Advanced 4-dimensional cone-beam computed tomography reconstruction by combining motion estimation, motion-compensated reconstruction, biomechanical modeling and deep learning

You Zhang*, Xiaokun Huang and Jing Wang

Abstract

4-Dimensional cone-beam computed tomography (4D-CBCT) offers several key advantages over conventional 3D-CBCT in moving target localization/delineation, structure de-blurring, target motion tracking, treatment dose accumulation and adaptive radiation therapy. However, the use of the 4D-CBCT in current radiation therapy practices has been limited, mostly due to its sub-optimal image quality from limited angular sampling of cone-beam projections. In this study, we summarized the recent developments of 4D-CBCT reconstruction techniques for image quality improvement, and introduced our developments of a new 4D-CBCT reconstruction technique which features simultaneous motion estimation and image reconstruction (SMEIR). Based on the original SMEIR scheme, biomechanical modeling-guided SMEIR (SMEIR-Bio) was introduced to further improve the reconstruction accuracy of fine details in lung 4D-CBCTs. To improve the efficiency of reconstruction, we recently developed a U-net-based deformation-vector-field (DVF) optimization technique to leverage a population-based deep learning scheme to improve the accuracy of intra-lung DVFs (SMEIR-Unet), without explicit biomechanical modeling. Details of each of the SMEIR, SMEIR-Bio and SMEIR-Unet techniques were included in this study, along with the corresponding results comparing the reconstruction accuracy in terms of CBCT images and the DVFs. We also discussed the application prospects of the SMEIR-type techniques in image-guided radiation therapy and adaptive radiation therapy, and presented potential schemes on future developments to achieve faster and more accurate 4D-CBCT imaging.

Keywords: Cone-beam computed tomography, Image reconstruction, Motion estimation, Biomechanical modeling, Deep learning

Introduction

Accurate tumor/target localization is key to safe, precise and effective radiotherapy [1]. Cone-beam computed tomography (CBCT) imaging has become a standard-of-care in a majority of the radiotherapy clinics, with its successful capture of volumetric anatomical information to guide accurate on-board target localization and setup correction [2, 3]. In principal, the 3-dimensional CBCT (3D-CBCT) technique acquires 2D cone-beam projections from varying beam angles, usually covering an angle span of at least 200° [4], to reconstruct 3D volumetric information. Compared with 2D projections, the overlaysing effects have been removed by 3D-CBCT to allow better soft-tissue contrast and more accurate structure localization in 3D [2]. However, there are multiple remaining issues with the current CBCT technique. One major issue is the motion-induced imaging artifacts and blurring for motion-involved sites such as lung and liver, which may severely reduce the tumor localization accuracy [5, 6]. The conventional 3D-CBCT technique acquires cone-beam projections and reconstructs them into a single CBCT volume without considering the motion of anatomical structures...
across the projections. Due to the position variations of anatomical structures, combining all projections into one reconstruction will lead to motion blurriness and inaccurate target localization. To effectively suppress the motion blurriness, methods have been proposed to develop the respiratory-correlated CBCT, also known as 4-dimensional CBCT (4D-CBCT), to capture the motion in the fourth dimension (in addition to the three spatial dimensions) [7–9]. The key idea of 4D-CBCT is phase-sorting [10], which sub-groups the cone-beam projections into different respiratory phase bins according to their relative positioning on a nominal respiratory cycle. The relative positioning can be determined via an external landmark, such as the real-time position management system or the Anzai Belt [11, 12]; or via an internal landmark, such as the diaphragm position or the fiducial marker position [7]. More advanced signals based on Fourier transform have also been investigated to be successful [10]. Based on these positioning signals, projections were sorted into different bins. Within each phase bin, the projections are semi-static with minimal intra-phase motion. CBCTs reconstructed from these phase-sorted projections will have the motion blurriness subdued, leading to more accurate delineation and localization of anatomical structures. Stacking these phase-specific CBCT images together will also reveal the full motion trajectory of anatomical structures to guide tumor targeting and organ-of-risk sparing. Though conceptually concise and convenient to implement, the use of 4D-CBCT in clinics is currently limited. One major concern is that the phase-sorting process employed by 4D-CBCT leads to insufficient projections inside each phase bin (angular under-sampling), causing severe imaging streaking artifacts to the clinical CBCT reconstruction technique, the Feldkamp-Davis-Kress (FDK) algorithm [13, 14].

One straightforward solution is to acquire more projections to achieve sufficient angular sampling even after phase-sorting [4]. Such a strategy, however, is not practical due to the accompanying excessive imaging dose, which may induce secondary cancers [15]. Acquiring more projections also prolongs the imaging and treatment time, which may increase patient on-board position deviations [16], and add additional logistic burdens to the clinics. Many groups have tried to develop new reconstruction techniques, mostly iterative in nature, to improve the 4D-CBCT image quality from under-sampled acquisition. Many of these techniques rely on regularization techniques via metrics like total variation [17, 18], non-local means [19, 20] or wavelet frames [21] to explore the data sparsity and reduce the noises/artifacts of under-sampling. Substantial image quality improvement has been observed, and quantitatively validated via metrics like signal-to-noise ratio, root-mean-squared-error (RMSE) or universal quality index (UQI) [22]. However, these techniques are susceptible to displace or smooth out fine details and low-contrast anatomies from the reconstructed images.

Another type of CBCT reconstruction technique tries to incorporate prior information into the reconstruction process, such as the 2D-3D deformation method [23–28]. Instead of directly reconstructing CBCT images from acquired projections, the 2D-3D deformation method views the new CBCT volume as a deformation of prior CT/CBCT images, and translates the CBCT reconstruction into a deformation-vector-field (DVF) optimization problem. For each to-be-deformed CT/CBCT image, the DVFs usually compose of three matrices in 3D, each matrix with the same dimension as the to-be-deformed CT/CBCT image. The three matrices indicate the deformation along the three Cartesian directions, x, y and z, respectively [25]. The 2D-3D deformation algorithm iteratively optimizes the DVF, such that it deforms the prior image until the digitally-reconstructed-radiographs (DRRs) of the deformed image match the acquired cone-beam projections. The 2D-3D deformation technique can not only generate 4D-CBCT images, but provide DVFs to potentially allow automatic tumor localization, target tracking, dose accumulation and adaptive radiation therapy [29–32]. The incorporation of high-quality prior information also introduces more accurate Hounsfield units [30], and allows further imaging dose reduction by acquiring fewer projections for reconstruction. However, this technique cannot reconstruct non-deformation-induced intensity changes in new CBCT volumes, since the new CBCT is simplified as a purely deformed volume of the prior image [33].

In addition to these two types of techniques, motion-compensated reconstruction is another reconstruction technique to address the 4D-CBCT under-sampling issue [34–37]. The motion-compensation technique reconstructs a reference phase CBCT image through an estimated inter-phase motion model [34]. The motion model relates other phases of the 4D-CBCT to the reference phase in the form of inter-phase DVFs. Through applying the inter-phase DVFs, motion-compensated reconstruction can combine projections from all phases to achieve sufficient angular sampling. After the reference phase is reconstructed, the other phases can be derived using corresponding inverse DVFs. Unfortunately, the motion-compensation technique is limited by the inter-phase motion model accuracy, since a priori-information-driven motion model can be invalidated after motion pattern changes [25, 34, 36]. To have the model up-to-date, studies have proposed to estimate an
on-board motion model directly between 4D-CBCT phase images reconstructed from under-sampled projections [35, 37]. However, the accuracy of these on-board motion models can still be impaired by the artifacts presented in the low-quality CBCT phase images.

In response to the current challenges of high-quality 4D-CBCT imaging, we have developed a simultaneous motion estimation and image reconstruction (SMEIR) technique [38–40], which combines total variation-based image regularization, 2D-3D deformation-driven motion model estimation, and motion-compensated reconstruction into a comprehensive reconstruction scheme. In comparison to the previous techniques, the SMEIR algorithm estimates an inter-phase motion model via 2D-3D deformation from a motion-compensated CBCT (mCBCT) and the phase-specific projections. The improved motion model is subsequently fed back into the motion-compensated reconstruction to update the mCBCT. Total variation-regularization is incorporated into the motion-compensated reconstruction to further improve the CBCT image quality. The resulting mCBCT, with improved accuracy and quality, is iteratively fed back into the 2D-3D deformation to dynamically update the motion model. The 2D-3D motion estimation and the motion-compensated image reconstruction form an iterative loop to continuously update the mCBCT image as well as the motion model until final convergence. The 4D-CBCT images at other phases are deformed from the mCBCT via the inverse DVFs simultaneously-optimized by the 2D-3D deformation algorithm. In this study, we detailed the general philosophy and workflow of the SMEIR algorithm, and introduced two new developments based on the original SMEIR algorithm: the biomechanical modeling-guided SMEIR (SMEIR-Bio) and the SMEIR algorithm with artificial intelligence (AI)-driven DVF fine-tuning (SMEIR-Unet) [41]. In comparison to the original SMEIR algorithm, the SMEIR-Bio algorithm introduced biomechanical modeling to improve the intra-lung DVF accuracy to better reconstruct the fine details in lung, and better capture their motion. Similar to SMEIR-Bio, the SMEIR-Unet algorithm was developed to fine-tune intra-lung DVFs, however via a deep learning-driven approach. Compared with SMEIR-Bio, SMEIR-Unet may allow improved efficiency with reduced complexity and computational load. Corresponding reconstruction results of both digital anthropomorphic phantoms and real lung patient data were presented to illustrate the strengths and weaknesses of different methods. In the end, a review was further included to discuss the application prospects of the developed algorithms, specifically in terms of real-time image-guided radiotherapy and adaptive radiotherapy.

Methods
The original SMEIR algorithm
The SMEIR algorithm integrates the motion estimation and image reconstruction to solve both the 4D-CBCT images and the corresponding inter-phase DVFs simultaneously. Compared with the conventional sequential scheme (Fig. 1), both the 4D-CBCT images and the DVFs are iteratively and dynamically updated by SMEIR to maximize the sharing of information to improve image and DVF accuracy.

The SMEIR algorithm is comprised of two major components, the motion estimation and the motion-compensated reconstruction. For motion estimation, the 2D-3D deformation technique is used, which is based on optimizing objective functions as shown in Eqs. 1 and 2:

\[
f_1(DVF^{0 \rightarrow t}) = \| p^t - A p^0 (x + DVF^{0 \rightarrow t}) \|^2_{L_2} + \beta \| E(DVF^{0 \rightarrow t}) \|
\]

\[
f_2(DVF^{t \rightarrow 0}) = \| p^0 - A p^t (x + DVF^{t \rightarrow 0}) \|^2_{L_2} + \beta \| E(DVF^{t \rightarrow 0}) \|
\]

\( p \) denotes the 2D, phase-binned cone-beam projections, with the superscript indicating the respiratory phase they belong to; “\( A \)” denotes the projecting matrix of the projections \( p \), which generates DRRs corresponding to the projections; “\( 0 \)” indicates the reference phase (motion-compensated phase); “\( t \)” denotes a general symbol representing each of all other phases; \( \mu \) denotes the CBCT image at the phase designated by its superscript; \( x \) denotes the 3D coordinates of the CBCT image; DVF denotes the deformation-vector-field; The \( \| * \|^2_{L_2} \) terms in Eqs. 1 and 2 calculate the sum of squared errors between the acquired projections \( p \) and the DRRs of the deformed image; \( E(\cdot) \) denotes a quadratic deformation energy term defined to regularize the DVF smoothness [25]; “\( \beta \)” balances the data fidelity term and the deformation energy term [24]. In general, the 2D-3D deformation algorithm is optimizing a DVF to deform 3D images based on matching 2D projections, as its name “2D-3D” suggests.

The objective functions of Eqs. 1 and 2 are designed as an inverse-consistent deformable registration to solve a pair of forward and inverse DVFs (\( DVF^{0 \rightarrow t} \) and \( DVF^{t \rightarrow 0} \)). The two objective functions are optimized sequentially and the output of one function is input into the other for cyclic iterations. The conversion between \( DVF^{0 \rightarrow t} \) and \( DVF^{t \rightarrow 0} \) is achieved through the following inverse-consistent constraint:
Based on the DVFs optimized by 2D-3D deformation, the SMEIR algorithm applies motion-compensated reconstruction through a modified simultaneous algebraic reconstruction technique (SART) [42], as shown in Eq. 5:

\[ \mu_{j}^{0,(k+1)} = \mu_{j}^{0,(k)} + \lambda \frac{\sum_{t,n} DVF_{jn}^{x \to 0} \left[ p_{jn} - \sum_{j'} a_{jn} \mu_{j'}^{0,(k)} \right]}{\sum_{t,n} DVF_{jn}^{x \to 0} a_{jn}} \]  

(5)

In Eq. 5, “ i ” and “ n ” denote voxels in the mCBCT image (reference phase) and the CBCT images at the other phases, respectively; “ i ” denotes the projecting ray at pixel location i within each projection; “ j ” denotes the total number of voxels intercepted by ray i; “ a ” denotes the intersection length of projecting ray i across each voxel; “ λ ” denotes the relaxation factor for SART; “ k ” indicates the iteration number. Compared with the conventional SART, the modified SART applies the voxel-wise deformation field \( DVF_{jn}^{x \to 0} \) to the correction term of each phase, which aligns these correction terms to the same coordinates as those of the reference image, to contribute them all towards the update of the reference image. Using all available information, the modified SART algorithm allows the reconstruction of a high-quality mCBCT with under-sampling streaking artifacts removed. After SART, we applied total-variation regularization on the motion-compensated image to further reduce the image noises and artifacts to improve its quality. The updated mCBCT, \( \mu^{0} \), is further fed as input into the 2D-3D deformation (Eqs. 1–4) to form an iterative loop until final convergence.
Figure 2 compares the reconstructed images by different techniques. The images shown in column (a) were reconstructed by FDK using non-phase-sorted projections from all respiratory phases. Though with minimal streaking artifacts, these FDK images presented prominent motion blurriness from mixed-phase reconstruction. By reconstruction using only the phase-sorted projections, the FDK images in column (b) display reduced motion blurriness, but with amplified under-sampling streaking artifacts and noises. With the de-noising effects from total-variation regularization, the images reconstructed by ART-TV substantially reduced the imaging noises and artifacts as compared to (b), however at the cost of missing fine details (as indicated by the arrows). In contrast, the CBCT images reconstructed by the SMEIR algorithm [column (d)] not only removed the noises/artifacts, but the fine details of the images were well preserved through motion-compensated reconstruction with a high-quality motion model.

The SMEIR-bio algorithm

Though the SMEIR algorithm has proven generally accurate in reconstructing geometrical and intensity information for 4D-CBCT images, some fine anatomical structures, which are small in size, were found not well reconstructed by SMEIR. The organ of lung includes many such structures, including vessels, small bronchi-oles, and small nodules. Though small in size, these fine structures can serve important landmarks for diagnosis and treatment toxicity evaluation [44]. Accurate reconstruction and presentation of these fine details can be critical to safe and accurate image-guided radiation therapy and adaptive radiation therapy, and pivotal to achieve the maximum therapeutic ratio. These fine details may also convey constructive information towards AI-driven data analysis, such as Radiomics, to predict patient-specific outcomes and treatment responses [45, 46]. It is therefore compelling to correctly reconstruct and display these fine details in lung 4D-CBCTs. However, the 2D-3D deformation technique employed by the original SMEIR algorithm is essentially an intensity-driven approach. The intensity-driven deformable registration techniques may not perform well for these fine structure regions, the deformation of which only lead to small image intensity variations. For 2D-3D deformation, the resulting intensity changes on the 2D cone-beam projections from the deformation of these fine details are even more obscure, leading to minimal changes to the objective functions of 2D-3D deformation (Eqs. 1, 2), especially for under-sampled acquisition scenarios. To improve the motion estimation accuracy of SMEIR of these fine details, we further introduced biomechanical modeling into the SMEIR algorithm (SMEIR-Bio), to solve the DVF via a physics-driven approach [41]. Biomechanical modeling-based deformable registration has been found effective, especially at solving the deformation at low-contrast regions with minimal intensity variation signals [47–51]. By finite element analysis, biomechanical modeling drives the deformation of organs/structures of interest to meet displacement-based or force-based boundary conditions. The elastic properties of these organs/structures are incorporated into the finite element analysis, through a material model, to drive physically-plausible deformation. With minimal dependence on image intensity differences, biomechanical modeling can potentially boost the accuracy of SMEIR in
motion estimation and reconstruction of small structures.

The biomechanical modeling process has been incorporated seamlessly into the SMEIR workflow, as illustrated by Fig. 3. To perform biomechanical modeling, we segmented the lung of the mCBCT after motion-compensated reconstruction, and automatically constructed a volumetric tetrahedral mesh using the ISO2MESH and Tetgen packages [52, 53]. The lung was segmented using the automatic “snake” method, which is based on an active contour model [54]. In detail, a few “seeds” were randomly placed within the lung based on intensity-thresholding and grew to fill up the whole lung region, which automatically defined the lung boundary. On the lung boundary, we used the lung surface DVF solved in the following SMEIR motion estimation (2D-3D deformation) step as displacement-based boundary conditions for biomechanical modeling. With high intensity contrast at the lung boundary, the 2D-3D algorithm could accurately solve lung surface DVF to provide adequate boundary conditions [27]. With the tetrahedral lung mesh and the boundary conditions, the intra-lung DVF were deduced via a material model on the relationship between the strain energy and the deformation [55]. Previous studies on lung biomechanical modeling adopted multiple different models, including the linear elastic model [56], the Neo-Hookean model [55], the Ogden model [57], the Marlow model [58], and the Mooney-Rivlin model [27]. Up until now there is no consensus over the most appropriate model, as different models may yield similarly accurate results by using customized material elastic parameters. In our study, we used the Mooney-Rivlin hyper-elastic material, a model usually used to describe soft tissues with relatively large deformation. Detailed information of the model can be found in previous publications [27, 32, 55]. In this study, we modeled the lung as a homogeneous organ with the same elastic parameters throughout the whole volume. Such a strategy has proved effective and efficient by previous studies [27, 47, 55, 56, 58, 59]. After finite element analysis [60], we combined the biomechanical modeling-corrected intra-lung DVF with outside-lung DVF, and fed the resulting DVF into a new motion-compensated reconstruction step before assessing the convergence. If not converged, the mCBCT as well as the inter-phase DVF would be fed back to initiate a new round of lung segmentation, 2D-3D deformation, biomechanical modeling and motion-compensated reconstruction, until the final convergence was reached.

Fig. 3 The general flow of the biomechanical modeling-guided SMEIR scheme. 4D-CBCT: 4-Dimensional cone-beam computed tomography; DVF: Deformation-vector-field; mCBCT: Motion-compensated CBCT
The SMEIR algorithm with deep learning-driven DVF fine-tuning (SMEIR-Unet)

The success of SMEIR-Bio requires careful implementation of biomechanical modeling, which involves organ segmentation, tetrahedral mesh generation, material model/parameter assignment and finite element analysis. The additional computational workload and the complexity may reduce the efficiency of 4D-CBCT reconstruction. Recent developments of AI have found many applications in medicine, to potentially improve the accuracy and efficiency over conventional methods [61]. Multiple AI-based studies have been reported in the field of medical image registration [62–65]. Deep convolutional neural networks have been trained to learn DVFs between fixed and moving image pairs to improve deformable registration efficiency. Inspired by the potentials of deep learning, we developed a convolution neural network featuring the U-net structure (SMEIR-Unet) [66], to simplify the SMEIR-Bio workflow and accelerate the computational speed. Instead of applying biomechanical modeling to fine-tune the intra-lung DVFs, we trained a supervised network to establish a direct conversion scheme between the 2D-3D DVFs and the high-quality DVFs. The high-quality DVFs for neural network training were obtained via direct Demons registration between 4D-CT images. In detail, we simulated limited-view phase-binned cone-beam projections from the 4D-CT images and reconstructed 4D-CBCT images and the inter-phase DVFs from the projections using the original SMEIR algorithm. And we used the Demons registration algorithm to register the other phase images of 4D-CT to the reference phase image (motion-compensated phase), to derive the corresponding high-quality Demons DVFs. We paired the 2D-3D DVFs solved by the original SMEIR algorithm with the Demons DVFs, and fed them to train a population-based neural network. To focus the neural network on fine-tuning the intra-lung regions, we cropped both the 2D-3D DVFs and the Demons DVFs using the segmented lung contours at the reference phase. It can be observed that the FDK image contained structures non-existent on the “Ground-truth” image. These structures were from the under-sampling artifacts and noises. For the ART-TV images, the fine details were over-smoothed and barely recognizable. The original SMEIR algorithm was similarly susceptible to the mis-match of the fine-detail regions from those of the “Ground-truth”, indicating incorrect inter-phase DVFs. In contrast, the

Evaluation

For evaluation, we simulated cone-beam projections from high-quality lung 4D-CT images for 4D-CBCT reconstruction, and used the original 4D-CT images as the “Ground-truth” for reference. In this study, we used a dataset including 11 lung patient 4D-CTs acquired on a 16-slice Philips Brilliance CT scanner (Philips Medical Systems, Cleveland, Ohio), which were collected in an academic medical center and not publicly assessable. All the CTs were of the same slice thickness (1.5 mm), while the intra-slice pixel resolutions ranged from 0.78 mm to 0.98 mm, and were uniformly re-sampled to 1.5 mm in our study. Thus all the CTs were of resolution 1.5 mm × 1.5 mm × 1.5 mm. The CTs were of slice dimension 512 × 512, with the slice number ranging from 190 to 270. We simulated 40 projections for each 4D-CT phase, with projection angles evenly distributed across a 360° scan angle, using the Siddon’s ray-tracing technique [68]. We reconstructed 4D-CBCT images from these projections via methods including FDK, algebraic reconstruction technique with total-variation regularization (ART-TV) [69], SMEIR, SMEIR-Bio and SMEIR-Unet for comparison.

To quantitatively assess the accuracy of the motion model (DVFs) solved by different methods, ~80 lung landmarks were manually identified by expert radiation oncologists on each phase of the 4D-CT images used in our study. The 3D location changes of the same landmark between different phases were used as the “gold-standard” DVF to calculate the DVF errors of different methods. We evaluated the DVF errors along each of the three Cartesian directions (X, Y, Z), and also in terms of the error vector length.

Results

Figure 5 compares the small lung details reconstructed by different methods at the reference phase. It can be observed that the FDK image contained structures non-existent on the “Ground-truth” image. These structures were from the under-sampling artifacts and noises. For the ART-TV images, the fine details were over-smoothed and barely recognizable. The original SMEIR algorithm was similarly susceptible to the mis-match of the fine-detail regions from those of the “Ground-truth”, indicating incorrect inter-phase DVFs. In contrast, the
SMEIR-Bio algorithm reconstructed small lung details to best match the “Ground-truth”, due to the efficacy of biomechanical modeling in correcting DVFs at these fine-detail regions.

To further evaluate the accuracy of the solved inter-phase DVFs, we used the Demons algorithm to register between the reference phase and the other phases of 4D-CT to generate Demons DVFs for comparison [70]. Since the Demons algorithm was applied directly between high-quality 4D-CT images, the resulting DVFs were of high fidelity to serve as references. Figure 6 shows a comparison between the SMEIR DVF, the SMEIR-Bio DVF and the Demons DVF of three different views. It can be observed that the SMEIR DVF only displayed substantial deformation around the high contrast lung surface. Due to the lack of sufficient intensity variations inside the lung, the intra-lung DVF was not correctly derived by the original SMEIR algorithm. After incorporating biomechanical modeling, SMEIR-Bio substantially improved the DVF accuracy to match with the Demons DVF. In our study, the
SMEIR-Bio DVF appeared smoother than the Demons DVF, since we modeled the lung as a homogeneous organ. Introducing heterogeneity during the construction of the biomechanical model may help to further improve the SMEIR-Bio DVF accuracy [51, 71], to potentially better match with the Demons DVF.

As demonstrated in Fig. 7a, the SMEIR-Bio DVFs closely matched the SMEIR-Unet DVFs, and both were substantially improved compared to the original SMEIR DVFs (Fig. 6) in terms of the resemblance to the high-quality Demons DVFs. Due to the homogeneous material modeling, the SMEIR-Bio DVFs were smooth with more...
gradual spatial variations. In contrast, the SMEIR-Unet DVF s were more heterogeneous, since they were directly learned from the heterogeneous Demons DVF s and inherited similar features. However, a comparison between the reconstructed CBCT images (Fig. 7b) shows the similarity between the reconstructed CBCT images by SMEIR-Bio and SMEIR-Unet, demonstrating that both methods were capable of reconstructing accurate and high-quality CBCT images. We also computed two image quality metrics, the RMSE [39] and the UQI [33], to compare SMEIR-Bio and SMEIR-Unet to the original SMEIR algorithm, based on 10 lung region-of-interests (ROIs) focusing on fine details. For SMEIR-Bio, the average RMSE of the 10 evaluated ROIs was 0.0033 and the average UQI was 0.87. The corresponding results were 0.0035 and 0.93 for SMEIR-Unet, and 0.0039 and 0.67 for the original SMEIR method. Both the SMEIR-Bio and SMEIR-Unet methods improved the accuracy of fine detail reconstruction in lung.

Table 1 shows the quantitative DVF errors for the SMEIR, SMEIR-Bio and SMEIR-Unet methods. Without DVF fine-tuning from biomechanical modeling or convolutional neural network-based conversion, the SMEIR algorithm suffered from the largest DVF error, with the mean vector DVF error around 5 mm. The SMEIR-Bio and SMEIR-Unet techniques achieved appreciable reductions of the DVF errors, with the mean vector DVF error < 3.5 mm for both methods. Compared with

| (Mean ± SD) DVF Error (mm) | SMEIR | SMEIR-Bio | SMEIR-Unet |
|---------------------------|-------|-----------|-------------|
| X                         | 0.79 ± 0.89 | 0.57 ± 0.78 | 1.00 ± 0.98 |
| Y                         | 1.55 ± 0.77 | 0.86 ± 0.62 | 0.61 ± 0.82 |
| Z                         | 4.31 ± 1.07 | 2.53 ± 1.00 | 2.77 ± 0.99 |
| Vector                    | 4.98 ± 1.11 | 3.11 ± 1.03 | 3.30 ± 1.08 |

X, Y, and Z indicate the DVF errors along the three Cartesian directions in space, respectively. The fifth row shows the vector length of the DVF error. SMEIR: Simultaneous motion estimation and image reconstruction; CBCT: Cone-beam computed tomography; SMEIR-Bio: Biomechanical modeling-guided SMEIR; SMEIR-Unet: SMEIR with U-net based DVF fine-tuning. DVF: Deformation-vector-field.
SMEIR-Bio, which took > 5 min in building and solving the biomechanical model (including mesh generation, boundary condition assignment, finite element analysis, etc.), SMEIR-Unet only took ~ 10 s to update and fine-tune the DVF for each phase.

Discussion

4D-CBCT imaging has many potential applications and benefits in radiation therapy. One major application is in image-guided radiation therapy to allow precise cancer treatment. Studies have found that the free-breathing 3D-CBCT imaging may underestimate the motion range of tumors, which can potentially affect the tumor targeting accuracy [72]. The capture of on-board 4D tumor motion by 4D-CBCT will enable better alignment between 4D-CTs and 4D-CBCTs to achieve more accurate and robust patient setup and tumor targeting [73]. The reconstruction accuracy offered by SMEIR-type algorithms fulfills the clinical needs of accurate 4D tumor localization. The ability of SMEIR to reconstruct high-quality images from few projections also renders 4D-CBCT more time-efficient and safe. In comparison to the original SMEIR algorithm, our newly-developed SMEIR-Bio algorithm substantially improved the accuracy of intra-lung DVFs solved using limited-view projections. The derivation of a physically-plausible DVF, from displacement-based boundary conditions, allows us to capture the motion of small, intricate structures within the lung. The biomechanical modeling step requires the input of a lung contour, which was automatically-segmented using the “snake” method in this study. An investigation comparing different lung segmentation methods towards the SMEIR-Bio accuracy was not included in this article. However, since the lung is of high contrast at its boundaries, auto-segmentation is relatively easy and straightforward, and variations among different methods will be small. Furthermore, since DVFs are mostly piecewise constant and smooth, small variations between the lung boundaries segmented by different methods will bear minimal impacts on the extracted boundary condition and the corresponding biomechanical modeling results. In comparison to SMEIR-Bio, which involved the complex biomechanical modeling process, SMEIR-Unet provided similar 4D-CBCT reconstruction and DVF accuracy while with reduced computational workload and improved efficiency.

Currently, CBCT imaging is mostly applied at the beginning of each radiation therapy fraction, prior to radiation starts. Acquiring more intra-treatment 4D-CBCT images, either between consecutive radiation beams or within each beam [74], will enable continuous updates of tumor motion information to allow treatment adjustments on-the-fly. The ability to closely monitor the motion of moving targets can also promote safety margin reduction to spare more normal structures from being damaged by radiation beams [75]. Our recent developments towards SMEIR acceleration make it feasible to acquire and reconstruct the 4D-CBCT images for treatment guidance in a clinically-acceptable time frame [40], to allow multiple acquisitions and reconstructions during the treatment for continuous target monitoring. The evaluation of the SMEIR algorithm finds it could reconstruct high-quality 4D-CBCT images using as few as 20 projections per phase [38, 39]. Further imaging time reduction by acquiring fewer projections will help approach the goal of real-time volumetric, 4D imaging. Introducing prior information into the reconstruction, like the previously-acquired CT/CBCT images, or a prior motion model, may help to further lower sampling requirements for even faster imaging [25, 28]. On the other hand, high frame-rate non-ionizing imaging signals, such as the surface-guided optical imaging, have also found their potential in helping to achieve real-time volumetric imaging [76]. We envision the next evolution of real-time 4D-CBCT imaging will maneuver the potential of combining the prior information, on-board x-ray imaging, and on-board non-ionizing imaging sources. Through feeding them into a system comprised of advanced reconstruction techniques like SMEIR, biomechanical modeling and AI, a high-quality real-time volumetric image can potentially be reconstructed to maximize the accuracy of radiotherapy (Fig. 8).

In addition to providing anatomical and geometrical information towards more precise radiotherapy targeting, another potential benefit of 4D-CBCT is to calculate delivered radiation doses, and accumulate the delivered doses throughout the treatment course to dynamically assess the need of adaptive radiation therapy, and to provide dosimetric data for dose outcome analysis to fine-tune the treatment prescription [30, 31, 77]. With highly-accurate inter-phase DVFs, the 4D-CBCT solved by SMEIR-type algorithms allows direct and accurate 4D dose accumulation to determine the dose delivered to the gross tumor volumes as well as the normal tissues to evaluate the true dose coverage and the normal tissue toxicity (Fig. 9). By linking 4D-CBCTs acquired at different sessions through inter-session DVFs, we can also accumulate the overall doses delivered during the treatment course for comprehensive treatment evaluation. Currently, one major limitation of applying the CBCTs for direct dose calculation is the inaccurate HU values of the CBCT images, mostly due to the amplified scatter from the cone-beam geometry, as well as the limited field-of-view [78, 79]. Some studies proposed to circumvent the issue through direct CT-CBCT registration to calculate treatment doses on the deformed CT image instead [80, 81]. However, such methods are susceptible to the registration errors between the planning CT and
The envisioned future developments towards real-time 4D-CBCT imaging, by combining real-time on-board information provided by non-ionizing imaging signals including the optical surface imaging, and prior information (prior CT/CBCT, prior motion model), and real-time on-board x-ray imaging. Through combining all these information sources, a reconstruction system comprising SMEIR, biomechanical modeling, and AI could potentially achieve real-time imaging to guide the most accurate treatment delivery, and allow on-the-fly plan adjustment to fit the patients’ daily variations. 4D-CBCT: 4-Dimensional cone-beam computed tomography; SMEIR: Simultaneous motion estimation and image reconstruction; AI: Artificial intelligence.

Proposed SMEIR-driven 4D dose accumulation and evaluation workflow for radiation oncology. After applying the rigid registration to align the 4D-CBCT to the 4D-CT coordinates, the 4D-CBCT uses the DVFs solved by SMEIR to deform 4D doses to a single reference phase, in a similar way as the dose accumulation process of 4D-CT. The 4D-CT uses a direct deformable registration algorithm to derive inter-phase DVFs, since the high-quality CT images allow direct inter-phase DVF solution. 4D-CBCT: 4-Dimensional cone-beam computed tomography; SMEIR: Simultaneous motion estimation and image reconstruction; DVF: Deformation-vector-field.
on-board CBCT images. It would be ideal to directly improve the CBCT HU accuracy to allow direct dose calculation on CBCT. Multiple methods have been developed to reduce and correct the scatter signal, including experimental measurements, hardware-based corrections and Monte-Carlo simulations [79, 82, 83]. Introducing such methods into the clinic will allow substantial improvements of the CBCT HU accuracy. On the other hand, the issue of limited field-of-view can be potentially addressed through incorporating prior information into the CBCT reconstruction, which also enjoys the benefit of introducing more accurate HUs [30]. Recent developments of AI-based techniques were found effective to convert the CBCT images directly to improve their quality to match that of planning CT images for accurate dose calculation [84]. Such techniques could be grafted with SMEIR-type algorithms to yield the final 4D-CBCT directly applicable for accurate dose calculation. Based on the DVF-accumulated doses on each treatment fraction and throughout all fractions, and based on the DVF-propagated tumor and normal tissue contours [32, 85], decisions can be rendered in regards to the necessity of adaptive radiation therapy. Based on high-quality CBCTs, radiotherapy plans can be adapted and optimized to meet the patient’s clinical needs to maximize the radiotherapy benefits.

Conclusions
The fast development of medical imaging has allowed more precise radiation therapy treatments and is continuously pushing the boundary of technology towards safer beam delivery and more effective normal tissue sparing. 4D-CBCT plays a key role in managing sites involving moving targets such as lungs and livers. Through combining motion estimation and motion-compensated reconstruction, the SMEIR algorithm allows low-dose and faster 4D-CBCT imaging to better meet the clinical demands. The introduction of biomechanical modeling and U-net based DVF fine-tuning further develops the original SMEIR algorithm to better reconstruct small, fine details in addition to bulky structures. Through introducing more information, including the prior information and the high frame-rate, non-ionizing imaging technique, we could potentially accelerate the 4D-CBCT acquisition to allow near real-time imaging to achieve the most precise tumor targeting with a minimal safety margin. In addition to image-guidance, the 4D-CBCT and the inter-phase DVFs solved by SMEIR-type algorithms also allow dose calculation and accumulation to guide intra-course adaptive radiation therapy to tailor the plan to better deliver patient-specific cancer care, and provides abundant information to assess the dose-treatment outcomes, on both tumor control and normal tissue toxicity.

Abbreviations
4D-CBCT: 4-dimensional cone-beam computed tomography; AI: Artificial intelligence; ART-TV: Algebraic reconstruction technique with total variation-regularization; DRR: Digitally-reconstructed-radiographs; DVF: Deformation-vector-field; FDK: Feldkamp-Davis-Kress; mCBCT: Motion-compensated CBCT; NCAT: Non-uniform rational B-spline cardiac and torso; RMSE: Root-mean-squared-error; ROIs: Region-of-interests; SART: Simultaneous algebraic reconstruction technique; SMEIR: Simultaneous motion estimation and image reconstruction; SMEIR-Bio: Biomechanical modeling-guided SMEIR; SMEIR-Unet: SMEIR with U-net based DVF fine-tuning; UQI: Universal quality index

Acknowledgements
The authors would like to thank Dr. Paul Segars from the Duke University Medical Center for sharing the NCAT phantom for this study.

Authors’ contributions
YZ, XKH and JW participated in the methodology and experiment design, conducted the experiments and drafted the manuscript; all authors read and approved the final manuscript.

Funding
This work was supported in part by grants from the US National Institutes of Health, Nos. R01 EB020366 and R01 EB027898; the Cancer Prevention and Research Institute of Texas, Nos. RP1 30109 and RP1 600651, and from the University of Texas Southwestern Medical Center (Radiation Oncology Seed Grant).

Availability of data and materials
Please contact the corresponding author for data requests.

Ethics approval and consent to participate
This is a retrospective study under a departmental umbrella protocol approved by the internal review board (IRB).

Competing interests
The authors declare that they have no competing interests.

Received: 5 September 2019 Accepted: 13 November 2019
Published online: 12 December 2019

References
1. Zelefsky MJ, Kolmmeier M, Cox B, Fidaleo A, Spoering D, Pei X et al (2012) Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer. Int J Radiat Oncol Biol Phys 84(1):125–129. https://doi.org/10.1016/j. i(jqobp.2011.11.047
2. Jaffray DA, Sieversdien JH (2000) Cone-beam computed tomography with a flat-panel imager: initial performance characterization. Med Phys 27(6):1311–1323. https://doi.org/10.1118/1.1599009
3. Jaffray DA, Sieversdien JH, Wong JW, Martinez AA (2002) Flat-panel cone-beam computed tomography for image-guided radiation therapy. Int J Radiat Oncol Biol Phys 53(5):1337–1349. https://doi.org/10.1016/S0360-3016(02)02884-5
4. Lu J, Guerrero TM, Munro P, Jeung A, Chi PCM, Balter P et al (2007) Four-dimensional cone beam CT with adaptive gantry rotation and adaptive data sampling. Med Phys 34(9):3520–3529. https://doi.org/10.1118/1.2767145
5. Sweeney RA, Seubert B, Stark S, Homann V, Muller G, Flentje M et al (2002) Accuracy and inter-observer variability of 3D versus 4D cone-beam CT based image-guidance in SBRT for lung tumors. Radiat Oncol 7:81. https://doi.org/10.1118/1.1748-717X-7-81
6. Kincaid RE Jr, Yorke ED, Goodman KA, Rimner A, Wu AJ, Mageras GS (2013) Investigation of gated cone-beam CT to reduce respiratory motion blurring. Med Phys 40(4):041717. https://doi.org/10.1118/1.4795336
7. Sonke JJ, Zijp L, Remeijer P, van Herk M (2005) Respiratory correlated cone beam CT: first experimental results. Phys Med Biol 51(11):2939–2952. https://doi.org/10.1088/0031-9155/51/11/017

Grant).
51. Li M, Castillo E, Zheng XL, Luo HY, Castillo R, Wu Y et al (2013) Modeling lung deformation: a combined deformable image registration method with spatially varying Young's modulus estimates. Med Phys 40(10):081902. https://doi.org/10.1118/1.4812419

52. Fang QQ, Boas DA (2009) Tetrahedral mesh generation from volumetric binary and grayscale images. In: Abstracts of 2009 IEEE international symposium on biomedical imaging, IEEE, Boston, 28 June-1 July 2009

53. Nairn AS, Patel RV, Samani A (2011) Measurement of lung hyperechoic properties using inverse finite element approach. IEEE Trans Biomed Eng 58(10):2852-2869. https://doi.org/10.1109/TBME.2011.2166636

54. Shridari Z, Sadeghi-Naini A, Samani A (2013) Toward in vivo lung’s tissue incompressibility characterization for tumor motion modeling in radiation therapy. Med Phys 40(5):051902. https://doi.org/10.1118/1.4798461

55. Tehrani JN, Moseley J, Brock KK (2008) Contact surface and material nonlinearity of organ lungs, Phys Med Biol 53(1):305–317. https://doi.org/10.1088/0031-9155/53/1/022

56. Werner R, Ehhardt J, Schmidt R, Handels H (2010) Patient-specific finite element model of respiratory lung motion using 4D CT image data. Med Phys 36(5):1500–1511. https://doi.org/10.1118/1.3310820

57. Si H (2015) TetGen, a Delaunay-based quality tetrahedral mesh generator. ACM Trans Math Softw (TOMS) 41(2):Artic No.11. https://doi.org/10.1145/2629697

58. Yang X, Kwitt R, Niethammer M (2016) Fast predictive image registration. In: Abstracts of the 2016 38th annual international conference of the IEEE engineering in medicine and biology society, IEEE, Orlando, 16-20 August 2016. https://doi.org/10.1109/EMBC.2016.7591443

59. Al-Mayah A, Moseley J, Velec M, Hunter S, Brock K (2010) Deformable image registration for contour propagation from CT to cone-beam CT scans. Int J Radiat Oncol Biol Phys 76(1):6-12. https://doi.org/10.1016/j.ijrobp.2009.05.094

60. Li M, Castillo E, Zheng XL, Luo HY, Castillo R, Wu Y et al (2013) Modeling lung deformation: a combined deformable image registration method with spatially varying Young's modulus estimates. Med Phys 40(10):081902. https://doi.org/10.1118/1.4812419

61. de Vos BD, Berendes FF, Viergever MA, Staring M, Illigum I (2017) Nongrid image registration using multi-scale 3D convolutional neural networks. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins DL, Duchesne S (eds) Medical image computing and computer-assisted intervention. 18th international conference, October, 2015. Lecture notes in computer science (lecture notes in artificial intelligence), vol 9531. Springer, Heidelberg, pp 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

62. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization arXiv:1412.6980

63. Sokooti H, de Vos B, Berendes F, Cardeira G, Mateus D, Peter L, Bradley A, Tavares JMRS, Belagiannis V, et al (2018) Artificial intelligence in radiation oncology: a specialty-wide disruptive transformation? Radiother Oncol 129(3):421–427. https://doi.org/10.1016/j.radonc.2018.04.003

64. Yang X, Kwitt R, Niethammer M (2016) Fast predictive image registration. In: Abstracts of the 2016 38th annual international conference of the IEEE engineering in medicine and biology society, IEEE, Orlando, 16-20 August 2016. https://doi.org/10.1109/EMBC.2016.7591443

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.