijp L, Remeijer P, van Herk M. Respiratory correlated cone beam CT. *Med Phys.* 2005;32:1176–1186.
- Bissonnette J-P, Franks KN, Purdie TG, et al. Quantifying interfraction and intrafraction tumor motion in lung stereotactic body radiotherapy using respiration-correlated cone beam computed tomography. *Int J Radiat Oncol.* 2009;75:688–695.
- Kincaid RE Jr., Yorke ED, Goodman KA, Rimner A, Wu AJ, Mageras GS. Investigation of gated cone-beam CT to reduce respiratory motion blurring. *Med Phys.* 2013;40:41717.
- Li T, Schreiber E, Yang Y, Xing L. Motion correction for improved target localization with on-board cone-beam computed tomography. *Phys Med Biol.* 2006;51:253–267.
- Rit S, Wolthaus J, Van Herk M, Sonke JJ. On-the-fly motion-compensated cone-beam CT using an a priori motion model. *Med Phys.* 2009;36:2283–2296.
- Dzyubak O, Kincaid R, Hertanto A, et al. Evaluation of tumor localization in respiration motion-corrected cone-beam CT: prospective study in lung. *Med Phys.* 2014;41:101918.
- Brehm M, Paysan P, Oelhafen M, Kunz P, Kachelrieß M. Self-adapting cyclic registration for motion-compensated cone-beam CT in image-guided radiation therapy. *Med Phys.* 2012;39:7603–7618.
- Brehm M, Paysan P, Oelhafen M, Kachelrieß M. Artifact-resistant motion estimation with a patient-specific artifact model for motion-compensated cone-beam CT. *Med Phys.* 2013;40:101913.
- Brehm M, Sawall S, Maier J, Sauppe S, Kachelrieß M. Cardiorespiratory motion-compensated micro-CT image reconstruction using an artifact model-based motion estimation. *Med Phys.* 2015;42:1948–1958.
- Wang J, Gu X. Simultaneous motion estimation and image reconstruction (SMEIR) for 4D cone-beam CT. *IEEE Nucl Sci Symp Conf Rec.* 2013;40:101912.
- Hertanto A, Zhang Q, Hu Y-C, Dzyubak O, Rimmer A, Mageras GS. Reduction of irregular breathing artifacts in respiration-correlated CT images using a respiratory motion model. *Med Phys.* 2012;39:3070–3079.
- Zhang Q, Hu YC, Liu F, Goodman K, Rosenzweig KE, Mageras GS. Correction of motion artifacts in cone-beam CT using a patient-specific respiratory motion model. *Med Phys.* 2010;37:2901–2909.
- Seppenwoolde Y, Shirato H, Kitamura K, et al. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int J Radiat Oncol.* 2002;53:822–834.
- Metz CT, Klein S, Schaap M, van Walsum T, Niessen WJ. Nonrigid registration of dynamic medical imaging data using nD + t B-splines and a groupwise optimization approach. *Med Image Anal.* 2011;15:238–249.
- Vandemeulebroucke J, Rit S, Kybic J, Clarysse P, Sarrut D. Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. *Med Phys.* 2011;38:166–178.
- Balik S, Weiss E, Jan N, et al. Evaluation of 4-dimensional computed tomography to 4-dimensional cone-beam computed tomography deformable image registration for lung cancer adaptive radiation therapy. *Int J Radiat Oncol Biol Phys.* 2013;86:372–379.
- Roman NO, Shepherd W, Mukhopadhyay N, Hugo GD, Weiss E. Interfractional positional variability of fiducial markers and primary tumors in locally advanced non-small-cell lung cancer during audiovisual biofeedback radiotherapy. *Int J Radiat Oncol Biol Phys.* 2012;83:1566–1572.
- Hugo GD, Weiss E, Sleeman WC, et al. A longitudinal four-dimensional computed tomography and cone beam computed tomography dataset for image-guided radiation therapy research in lung cancer. *Med Phys.* 2017;44:762–771.
- Zijp L, Sonke J, van Herk M. Extraction of the respiratory signal from sequential thorax cone-beam X-ray images. *Int Conf Use Comput Radiat Ther.* 2004;63:507–509.
- Rit S, Vila Oliva M, Brousmiche S, Labarbe R, Sarrut D, Sharp GC. The reconstruction toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the insight toolkit (ITK). *J Phys Conf Ser.* 2014;489:12079.
- Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. *J Opt Soc Am A.* 1984;1:612–619.
- Rueckert D, Sonoda LI. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging.* 1999;18:712–721.

27. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29:196–205.
28. Shamonin DP, Bron EE, Lelieveldt BPF, Smits M, Klein S, Staring M. Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease. *Front Neuroinform*. 2014;7:1–15.
29. Shieh C-C, Kipritidis J, O'Brien RT, Keall PJ. Image quality in thoracic 4D cone-beam CT: a sensitivity analysis of respiratory signal, binning method, reconstruction algorithm, and projection angular spacing. *Med Phys*. 2014;41:1–18.
30. Shieh C-C, Kipritidis J, O'Brien RT, Cooper BJ, Kuncic Z, Keall PJ. Improving thoracic four-dimensional cone-beam CT reconstruction with anatomical-adaptive image regularization (AAIR). *Phys Med Biol*. 2015;60:841–868.
31. Chen G-H, Tang J, Leng S. Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med Phys*. 2008;35:660–663.