Improving image-guided radiation therapy of lung cancer by reconstructing 4D-CT from a single free-breathing 3D-CT on the treatment day

Guorong Wu[a)]
BRIC and Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Jun Lian[b)]
Department of Radiation Oncology, University of North Carolina at Chapel Hill, North Carolina 27599

Dinggang Shen[b)]
BRIC and Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

(Received 29 August 2012; revised 13 October 2012; accepted for publication 27 October 2012; published 3 December 2012)

Purpose: One of the major challenges of lung cancer radiation therapy is how to reduce the margin of treatment field but also manage geometric uncertainty from respiratory motion. To this end, 4D-CT imaging has been widely used for treatment planning by providing a full range of respiratory motion for both tumor and normal structures. However, due to the considerable radiation dose and the limit of resource and time, typically only a free-breathing 3D-CT image is acquired on the treatment day for image-guided patient setup, which is often determined by the image fusion of the free-breathing treatment and planning day 3D-CT images. Since individual slices of two free breathing 3D-CTs are possibly acquired at different phases, two 3D-CTs often look different, which makes the image registration very challenging. This uncertainty of pretreatment patient setup requires a generous margin of radiation field in order to cover the tumor sufficiently during the treatment. In order to solve this problem, our main idea is to reconstruct the 4D-CT (with full range of tumor motion) from a single free-breathing 3D-CT acquired on the treatment day.
Methods: We first build a super-resolution 4D-CT model from a low-resolution 4D-CT on the planning day, with the temporal correspondences also established across respiratory phases. Next, we propose a 4D-to-3D image registration method to warp the 4D-CT model to the treatment day 3D-CT while also accommodating the new motion detected on the treatment day 3D-CT. In this way, we can more precisely localize the moving tumor on the treatment day. Specifically, since the free-breathing 3D-CT is actually the mixed-phase image where different slices are often acquired at different respiratory phases, we first determine the optimal phase for each local image patch in the free-breathing 3D-CT to obtain a sequence of partial 3D-CT images (with incomplete image data at each phase) for the treatment day. Then we reconstruct a new 4D-CT for the treatment day by registering the 4D-CT of the planning day (with complete information) to the sequence of partial 3D-CT images of the treatment day, under the guidance of the 4D-CT model built on the planning day.
Results: We first evaluated the accuracy of our 4D-CT model on a set of lung 4D-CT images with manually labeled landmarks, where the maximum error in respiratory motion estimation can be reduced from 6.08 mm by diffeomorphic Demons to 3.67 mm by our method. Next, we evaluated our proposed 4D-CT reconstruction algorithm on both simulated and real free-breathing images. The reconstructed 4D-CT using our algorithm shows clinically acceptable accuracy and could be used to guide a more accurate patient setup than the conventional method.
Conclusions: We have proposed a novel two-step method to reconstruct a new 4D-CT from a single free-breathing 3D-CT on the treatment day. Promising reconstruction results imply the possible application of this new algorithm in the image guided radiation therapy of lung cancer. [http://dx.doi.org/10.1118/1.4768226]

Key words: 4D-CT reconstruction, spatial-temporal registration, super-resolution 4D-CT model, image-guided radiation therapy, lung cancer

I. INTRODUCTION

Lung cancer is the leading lethal cancer type worldwide. For people who are diagnosed with nonsmall cell lung cancer, more than half of them will receive radiation therapy on the treatment.[1,2] Image guided radiation therapy (IGRT) is the state of the art technology to help better deliver radiation therapy to cancerous tissue, while minimize the dose to the nearby normal structures. IGRT of lung tumor is a highly challenging problem since the lung tumors move with the respiratory cycle, and it may change size, location, and shape during the course of daily treatment.[3,4] Recently, 4D-CT imaging, with

respiration phase as the fourth dimension, has been developed and becomes more and more popular since it is able to capture full range of respiration motion of lung.[5,6] It has been shown that 4D-CT imaging helps reduce artifacts from patient motion and provides more accurate delineation of tumor and other critical structures.[6–8]

For treating lung cancer with radiation therapy, generally a 4D-CT and a free-breathing 3D-CT are acquired on the planning day.[9] With these images, physician can manually delineate the outlines of tumor, lung, heart, and other critical structures from the free-breathing 3D-CT, 4D-CT or sometimes maximum intensity projection of 4D-CT.[10,11] Automatic methods to determine the lung motion envelope by using deformable registration methods have also been developed in Refs. 5, 12, and 13. With the estimated motion information, an appropriate margin of planning target volume (PTV) can be estimated to assure that the moving tumor is always covered by the radiation field. Then, the intensity of radiation field can be optimized to have PTV sufficiently treated and the dose of nearby normal structures minimized. On the treatment day, a free-breathing 3D-CT image will be acquired for image-guided patient setup. Due to CT dose and the limit of technology and time in treatment room, 4D-CT scanning normally is not performed daily before the treatment of patient. Instead, a free-breathing 3D-CT is acquired and registered with the free-breathing 3D-CT acquired on the planning day. The image registration is often rigid, mostly focusing on the alignment of bony landmarks around tumor or even tumor itself if it is visible. The result of image registration will be finally used to shift the treatment plan, so that the original treatment fields designed in the planning day can target at the right anatomy of patient, for accurate delivery of radiation therapy.

Accurate patient setup is crucial for the quality of radiation therapy. Unfortunately, the setup of lung patient has been one of the most challenging tasks in radiation therapy due to respiratory motion. The commonly used imaging technology is 2D radiograph based or 3D free-breathing based method. 3D CT has been more and more adopted because soft tissue like tumor can be imaged much clearly than 2D films. However, it is not straightforward to register two free-breathing 3D-CT images. The difficulty mainly comes from the inconsistent image information included in the two free-breathing 3D-CT images, since their corresponding slices could be acquired from the completely different respiratory phases on the planning and treatment days. Also, since the free-breathing 3D-CT is actually a mixed-phase image that never exists in the reality, it could provide misleading information to the observers. For example, as demonstrated in Refs. 14 and 15, a moving spherical object could be acquired with different shapes under the free-breathing condition, depending on which phase the scanner starts to acquire the images. 4D-CT, which can capture the motions of tumor and anatomic structures, can be a very useful tool to help the patient setup. However, in reality, 4D-CT is often not used on the treatment days due to various reasons: (1) The scheduled short treatment time for each patient does not allow the extended time for 4D-CT acquisition and reconstruction; (2) Daily 4D-CT yields significant dose on patient; (3) 4D-CT is not available in most treatment rooms. Therefore, 4D-CT is not used for the routine patient setup, for helping accurate delivery of radiation treatment.

In this paper, we propose a novel method to reconstruct a new 4D-CT from a single free-breathing 3D-CT on the treatment day, with the guidance from the super-resolution 4D-CT model initially built from the planning day 4D-CT and later adapted to estimate new motions on the treatment day. Thus, the planning day 4D-CT can be fully utilized to guide patient setup on treatment day, and finally the calculated treatment dose can be more accurately delivered to the patient. Specifically, our method consists of two steps as detailed below.

In the first step of our method (planning day), we apply a novel spatiotemporal registration algorithm[16] to simultaneously register all phase images of the 4D-CT acquired on the planning day, for building a super-resolution 4D-CT model. Specifically, all phase images are simultaneously deformed from their own domains toward the group-mean image sitting in a common space,[32] with the temporal correspondences across phases being consistent with the respiration. Here, the group mean acts as the unbiased representative image and also a common space, to which all phase images in 4D-CT will be aligned. To enhance the poor superior-inferior resolution in 4D-CT, super-resolution technique is then utilized to integrate anatomical information from different phase images for building a super-resolution group-mean image, which encompasses more anatomical details than any individual phase image. Finally, the super-resolution 4D-CT model, which consists of (1) the temporal correspondences across respiratory phases and (2) the spatiotemporal-resolution-enhanced planning day 4D-CT, is built to guide the reconstruction of a new 4D-CT on the treatment day.

In the second step of our method (treatment day), a new 4D-CT on the treatment day will be reconstructed from a free-breathing 3D-CT image by a novel 4D-to-3D image registration method. The main challenge comes from the mixed-phase image information in the free-breathing 3D-CT image which might mislead the image registration with the planning day 4D-CT. To solve this problem, we present an iterative method which consists of two steps: (1) we determine an optimal phase for each image patch in the free-breathing (mixed-phase) 3D-CT of the treatment day and thus obtain a sequence of partial 3D-CT images on the treatment day, where each new 3D-CT image has only partial image information in certain slices; and (2) we reconstruct a new 4D-CT for the treatment day by warping the 4D-CT of the planning day to the sequence of partial 3D-CT images on the treatment day, with the guidance from the 4D-CT model built on the planning day.

Each of the two above-described steps has been evaluated in our experiments. Specifically, the super-resolution 4D-CT model built in the first step has been tested on a dataset with ten lung 4D-CT images,[17] each having a set of manually identified landmarks in both selected phase images and across all phase images. The respiratory motion estimated by our algorithm achieves the best accuracy, compared to diffeomorphic Demons (Ref. 18) and SyN,[19] which are the state-of-the-art deformable registration algorithms among 14 registration

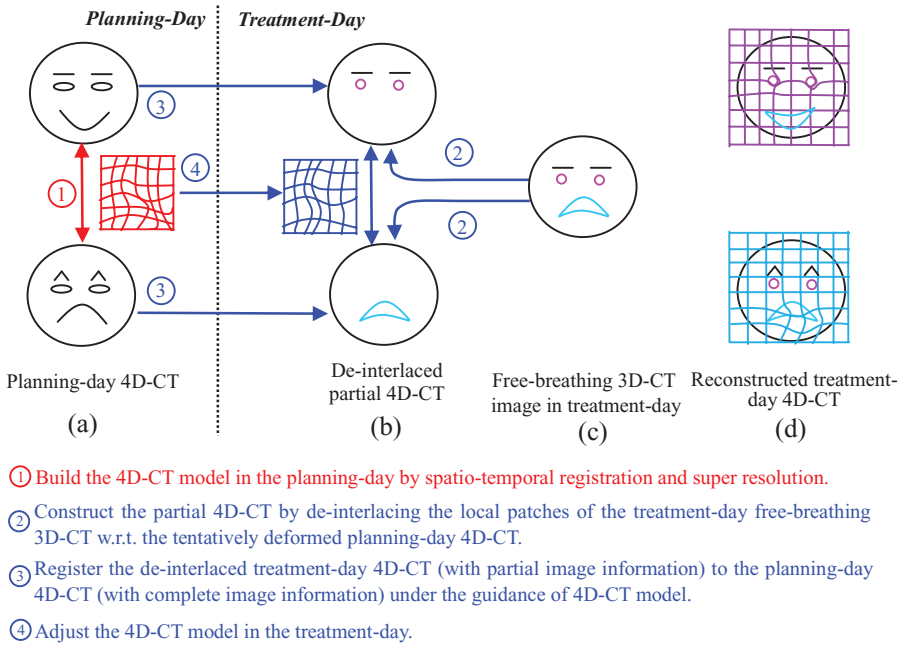


FIG. 1. The overview of our proposed method for estimating the 4D-CT images in the treatment day (d) from a free-breathing image in the treatment day (c), by using the 4D-CT model built in the planning day (a). Our method first deinterlaces the free-breathing image into the respective phase images (b), and then reconstructs a new 4D-CT in the treatment day (d) by warping the planning day 4D-CT onto the deinterlaced images, with the guidance of the 4D-CT model built in the planning day.

algorithms evaluated in Ref. 20. Our 4D-CT reconstruction algorithm for the treatment day has been evaluated on both simulated data and the real patient data. The result indicates its excellent performance in reconstructing 4D-CT from a single free-breathing 3D-CT.

II. METHODS

The goal of our method is to reconstruct a 4D-CT $\boldsymbol{T} = \{T_s | s = 1, \ldots, N\}$ with N phases for the treatment day, based on a single free-breathing 3D-CT image I taken on the treatment day, and also to build a super-resolution 4D-CT model $\boldsymbol{M}$ from the planning day 4D-CT $\boldsymbol{P} = \{P_s | s = 1, \ldots, N\}$ to guide the described 4D-CT reconstruction. Therefore, our whole method consists of two steps: (1) build a 4D-CT model $\boldsymbol{M}$ by registering all phase images of the planning day 4D-CT to the common space $\mathcal{C}$; and (2) reconstruct the 4D-CT $\boldsymbol{T}$ for the treatment day by utilizing the 4D-CT model $\boldsymbol{M}$ built on the planning day.

Figure 1 gives an overview of our method, and also explains the difficulty in registration of the free-breathing 3D-CT images. For simplicity and easy demonstration, we use face expression, instead of lung motion, as an example, as shown in Fig. 1(a). Here, we assume only two phases, i.e., smiling and angry phases, in the whole cycle of face expression change. It can be observed that the shapes of mouth and eyebrows are the keys to distinguish the smiling and angry phases. Assume that a subject's face turns from smile to anger during the image acquisition, which can be used to simulate the effect of respiration in the lung imaging when patient breathes freely. Thus, the top half face (i.e., eyes and eyebrows) can be scanned with smiling expression, while the bottom half face (i.e., mouth) can be scanned with angry expression. As the result, a mixed-expression face is obtained in Fig. 1(c). To simulate the possible shape changes during free-breathing acquisition (e.g., tumor shrinkage after treatment), we also simulate some local distortions around eyes (i.e., deformed from ellipse to circle) and mouth (i.e., deformed from close to open), as shown in pink and cyan in Fig. 1(c), respectively.

Assume we can align either smiling or angry face to the mixed-expression face [half smiling and half angry face shown in Fig. 1(c)] with perfect registration algorithm, the registered phase images will be all half smiling and half angry expression faces which may never happen in the reality. The reason of this unreasonable 4D-CT reconstruction is that the reference image is already biased by the mixed-phase information.

To address this difficulty, we first build a 4D-CT model on the planning day by a robust spatiotemporal registration algorithm, which we will explain later in Sec. II.A. After that, the temporal transformations along all phases can be obtained, as displayed by the deformation grid in Fig. 1(a). On the treatment day, we first determine the appropriate phase for each image patch of the free-breathing image by looking for the best matched patch in the corresponding phase image of the 4D-CT, as shown in Fig. 1(b). Thus, the whole free-breathing image can be deinterlaced into several 3D-CT images, denoted as $\boldsymbol{D} = \{D_s | s = 1, \ldots, N\}$, where the image information in each D_s is not complete, i.e., mouth and eyes are missing in the smile and angry phases, respectively. Then, the spatial deformation fields $\boldsymbol{H} = \{h_s(x) | x \in \Omega_{P_s \to D_s}, s = 1, \ldots, N\}$ between the 4D-CT on the planning day and the image sequence $\{D_s\}$ can be estimated in two ways: (1) For the existing image patches in D_s, their correspondences can be determined with P_s directly. (2) For other missing image patches in D_s,

TABLE I. Summary of important notations.

Symbol	Description
I	Free-breathing 3D-CT image.
$\boldsymbol{P}$	4D-CT on the planning day, $\boldsymbol{P} = \{P_s \mid s = 1, \ldots, N\}$.[a]
$\boldsymbol{T}$	The reconstructed 4D-CT on the treatment day, $\boldsymbol{T} = \{T_s \mid s = 1, \ldots, N\}$.
$\boldsymbol{F}$	The deformation field from group-mean image to each phase image P_s in constructing the 4D-CT model on the planning day.
$\boldsymbol{D}$	The partial 4D-CT after phase deinterlace, $\boldsymbol{D} = \{D_s \mid s = 1, \ldots, N\}$.
$\boldsymbol{H}$	The collection of deformation fields from each D_s to P_s, $\boldsymbol{H} = \{h_s \mid s = 1, \ldots, N\}$.
$\boldsymbol{\Phi}$	The collection of temporal deformation fields between any planning day phase images.[a]
$\boldsymbol{\Psi}$	The collection of temporal deformation fields between any treatment day phase images.
$\boldsymbol{X}^s$	The set of Q_s key points in particular partial phase image D_s, i.e., $\boldsymbol{X}^s = \{x_i^s \mid i = 1, \ldots, Q_s\}$.

[a]$\boldsymbol{P}$ and $\boldsymbol{\Phi}$ represent the roughly aligned planning day 4D-CT and the planning day temporal deformation pathway by bone alignment in Sec. II.B.

their correspondences can be compensated from other phase images according to the 4D-CT model built on the planning day. In the face example, the deinterlaced results and the final reconstructed results are shown in Figs. 1(b) and 1(d), respectively. It can be seen that the final reconstruction results are reasonable (i.e., the smiling face is still the smiling face, and the angry face is still the angry face), and also the local distortions around the mouth and eyebrows are well preserved in each phase image.

In the following, we first present the construction of super-resolution 4D-CT model $\boldsymbol{M}$ on the planning day. Then, we detail an algorithm for reconstruction of a new 4D-CT $\boldsymbol{T}$ from a free-breathing image I on the treatment day. Before we introduce the whole method in Sec. II.A, important notations used in this paper are summarized in Table I.

II.A. Construction of super-resolution 4D-CT model on the planning day

Given the 4D-CT image $\boldsymbol{P}$ on the planning day, we apply our developed spatiotemporal registration method[16] to simultaneously register all phase images P_s to the group-mean image G in the common space by estimating the deformation fields $\boldsymbol{F} = \{f_s(x) \mid x \in \Omega_{\mathcal{C}}\}$, where each f_s denotes the deformation field from group-mean image G to P_s. Specifically, key points with distinctive image features are hierarchically selected to represent the shape of group-mean image, which are used to guide the registration with respect to each phase image by robust feature matching. Since the inferior-superior resolution (usually at 2–3 mm) in 4D-CT is much lower than intraslice resolution (usually at 1×1 mm^2), motion artifacts, such as blur, overlap, and gapping,[8,21] are obvious in 4D-CT, which challenges the image registration in establishing accurate correspondences. To alleviate this issue, super-resolution technique is utilized here to construct the super-resolution group-mean image (shown in the top of Fig. 2) by integrating the complementary information of all tentatively aligned phase images. Thus, the registration between the group-mean image and each phase image become relatively easy by taking advantage of the clear anatomy in the super-resolution group-mean image. Meanwhile, by

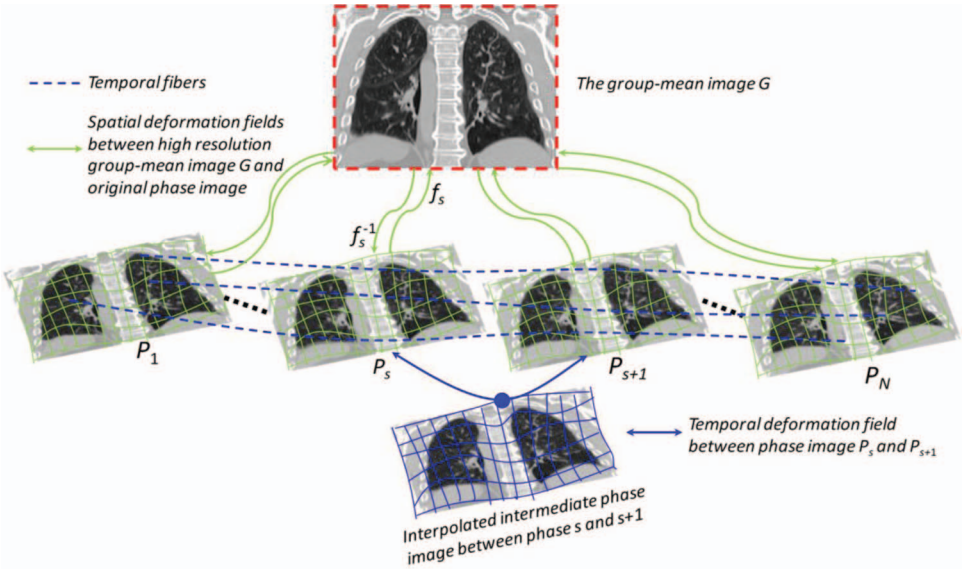


FIG. 2. The schematic illustration of our spatiotemporal registration on 4D-CT. The deformation fields $\{f_s\}$ are estimated based on the correspondences between key points by robust feature matching with respect to their shape and appearance. Meanwhile, the temporal continuity is preserved by performing kernel regression along the temporal fiber bundles $\boldsymbol{\Phi}$ (see the blue dashed curves).

mapping the group-mean image onto the domain of each phase image, every key point in the group-mean shape has several warped landmarks in different phase images, which can be assembled into a time sequence (with respect to respiratory phase) to form a virtual temporal fiber (displayed by dashed curves in Fig. 2). Therefore, the temporal coherence of 4D-CT registration can be maintained by applying kernel smoothing along all these temporal fibers. For more information, please refer to our previous paper.[16]

To make each deformation field f_s smooth and invertible, we further model f_s by an exponential model with constant velocity field to bring forth the diffeomorphism[18,22,23] which we will explain at Step 5 in Sec. II.B. Hereafter, we use f_s^{-1} to denote the inverse of deformation field f_s. Thus, the temporal correspondence between any two phases can be obtained through the domain of group-mean image. For example, the deformation pathway $\varphi_{s,\,s+1}$ (curve in Fig. 2) from phase s to the next phase $s+1$ can be calculated as $\varphi_{s,s+1} = f_s \circ f_{s+1}^{-1}$, where "$\circ$" denotes the operation of deformation composition (in this paper, we follow the definition of deformation composition in Ref. 23). In the following, we use $\mathbf{\Phi} = \{\varphi_{s,\,t} | s = 1, \ldots, N, t = 1, \ldots, N, s \neq t\}$ to denote all possible temporal deformation field between two phases s and t.

Generally, only a limited number of phase images are collected in the 4D-CT scan. For example, 4D-CT usually consists of ten phases for the complete respiration cycle or six phases from maximum inhale to maximum exhale stages. However, the temporal phase-to-phase transition might not be continuous in the original 4D-CT, especially for the fast-motion area, e.g., diaphragm. In order to remedy this motion judder, we propose to interpolate intermediate phase images between phase s and next phase $s+1$ by following the temporal deformation pathway $\varphi_{s,\,s+1}$ and deform the phase image P_s to the middle point of $\varphi_{s,\,s+1}$, as shown in the bottom of Fig. 2. After we perform the intermediate phase image interpolation within any possible neighboring phases, we can construct a temporal-resolution-enhanced 4D-CT on the planning day. To be simple, we still use $\boldsymbol{P}$ to denote the 4D-CT image sequence with N phases; however, the set of phase images has been augmented by including intermediate phase images. The advantage of phase interpolation not only lies in making the lung motion visually fluid but also brings forth more complete anatomical information of lung with respect to respiration, which is very important to clean up the mixed phases in the free-breathing 3D-CT during the reconstruction of treatment day 4D-CT. We will show this point in Sec. II.B.

Therefore, a super-resolution 4D-CT model $\boldsymbol{M} = \{\boldsymbol{P}, \mathbf{\Phi}\}$ can be built on the planning day, which consists of temporal resolution enhanced planning day 4D-CT $\boldsymbol{P}$ and the set of temporal deformation pathways between any two phases $\mathbf{\Phi}$.

II.B. Reconstruction of 4D-CT on the treatment day

The goal here is to reconstruct the 4D-CT $\boldsymbol{T}$ from a single free-breathing image I acquired on the treatment day, based on the 4D-CT model $\boldsymbol{M}$ built on the planning day. The principle behind our 4D-CT reconstruction method is that the free-breathing 3D-CT does convey the information of respiratory motion, although it is mixed and incomplete. Thus, the key point here is how to utilize the limited motion information hidden in the free-breathing 3D-CT to reconstruct the 4D-CT on the treatment day. Although it is straightforward to obtain the treatment day 4D-CT by registering each phase image P_s to the free-breathing 3D-CT I, the registration is difficult due to the mixed phase information in the free-breathing 3D-CT.

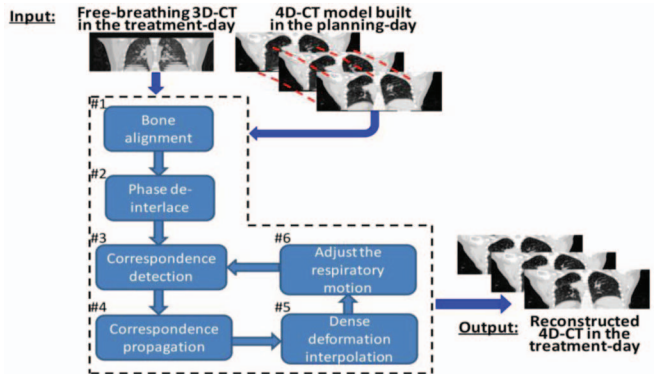


FIG. 3. The pipeline of reconstructing 4D-CT in the treatment day. First, the free-breathing 3D-CT is roughly aligned with planning day 4D-CT by bone alignment. Then, the mixed phase information in I will be deinterlaced into partial 4D-CT $\boldsymbol{D}$ according to the respiratory motion. Next, 4D-to-4D image registration is performed between the partial 4D-CT $\boldsymbol{D}$ and the complete 4D-CT $\boldsymbol{P}$ to reconstruct the treatment day 4D-CT $\boldsymbol{T}$ by iteratively establishing spatial correspondences on key points, propagating correspondences along the respiratory motion, interpolating the dense deformation field, and adjusting the treatment day respiratory motion.

The overview of our 4D-CT reconstruction method is shown in Fig. 3, which consists of six steps. First, each planning day phase image P_s is roughly registered with free-breathing 3D-CT I by the alignment of bones since the bone will almost not move with lung. Following the estimated transformations of bones, the temporal deformation pathway $\mathbf{\Phi}$ can be approximately transformed to the domain of free-breathing image I. To be simple, we still use $\boldsymbol{P}$ and $\mathbf{\Phi}$ to denote the roughly aligned planning day 4D-CT and the temporal deformation pathway, respectively. After that, we partition the free-breathing 3D-CT I into image patches by Oct-tree technique[24] and then determine the appropriate phase for each image patch by measuring the patch similarity across respiratory phases. By assigning each image patch into the corresponding phase images, we can obtain a partial 4D-CT $\boldsymbol{D}$ with limited image content in each phase image. This procedure is called as "phase deinterlace" in this paper. Next, the problem of 4D-CT reconstruction turns to the deformable registration between the partial 4D-CT $\boldsymbol{D}$ and the complete 4D-CT $\boldsymbol{P}$. Specifically, for each phase image D_s, a small number of key points will be detected to drive the estimation of entire deformation field h_s. For each key point x_i^s in D_s, robust correspondence detection will be performed with respect to P_s. Note that the temporal corresponding locations of x_i^s in all other phase images D_t ($t \neq s$) can be designated by following the respiratory motion $\mathbf{\Psi} = \{\psi_{s,t} | s = 1, \ldots, N, t = 1, \ldots, N, s \neq t\}$ on the treatment day, where each $\psi_{s,t}$ expresses the temporal correspondence between any two phase images. Thus,

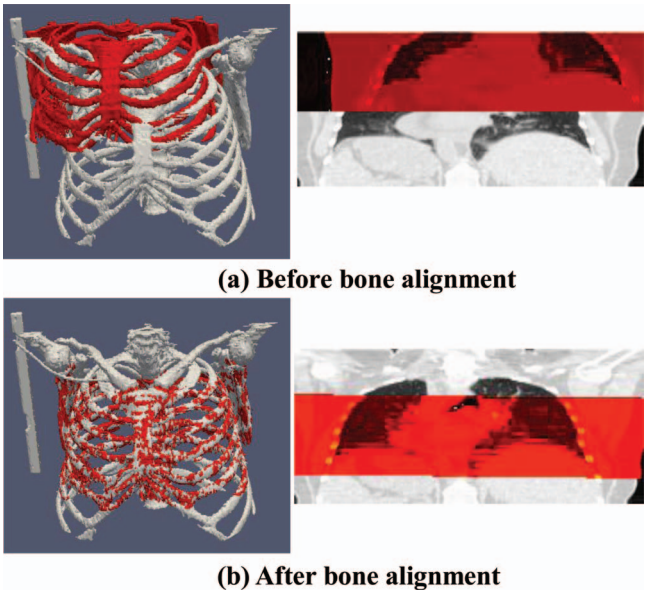


FIG. 4. The result of bone alignment. (a) shows the fusion of bone in treatment-day free-breathing 3D-CT (with smaller rib cage) over the particular phase image (with longer rib cage) before bone alignment. (b) shows the same fusion result after bone alignment. It is clear that the bones in and are approximately registered.

the established spatial correspondence on x_i^s can be propagated to other phases by requiring each corresponding location in other phase image D_t have the same spatial correspondence as x_i^s but toward P_t. In this way, the image-content-missing area in D_t will have the correspondence toward P_t consequently. Thin-plate splines (TPS) will be then used to interpolate the dense deformation h_s from the sparsely established correspondences in each phase image domain, where we further model the deformation field in the space of diffeomorphism.[18] Finally, the respiratory motion $\boldsymbol{\Psi}$ on the treatment day is adjusted according to the temporal deformation pathway $\boldsymbol{\Phi}$ on the planning day and also the tentatively estimated deformation field $\boldsymbol{H}$ to guide the next correspondence propagation step. In the end of reconstruction, each treatment day phase image T_s can be obtained by deforming the roughly aligned planning day phase image P_s according to the deformation field h_s, i.e., $T_s = h_s(P_s)$. In the following, we will explain these six steps one by one.

II.B.1. Step 1: Bone alignment

Bone is easy to extract in CT images by just keeping the image points with large intensity value. Then, we use FLIRT (the affine registration program in FSL software package developed by the Analysis Group, FMRIB, Oxford, UK) to estimate the 4×4 affine transformation matrix between free-breathing 3D-CT I and each phase image P_s of 4D-CT on the planning day. The results before and after bone alignment are shown in Figs. 4(a) and 4(b), respectively, where the bone in I and particular P_s are displayed in red and gray, respectively. Since bone will not move a lot with lung during respiration, we can roughly transform each phase image P_s of the planning day 4D-CT to the domain of the free-breathing 3D-CT I by the bone alignment. Also, we apply the affine transformation matrix to each temporal deformation field in $\boldsymbol{\Phi}$ to roughly transfer the planning day respiratory motion into the domain of the free-breathing 3D-CT.

II.B.2. Step 2: Phase deinterlace

Phase deinterlace is a very important step before image sequence registration since it cleans up the mixed phase information in the free-breathing 3D-CT I. The assumption here is that we regard lung as a moving 3D object when taking CT scans from patients. Therefore, each slice in I may contain image information from different respiratory phases. Moreover, the mixture image content could also be observed within the single slice. Fortunately, we can determine the appropriate phase for each image patch in I by looking for the similar patch across the roughly registered phase images $\boldsymbol{P}$. Since only very few phases are sampled in the original 4D-CT during entire respiration cycle ($\sim$5 s), it may be difficult for particular image patch in the free-breathing 3D-CT to find its matched patch in the existing phase images of the original 4D-CT. To alleviate this issue, as mentioned, we interpolated intermediate phase images when building the planning day 4D-CT model, in order to improve the temporal resolution and facilitate the deinterlace procedure. It is worth noting that only 2D patch is used here since the interslice thickness (>2 mm) is often larger than the intraslice resolution (<1 mm).

To clean up the misleading phase information in I, we first adaptively partition the entire image I into a number of 2D nonoverlapped image patches by Oct-tree technique.[24] Here, for each image patch, we use its intensity entropy as the complexity measurement to decide whether necessary to further divide it into smaller patches. The partition result on a free-breathing 3D-CT is shown in Fig. 5(a). Next step is to construct the partial 4D-CT $\boldsymbol{D}$ by assigning each patch into the appropriate phase image D_s, as demonstrated in Fig. 5(b). Specifically, we use $\theta_I(x, l)$ to denote the image patch in 3D-CT I with top-left position x and patch size $l \times l$, which is determined by Oct-tree partition. Then, the appropriate phase s^* for each image patch $\theta_I(x, l)$ is determined by

$$s^* = arg \min_s \delta(s), \delta(s) = \min_{y \in n_1(x)} \|\theta_I(x, l) - \theta_{P_s}(y, l)\|^2, \quad (1)$$

where $\theta_{P_s}(y, l)$ is the image patch in P_s with top-left position y. $n_1(x)$ denotes a small search neighborhood (with radium 1 mm in this paper) for finding the best match with minimal intensity difference. Finally, we move image patch $\theta_I(x, l)$ to the phase image D_{s^*} by letting $\theta_{D_{s^*}}(x, l) = \theta_I(x, l)$. Thus, each image patch will be assigned to only one particular phase image D_s. Since there is no overlap between any two image patches, the combination of existing image information in all D_ss turns to be the free-breathing 3D-CT image I, i.e., $\cup_{s=1}^N D_s = I$. As shown in Fig. 5(b), each D_s only contains limited image content after we disseminate each image patch $\theta_I(x, l)$ to the particular D_{s^*}. After obtaining $\boldsymbol{D}$, we will repeat the following four steps to reconstruct the treatment day

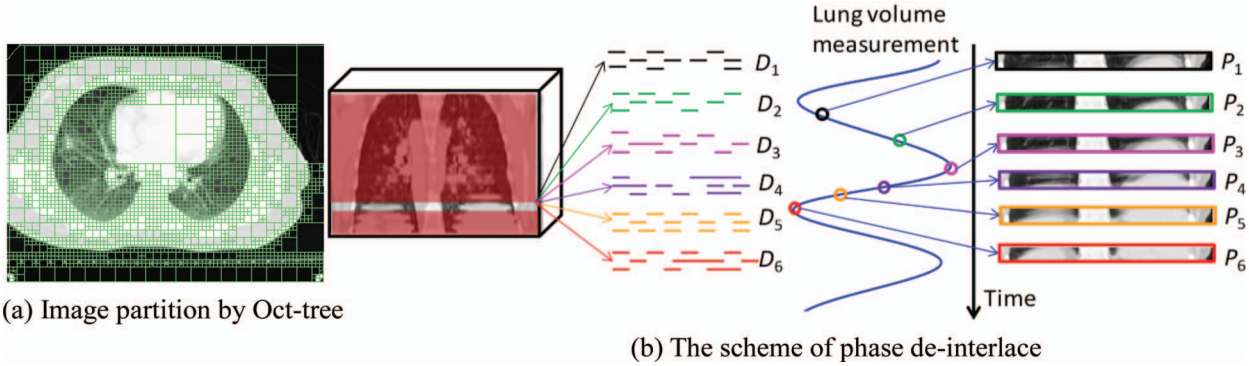


FIG. 5. The scheme of constructing partial 4D-CT $\boldsymbol{D}$ from a free-breathing 3D-CT I by phase deinterlace. First, the entire image I is portioned into a number of local 2D patches by Oct-tree technique. Then we disseminate each local patch $\theta_I(x, l)$ to the particular phase image D_{s*}, which hold the best matched local patch with respect to $\theta_I(x, l)$.

4D-CT T by gradually estimating the dense transformation field h_s between each pair of roughly aligned P_s and D_s. Since the transformation field $\boldsymbol{H}$ is iteratively estimated during 4D-CT reconstruction, we use k to denote the kth iteration in the following text.

II.B.3. Step 3: Correspondence detection by robust feature matching

The goal of this step is to establish the correspondence for any existing image patch $\theta_{D_s}(x, l)$ in the partial phase image D_s with respect to the roughly registered phase image P_s. Here, we follow the hierarchical registration framework in Ref. 16 to detect correspondence by robust feature matching.

For each image point in the partial phase image D_s and in the complete phase image P_s, we incorporate both local image appearance (i.e., image intensity) and edge information (i.e., gradients) in a 2D circular neighbor (with radius r, i.e., $r = 3$ in our experiments), into the attribute vector $\vec{a}$. Then, the normalized cross correlation is used to evaluate the similarity between point x in D_s and point y in P_s, which is denoted as $NCC(\vec{a}^s_D(x), \vec{a}^s_P(y))$. (Note that, in $\vec{a}^s_D$, we use D instead of D_s since s is already included as superscript.) In order to measure the difference between two points, we further define the feature discrepancy as $\eta(\vec{a}^s_D(x), \vec{a}^s_P(y)) = \frac{1-NCC(\vec{a}^s_D(x), \vec{a}^s_P(y))}{2}$, which ranges from 0 to 1.

We use the importance sampling strategy[25] to hierarchically select key points in each D_s. Specifically, we smooth and normalize the gradient magnitude values over the whole image domain and use the obtained values as the importance (or probability) to select key points. For example, a set of key points $\boldsymbol{X}^s = \{x^s_i | i = 1, \ldots, Q_s\}$ can be sampled via Monte Carlo simulation, with higher importance value indicating the higher likelihood of the underlying location being selected in the nonuniform sampling. Note that Q_s is the total number of key points in D_s, which is small in the beginning of reconstruction and then increases gradually until all image points are considered as key points in the end of reconstruction.

In $(k + 1)$th iteration, we will estimate the incremental deformation ϕ^{k+1}_s for each partial phase image D_s. As we will explain in Step 5, the refined deformation field h^{k+1}_s in the end of $(k + 1)$th iteration can be obtained by integrating the incremental deformation field ϕ^{k+1}_s and the latest estimated transformation field h^k_s. Exhaustive search is performed to refine the correspondence of each x^s_i with respect to the candidate point u in a search neighborhood $n_2(h^k_s(x^s_i))$, according to two criteria: (1) the feature discrepancy should be as small as possible between $\vec{a}^s_D(x)$ $(x \in \Omega_{D_s})$ and $\vec{a}^s_P(u)$ $(u \in \Omega_{P_s})$; (2) the spatial distance between candidate point u and the tentatively estimated correspondence location $\phi^{k+1}_s(x^s_i)$, i.e., $\|\phi^{k+1}_s(x^s_i) - u\|^2$, should be as close as possible.

Since there are a lot of uncertainties in correspondence matching, encouraging multiple correspondences is proven effective to alleviate the ambiguity issue.[26–28] Thus, for a particular key point x^s_i, a probability $\pi^s_{i,u}$ (called as spatial assignment) is assigned to each candidate point $u \in n_2(h^k_s(x^s_i))$ during the correspondence matching. For the sake of robustness, the candidate points even with large matching discrepancy still might have the chance to contribute to the correspondence matching in the beginning of registration, in order to encourage multiple correspondences. As the registration progresses, only the candidates with the most similar attribute vectors will be considered until the exact one-to-one correspondence is considered in the end of registration for achieving the registration specificity. This dynamic procedure can be encoded with the entropy term on the probability, i.e., $\pi^s_{i,u} \cdot \log \pi^s_{i,u}$. Here, high degree of entropy implies the fuzzy assignment while low degree means almost binary matching. We use a scalar value σ^{k+1} to act as the temperature to enforce the dynamic change on correspondence assignment. Thus, the total energy function in estimating the incremental deformation field ϕ^{k+1}_s is given as

$$E(\phi^{k+1}_s) = \sum_{i=1}^{Q_s} \sum_{u \in n_2(h^k_s(x^s_i))} \times \left\{ \pi^s_{i,u} \cdot \left[\eta\left(\vec{a}^s_D(x^s_i), \vec{a}^s_P(u)\right) + \left\| \phi^{k+1}_s(x^s_i) - u \right\|^2 \right] + \sigma^{k+1} \cdot \left[\pi^s_{i,u} \cdot \log\left(\pi^s_{i,u}\right) \right] \right\} + \beta \cdot B\left(\phi^{k+1}_s\right), \quad (2)$$

where $B(.)$ measures the bending energy[29] of incremental deformation field ϕ_s^{k+1}.

The spatial assignment $\pi_{i,u}^s$ can be calculated by letting $\partial E/\partial \pi_{i,u}^s = 0$

$$\pi_{i,u}^s = \exp\left\{-\frac{\eta\left(\bar{a}_D^s\left(x_i^s\right), \bar{a}_P^s(u)\right) + \left\|\phi_s^{k+1}\left(x_i^s\right) - u\right\|^2}{\sigma^{k+1}}\right\}. \tag{3}$$

It is clear that the spatial assignment $\pi_{i,u}^s$ is penalized in the exponential way. Notice that the temperature σ^{k+1} is the denominator of the exponential function in Eq. (3). Therefore, when σ^{k+1} is very high in the beginning of registration, even though the discrepancy η might be large or the candidate location is far away, the candidate point still might have the contribution to the correspondence detection procedure. As registration progresses, the specificity of correspondence will be encouraged by gradually decreasing the temperature σ^{k+1} to a small degree, until only the candidate point with the smallest discrepancy being selected as the correspondence in the end of registration.

After obtaining $\pi_{i,u}^s$ for each candidate u, the estimated incremental deformation on key point x_i^s can be computed by optimizing energy function E in Eq. (2) with respect to $\phi_s^{k+1}(x_i^s)$

$$\hat{\phi}_s^{k+1}\left(x_i^s\right) = \sum_{u \in n_2\left(h_s^k\left(x_i^s\right)\right)} \pi_{i,u}^s \cdot u \Big/ \sum_{u \in n_2\left(h_s^k\left(x_i^s\right)\right)} \pi_{i,u}^s. \tag{4}$$

The estimated incremental deformation field $\hat{\phi}_s^{k+1}$ is too sparse to derive the accurate deformation between D_s and P_s since there are no key points located at the image-content-missing areas after phase deinterlace. In order to compensate the deformations at the image-content-missing areas in D_s, we propose to propagate the spatial correspondence on x_i^s to all other phases by following the tentatively estimated treatment day respiratory motion.

II.B.4. Step 4: Correspondence propagation along respiratory motion

Given the previous estimated treatment day respiratory motion $\boldsymbol{\Psi}^k$, we are able to find the corresponding locations of each x_i^s in all other phases. For example, the corresponding location of x_i^s in phase t is $\psi_{s,t}^k(x_i^s)$, where $\psi_{s,t}^k$ is the temporal deformation field from phase s to phase t. Since these temporal corresponding locations have the same anatomical structure, it is reasonable to require they hold the same local deformations by

$$\hat{\phi}_t^{k+1}\left(\psi_{s,t}^k\left(x_i^s\right)\right) = \hat{\phi}_s^{k+1}\left(x_i^s\right). \tag{5}$$

The advantage of correspondence propagation is obvious in registering the partial 4D-CT with the complete 4D-CT, so that the image-content-missing area can also have the motion immediately. Taking the angry and smiling phases in Fig. 1 as an example, there is no mouth in the smiling face and no eyes in the angry face. If we only register the deinterlaced smiling face and angry face to the smiling and angry faces on the planning day, the local distortions at mouth (from close to open) and eye (from ellipse to circle) are unable to capture. However, if we know the temporal correspondence between the angry and smiling faces on the treatment day, the local deformation estimated from the existing anatomical structures in a particular phase can be propagated to other phases to compensate the motion for the same anatomical structure in other phases. As a result, the mouth in the reconstructed smiling face is open although it is impossible to detect this local change between the planning day smiling face and the deinterlaced treatment day smiling face. Similar results can also be observed at the eyes and eyebrows of the reconstructed angry face in Fig. 1.

II.B.5. Step 5: Interpolate the dense deformation field

For each D_s, it not only has the spatial correspondences on key points by robust feature matching in Step 3, but also receives the compensated displacements at the image-content-missing areas from the key points in other phase images in Step 4. Then TPS (Ref. 29) can be used to interpolate dense incremental deformation field $\hat{\phi}_s^{k+1}$ for the partial phase image D_s by considering the key points $\boldsymbol{X}^s$ and the points receiving displacements in Step 4 as control points.

In order to ensure the invertability of deformation field h_s^{k+1} between each pair of D_s and P_s, we follow an efficient nonparametric diffeomorphic approach[18] to adapt the optimization of h_s^{k+1} to the space of diffeomorphic transformation. The basic idea here is to consider the incremental deformation fields $\hat{\phi}_s^{k+1}$ in the vector space of velocity fields and then map them to the space of diffeomorphism through the exponentials, i.e., $\exp(\hat{\phi}_s^{k+1})$. Specifically, the following steps will be applied to calculate the deformation field $\hat{h}_s^{k+1}$ under the framework of diffeomorphism: (1) compute the exponential of incremental deformation field $\exp(\hat{\phi}_s^{k+1})$ by the scaling and squaring method;[18] (2) compose the exponential with the previously estimated deformation field by $\hat{h}_s^{k+1} = h_s^k \circ \exp(\hat{\phi}_s^{k+1})$; (3) the inverse deformation field can be computed by $(\hat{h}_s^{k+1})^{-1} = \exp(-\hat{\phi}_s^{k+1}) \circ (h_s^k)^{-1}$.

II.B.6. Step 6: Adjust the respiratory motion on the treatment day

Recall that we have built the 4D-CT model $\boldsymbol{M}$ on the planning day which includes the respiratory motion $\boldsymbol{\Phi}$ (temporal correspondences between any two phases) on the planning day estimated from the planning day 4D-CT $\boldsymbol{P}$ by spatiotemporal registration. Obviously, the respiratory motion $\boldsymbol{\Phi}$ on the planning day cannot be directly used on the treatment day, since the patient may breathe differently between planning and treatment days. Our solution to find the treatment day respiratory motion $\boldsymbol{\Psi}^{k+1}$ in the $(k+1)$th iteration is summarized in Fig. 6. It is apparent that the deformation fields between the roughly registered planning day phase images, the respiratory motion on the planning day, and the respiratory motion on the treatment day form a closed loop. Then, given the tentatively estimated deformation fields $\boldsymbol{H}^{k+1}$ in the $(k+1)$th iteration,

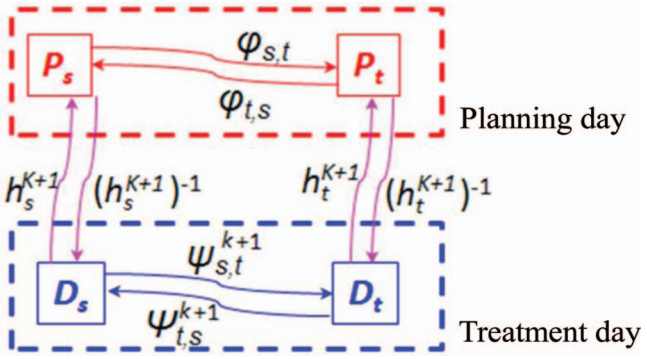


FIG. 6. The scheme of adjusting the treatment day respiratory motion between any phases s and t.

the temporal correspondence from the arbitrary phase s to another phase t can be computed as

$$\psi_{s,t}^{k+1} = \left(h_t^{k+1}\right)^{-1} \circ \varphi_{s,t} \circ h_s^{k+1}. \quad (6)$$

On the other hand, the temporal correspondence from phase t to s can be calculated by

$$\psi_{t,s}^{k+1} = \left(h_s^{k+1}\right)^{-1} \circ \varphi_{t,s} \circ h_t^{k+1}. \quad (7)$$

After we have obtained Ψ^{k+1}, we will use it in Step 4 again to propagate the spatial correspondence established in one phase to the image-content-missing areas of all other phases.

II.C. Summary of our 4D-CT reconstruction algorithm

Our 4D-CT reconstruction method can be summarized as below:

Input: Planning day 4D-CT $\boldsymbol{P}$ and the treatment day free-breathing 3D-CT I.

0. Set $k = 1$ and all deformation fields h_s^k (from deinterlace phase image D_s to phase image P_s) to identity.
1. Build the super-resolution 4D-CT model $\boldsymbol{M} = \{\boldsymbol{P}, \boldsymbol{\Phi}\}$ by spatiotemporal registration method in Ref. 16.
2. Extract the bone in I and each phase image P_s of the planning day 4D-CT.
3. Register the bone in P_s with the bone in I by FLIRT (Ref. 30) and obtain the roughly aligned planning day phase image.
4. Integrate the affine transformation matrix to each temporal deformation field $\varphi_{s,t}$ and obtain the roughly aligned planning day respiratory motion $\boldsymbol{\Phi}$ which sits in the free-breathing 3D-CT image space.
5. Use the mixed phase information in the free-breathing 3D-CT I to construct the deinterlaced 4D-CT $\boldsymbol{D}$, with each phase image only having limited image content.
6. Calculate the image attribute for each point of D_s and the roughly aligned P_s, and then select the key point $\boldsymbol{X}^s$ in each D_s.
7. Establish the correspondence for each key point x_i^s with respect to the roughly aligned phase image P_s.
   7.1. Compute the spatial assignment for each candidate point in a search neighborhood of key point x_i^s by Eq. (3);
   7.2. Compute incremental correspondence of x_i^s by Eq. (4).
8. Propagate the correspondence on each x_i^s from phase s to all other phases [Eq. (5)] by following the respiratory motion $\boldsymbol{\Psi}^k$ on the treatment day.
9. Compute the dense deformation field $\boldsymbol{H}^k$.
   9.1. Interpolate each dense incremental deformation field ϕ_s^k by TPS by considering both the key points and the points receiving the displacements from other phases in Step 8 as control points;
   9.2. Compute the exponential of ϕ_s^k and update the deformation field h_s^k by $h_s^k = h_s^k \circ \exp(\phi_s^k)$;
   9.3. Compute the inverse deformation field by $(h_s^k)^{-1} = \exp(-\phi_s^k) \circ (h_s^k)^{-1}$.
10. Adjust the respiratory motion $\boldsymbol{\Psi}^k$ on the treatment day by Eqs. (6) and (7).
11. $k \leftarrow k + 1$.
12. If not converged (e.g., k is less than the total number of iteration K), increase the number of key points Q_s and go to Step 6.
13. Obtain the phase image on the treatment day by deforming each roughly aligned planning day phase image P_s according to the estimated deformation field h_s^K, i.e., $T_s = h_s^K(P_s)$. Then, $T = \{T_s | s = 1, \ldots, N\}$ is the finally reconstructed 4D-CT on the treatment day.

III. RESULTS AND DISCUSSION

To demonstrate the performance of our proposed method, we first evaluate the accuracy of respiratory motion in building 4D-CT model on DIR-lab data,[17] with comparison to the state-of-the-art diffeomorphic Demons algorithm[18] and SyN (Ref. 19) registration method. Then, we validate our algorithm in the reconstruction of 4D-CT for the treatment day from both the simulated and real free-breathing 3D-CT data.

III.A. Evaluation of 4D-CT model on the planning day

The respiratory motion $\boldsymbol{\Phi}$ estimated on the planning day is very important to guide the reconstruction of 4D-CT on the treatment day. In this experiment, we demonstrate the accuracy of estimated respiratory motion in 4D-CT on DIR-lab dataset[17] (the abbreviation of DIR stands for deformable image registration), which has ten cases with each having a 4D-CT of six phases. The intraslice resolution is around $1 \times 1\,\text{mm}^2$, and the slice thickness is 2.5 mm. For each case, 300 corresponding landmarks in the maximum inhale (MI) and the maximum exhale (ME) phases are manually delineated by the expert. Also, correspondences of 75 landmarks are provided for each phase. Thus, the registration accuracy can be evaluated by the Euclidean distance between the expert-placed and computer-estimated landmark points. For comparison, we use diffeomorphic Demons (Ref. 18) and SyN (Ref. 19) as the comparison methods.

The registration results by diffeomorphic Demons,[33] SyN, and our method on 300 landmarks between MI and ME phases are shown in Table II. Here, the MI phase image is selected as the reference image for diffeomorphic Demons and

TABLE II. The mean and standard deviation of registration errors (mm) on 300 landmark points between maximum inhale and exhale phases.

	300 landmark points between MI and ME			
#	Initial	Diffeomorphic Demons	SyN	Our method
1	3.89 ± 2.78	2.14 ± 1.69	1.82 ± 1.43	0.64 ± 0.61
2	4.34 ± 3.90	2.10 ± 1.71	1.83 ± 1.06	0.56 ± 0.63
3	6.94 ± 4.05	2.58 ± 1.45	2.67 ± 1.25	0.70 ± 0.68
4	9.83 ± 4.85	4.81 ± 4.26	2.36 ± 1.35	0.91 ± 0.79
5	7.48 ± 5.50	1.99 ± 1.32	2.12 ± 1.37	1.10 ± 1.14
6	10.89 ± 6.97	7.94 ± 5.21	6.67 ± 5.25	3.28 ± 3.45
7	11.03 ± 7.42	6.79 ± 4.76	6.15 ± 3.34	1.68 ± 1.22
8	14.99 ± 9.01	6.30 ± 3.71	6.69 ± 3.15	1.70 ± 1.69
9	7.92 ± 3.98	4.47 ± 2.02	4.64 ± 2.14	1.72 ± 1.32
10	7.30 ± 6.35	3.67 ± 2.96	2.95 ± 2.05	1.48 ± 1.84

TABLE III. The mean and standard deviation of registration errors (mm) on 75 landmark points across all six phases.

	75 landmark points across all six phases			
#	Initial	Diffeomorphic Demons	SyN	Our method
1	2.18 ± 2.54	1.31 ± 0.99	1.29 ± 1.05	0.51 ± 0.39
2	3.78 ± 3.69	1.58 ± 1.20	1.44 ± 0.89	0.47 ± 0.34
3	5.05 ± 3.81	1.45 ± 0.85	1.56 ± 0.65	0.55 ± 0.32
4	6.69 ± 4.72	2.63 ± 2.01	2.08 ± 1.12	0.69 ± 0.49
5	5.22 ± 4.61	1.28 ± 0.85	1.24 ± 0.87	0.82 ± 0.71
6	7.42 ± 6.56	4.49 ± 3.06	3.87 ± 2.42	1.72 ± 1.83
7	6.66 ± 6.46	3.59 ± 2.56	2.65 ± 1.54	0.97 ± 0.70
8	9.82 ± 8.31	3.75 ±2.32	3.67 ± 2.05	1.70 ± 1.69
9	5.03 ± 3.79	2.45 ± 1.30	2.52 ± 1.24	1.15 ± 0.78
10	5.42 ± 5.84	2.29 ± 1.78	1.85 ± 1.51	1.06 ± 1.22

SyN. It can be observed that our method achieves the lowest mean registration errors. Table III shows the mean and standard deviation on 75 landmark points over all 6 phases by diffeomorphic Demons, SyN, and our method. Again, our method achieves the lowest registration errors. It is worth noting that the maximum registration errors among 300 landmarks are 6.08 mm by diffeomorphic Demons, 5.81 mm by SyN, and 3.67 mm by our method, respectively. These results demonstrate the accuracy of respiratory motion estimated by our method on the planning day, which will be used to guide the 4D-CT reconstruction on the treatment day.

III.B. Evaluation of 4D-CT reconstruction on the treatment day

III.B.1. Simulated dataset

Given the planning day 4D-CT [shown in Fig. 7(a), with the inhale, middle, and exhale phases displayed from left to

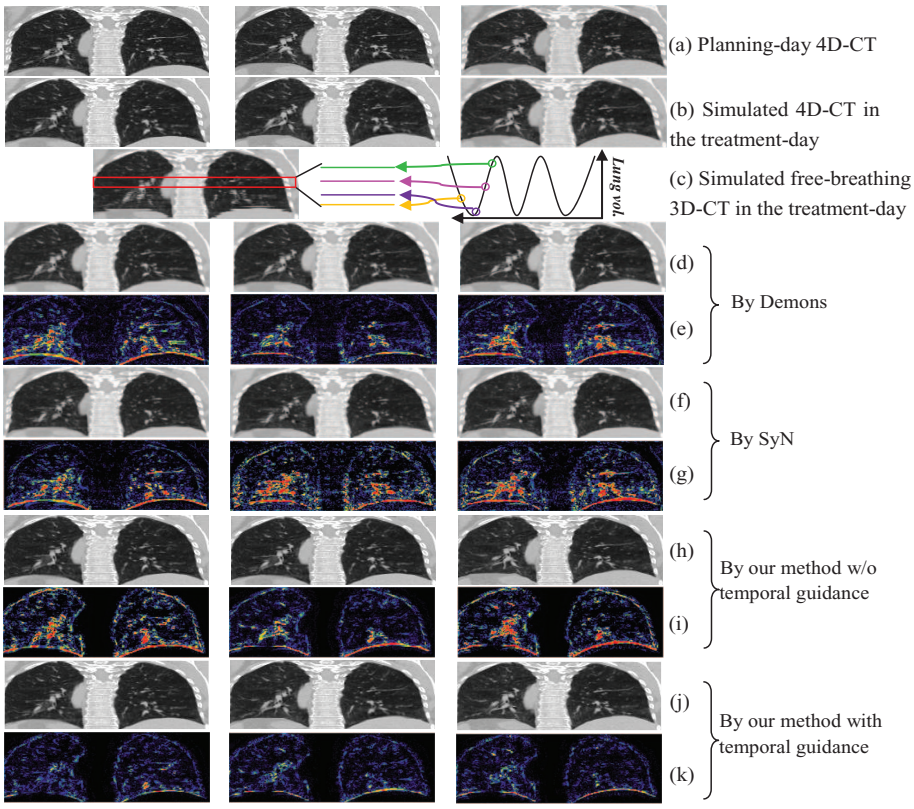


FIG. 7. The performance of 4D-CT reconstruction from a simulated free-breathing 3D-CT. A planning day 4D-CT is first obtained, with its maximum inhale, middle, and maximum exhale phases shown in (a). The simulated treatment day 4D-CT is displayed in (b), along with a simulated free-breathing 3D-CT of the treatment day shown in (c). The registration results between the planning day 4D-CT and the free-breathing 3D-CT by diffeomorphic Demons and SyN are shown in (d) and (f), with their difference images with respect to ground-truth (b) displayed in (e) and (g), respectively. The reconstruction results by our method without/with temporal guidance are displayed in (h) and (i), along with their corresponding difference images with respect to the ground-truth (b) shown in (i) and (k), respectively.

right], we first simulate the treatment day 4D-CT [shown in Fig. 7(b)] with following three steps: (1) we take the phase image at maximum inhale stage as the reference image and register all other phase images to this reference image; (2) we simulate the local distortions on the reference image by randomizing the parameters of B-spline control points, and apply this simulated deformation field (by B-spline interpolation) upon the reference image, thus obtaining the deformed reference image in the treatment day; (3) for other phases, we concatenate the deformation fields in Steps (1) and (2) to deform other phase images of planning day 4D-CT to the treatment day space. In this way, we are able to obtain the simulated 4D-CT on the treatment day, which consists of the deformed phase images in Steps (2) and (3).

Next, we need to simulate the free-breathing 3D-CT based on the simulated treatment day 4D-CT in Fig. 7(b). By mimicking the free-breathing scan procedure, we can assemble the free-breathing 3D-CT on the treatment day [Fig. 7(c)] by extracting the image slice from particular phase image of the simulated treatment day 4D-CT at the same couch table position, where the corresponding phase is determined by sequentially sampling along respiration [the demonstration of simulation in the box is shown in the right of Fig. 7(c)].

To validate the performance of our reconstruction method, we can now estimate the treatment day 4D-CT from this free-breathing 3D-CT [Fig. 7(c)] by (1) directly registering each phase image on the planning day 4D-CT with the free-breathing 3D-CT by diffeomorphic Demons, (2) directly registering each phase image on the planning day 4D-CT with the free-breathing 3D-CT by SyN, (3) reconstruction by our method without temporal guidance (Step 4: temporal correspondence propagation in Sec. II.B), and (4) reconstruction by our method with temporal guidance. The respective results are reported in Figs. 7(d), 7(f), 7(h), and 7(j), respectively, and also their difference images with respect to the ground-truth treatment day 4D-CT images [Fig. 7(b)] are given in Figs. 7(e), 7(g), 7(i), and 7(k), respectively. It can be seen from the difference images that our full method (with temporal guidance) achieves the best reconstruction result, which suggests the importance of temporal guidance from the 4D-CT model built on the planning day.

III.B.2. Real clinic data from lung cancer treatment

Ten lung cancer cases are evaluated in this experiment. Each case includes the 4D-CT on the planning day and the free-breathing 3D-CT in the followup treatment days. The CT images on the planning day and treatment days of a typical case are shown in Fig. 8. As shown in the top of Fig. 8, 4D-CT is able to provide the full range of lung motion by sorting the CT segments according to the respiration at each couch table position. Thus, it is useful for radiation oncologist to determine the range of tumor motion when designing the treatment plan. On the other hand, contouring the tumor on the free-breathing 3D-CT is much more difficult because of a lot of motion induced image artifacts. For example, we show three consecutive slices of the free-breathing 3D-CT taken in the second treatment day in the bottom of Fig. 8. It is apparent

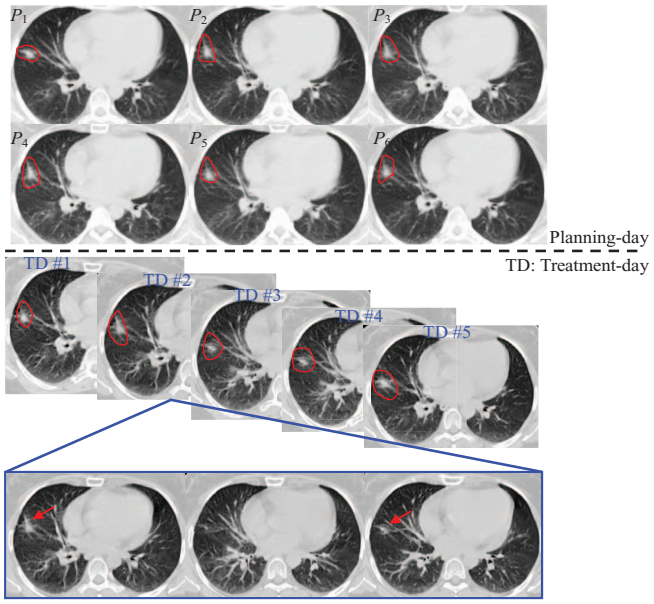
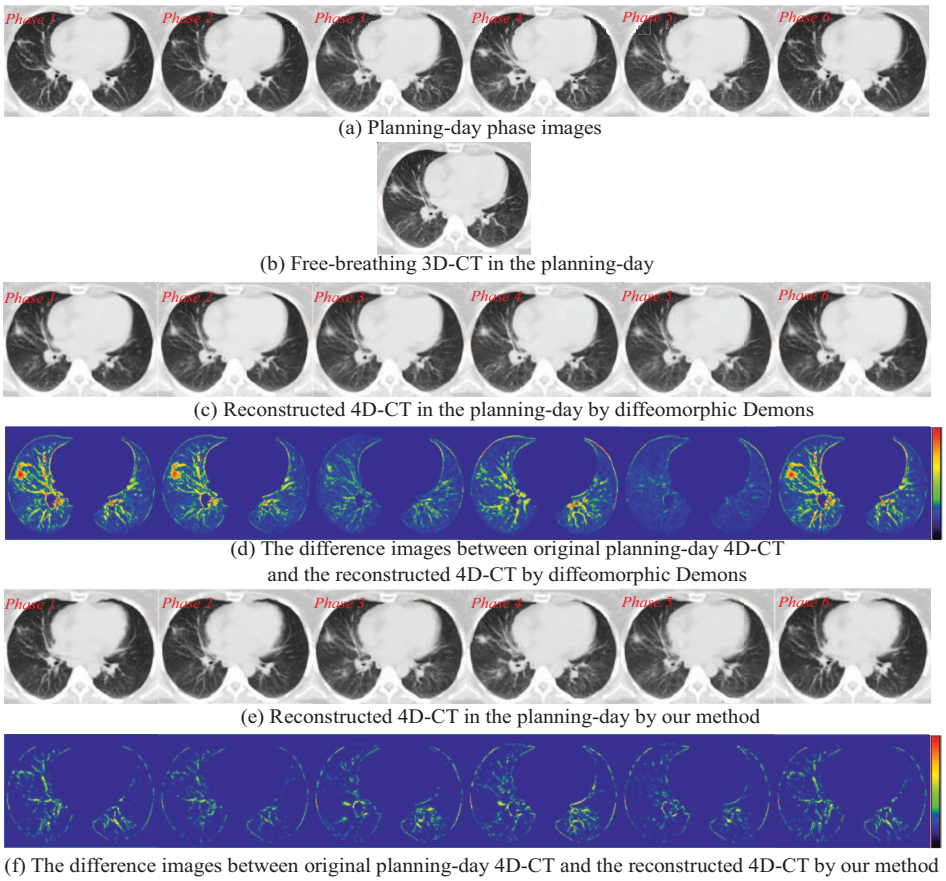


FIG. 8. The CT images scanned in a typical subject during lung cancer radiation therapy. The 4D-CT (with 6 phase images $P_1, \ldots, P_6$) scanned in the planning-day and the free-breathing 3D-CT images taken in the five treatment days are displayed in the top and middle panels of figure, respectively, with tumor also delineated by contours. Three consecutive slices of the second-treatment-day free-breathing 3D-CT are shown in the bottom for demonstrating the inconsistency for the tumor regions (indicated by arrows).

that the tumor appears at superior and inferior slices (indicated by arrows in Fig. 8) but disappears in the middle slice. Therefore, the radiation margins will have to be expanded in all directions, in order to cover the whole tumor in the radiation therapy. In the following, we will demonstrate the reconstruction of 4D-CT on the treatment day from the single free-breathing 3D-CT on the treatment day, which shows the inconsistency of tumor in the free-breathing 3D-CT.

III.B.2.a. Evaluation of 4D-CT reconstruction on the planning day. Since some patients have taken both 4D-CT and free-breathing 3D-CT on the planning day, we can evaluate the reconstruction performance of our proposed method by reconstructing 4D-CT from the free-breathing 3D-CT on the planning day. A 4D-CT and a free-breathing 3D-CT of the planning day are displayed in Figs. 9(a) and 9(b), respectively. The reconstructed 4D-CT by simply registering each phase image P_s [Fig. 9(a)] with the free-breathing 3D-CT I [Fig. 9(b)] by diffeomorphic Demons (where bone alignment is also performed first to remove the global difference) is displayed in Fig. 9(c). Obviously, all the phase images in the reconstructed 4D-CT are similar to the free-breathing 3D-CT. However, in this case, the information of respiratory motion is almost lost completely, i.e., the lung does not move from phase to phase. The reconstructed 4D-CT by our proposed method is shown in Fig. 9(e). First of all, there are obvious respiratory motions across different phases. Second, since the anatomical structures do not change for the images acquired at the same planning time, each image patch in the particular phase image of the reconstructed 4D-CT should be very similar to the corresponding one in the planning day



(a) Planning-day phase images

(b) Free-breathing 3D-CT in the planning-day

(c) Reconstructed 4D-CT in the planning-day by diffeomorphic Demons

(d) The difference images between original planning-day 4D-CT and the reconstructed 4D-CT by diffeomorphic Demons

(e) Reconstructed 4D-CT in the planning-day by our method

(f) The difference images between original planning-day 4D-CT and the reconstructed 4D-CT by our method

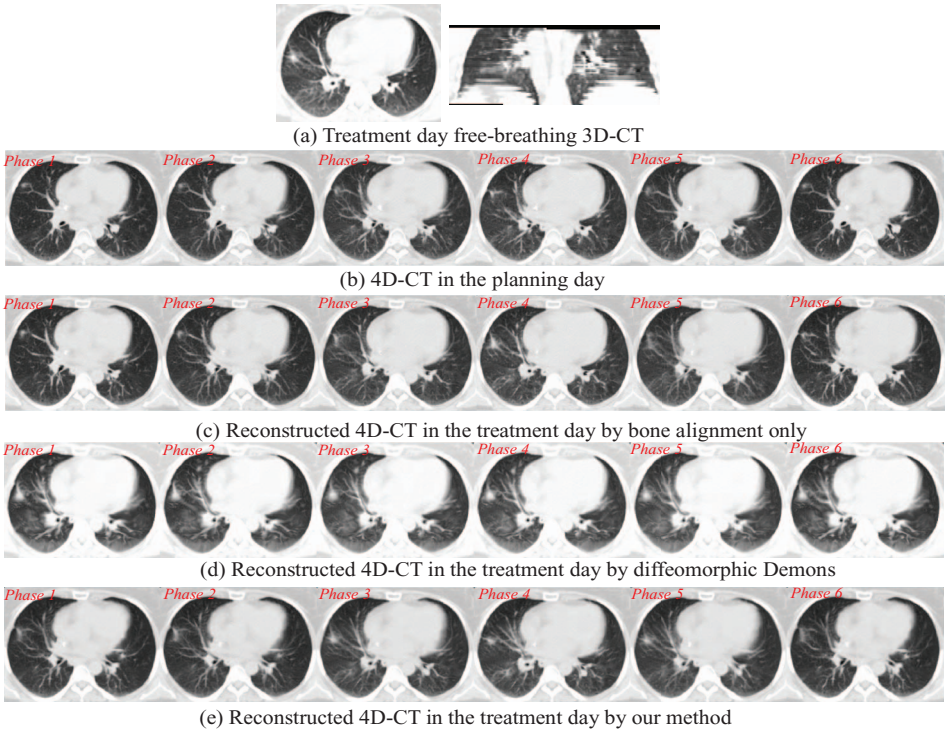
FIG. 9. Evaluation of the reconstructed 4D-CT from a single free-breathing 3D-CT in the planning day. The 4D-CT $\boldsymbol{P}$ and free-breathing 3D-CT in the planning day are shown in (a) and (b), respectively. The reconstructed 4D-CT $\boldsymbol{T}$ by diffeomorphic Demons and our method are shown in (c) and (e), with their difference images with respect to the planning day 4D-CT phase images displayed in (d) and (f), respectively.

free-breathing 3D-CT. Thus, we evaluate the image difference between the planning day 4D-CT and the reconstructed 4D-CT at every respiratory phase. The difference images by diffeomorphic Demons and our method are shown in Figs. 9(d) and 9(f), respectively. According to the color bar shown in the right, the reconstruction error by our method is much smaller than that by diffeomorphic Demons, indicating better performance of our reconstruction method.

III.B.2.b. Reconstruction of 4D-CT from a single free-breathing 3D-CT on the treatment day. We have performed our reconstruction over ten patients, each one with the 4D-CT on the planning day and the free-breathing 3D-CT on the followup treatment days. In clinical treatment, the therapists usually set up the patient by aligning the bones manually or semiautomatically. In Fig. 10(c), we show the reconstructed 4D-CT by applying bone alignment (as explained in Step 1 in Sec. II.B) between the free-breathing 3D-CT I [Fig. 10(a)] on the treatment day and each phase image of the planning day 4D-CT [Fig. 10(b)].

Although conventional deformable registration methods, e.g., diffeomorphic Demons, can be used to register each roughly aligned planning day phase image P_s (after bone alignment with respect to the free-breathing 3D-CT I), the reconstruction result is limited by the mixed phase information in the free-breathing 3D-CT. The reconstructed 4D-CT by performing diffeomorphic Demons in the conventional way is shown Fig. 10(d). Since the free-breathing 3D-CT is always used as the template image during registration, the reconstructed phase images in Fig. 10(d) are biased by the misleading image content in the free-breathing 3D-CT I. That is, the reconstructed phase images are too similar to perceive the respiratory motion in 4D-CT, which is not physically reasonable in real 4D-CT. The reconstructed 4D-CT by our method is demonstrated in Fig. 10(e). As we can observe, first, each phase image of reconstructed 4D-CT keeps the major anatomical structures as the planning day 4D-CT, which makes the reconstructed 4D-CT on the treatment day clinically reasonable. Second, the respiratory motion can be clearly observed in our reconstructed 4D-CT. We will evaluate the reconstruction accuracy of our method as follows.

The tumor on the treatment day free-breathing 3D-CT I is first manually delineated by radiologist. Similarly, we can obtain the tumor contours in all phase images of planning day 4D-CT $\boldsymbol{P}$. Then, we deform these tumor masks to the treatment day by using either the conventional bone alignment or deformation fields $\boldsymbol{H}$ estimated by our reconstruction method. The tumor movement field can finally be calculated by combining the deformed tumor masks across all phases. In this way, we obtain the tumor movement field in the reconstructed 4D-CT on the treatment day by the bone alignment and



(a) Treatment day free-breathing 3D-CT

(b) 4D-CT in the planning day

(c) Reconstructed 4D-CT in the treatment day by bone alignment only

(d) Reconstructed 4D-CT in the treatment day by diffeomorphic Demons

(e) Reconstructed 4D-CT in the treatment day by our method

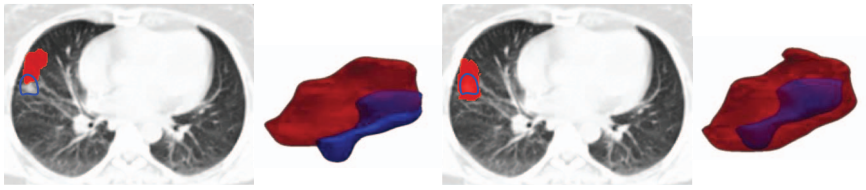
FIG. 10. Reconstructed 4D-CT from a single free-breathing 3D-CT in the treatment day. The free-breathing 3D-CT in the treatment day is shown in (a). The reconstructed 4D-CTs from (a) by bone alignment, diffeomorphic Demons, and our reconstruction method are displayed in (b), (c), and (d), respectively.

our reconstruction method, as shown with larger surface in Figs. 11(a) and 11(b), respectively.

Since the tumor usually could move with lung freely, it is very difficult to align the tumor seen in the planning CT with the tumor seen in the free-breathing treatment day 3D-CT, if aligned only with bones. As shown in Fig. 11(a), the estimated tumor movement field is unable to cover the entire tumor in the free-breathing 3D-CT on the treatment day due to the insufficient alignment. To address this issue, the treatment margin has to be expanded on the treatment-plan design stage in all directions.

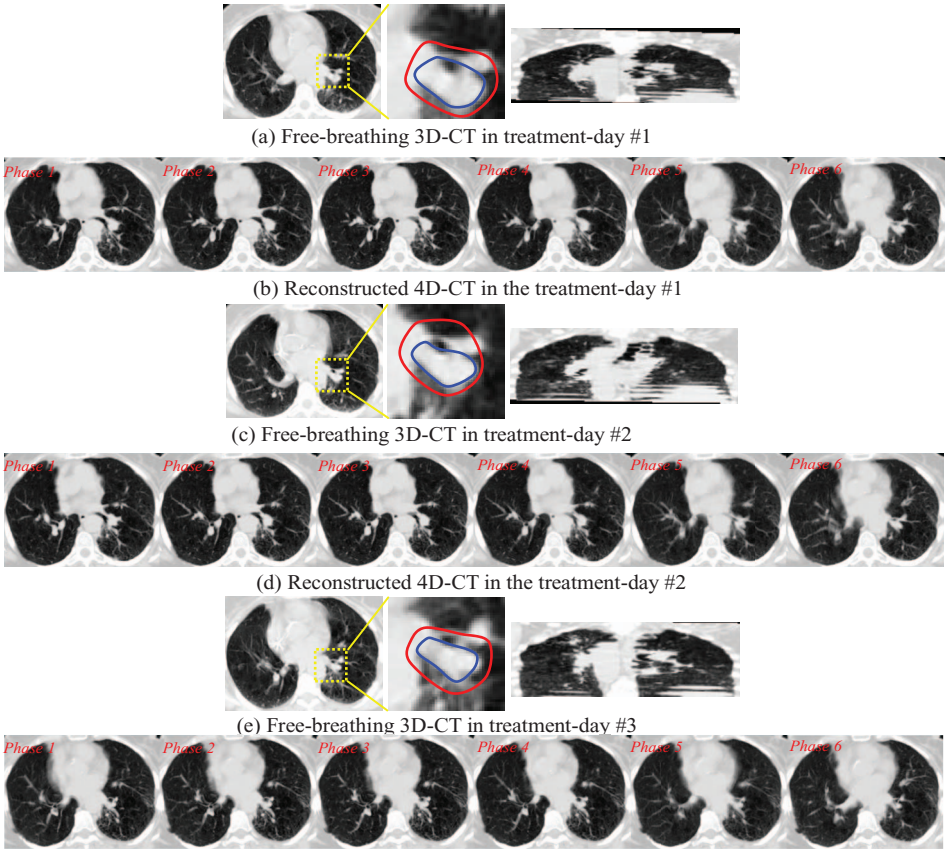
On the contrary, by using our reconstruction method, the tumor extracted in the free-breathing 3D-CT is always within the estimated tumor movement field [Fig. 11(b)]. With better estimation of 4D lung/tumor motion, we will be able to deliver the radiation dose more accurately onto the tumor. Moreover, it can allow the reduction of the margin of radiation treatment field. This makes the dose escalation of tumor possible which often means better tumor control.[3,11] Meanwhile, the critical structures around the tumor can be better spared as the result of tighter field margin. Consequently, we can make the margin of radiation treatment filed smaller than the conventional patient setup methods, which implies the possible dose escalation of tumor and fewer side effects to the normal structures.

The reconstructed 4D-CT by our method on another patient in three treatment days are also shown in Fig. 12, with the free-breathing 3D-CT in three treatment days shown in (a), (c), and (e), respectively. The reconstructed 4D-CTs are displayed in (b), (d), and (f) for each treatment day. To further demonstrate the performance of our reconstruction method with respect to the tumor motion, we manually extract the contour of tumors on each treatment day free-breathing 3D-CT and each phase image of the planning day 4D-CT. Then, the tumor movement field in each reconstructed treatment day 4D-CT can be obtained by integrating the deformed planning day tumor masks across all respiratory phases. The contours of tumor on the treatment day 3D-CT and the estimated tumor movement fields on the reconstructed 4D-CT are displayed in



(a) By bone alignment only. (b) By our reconstruction method.

FIG. 11. The overlap of tumor contour drawn on the free-breathing 3D-CT and the tumor movement field in the reconstructed treatment-day 4D-CT. The shape of tumor in the treatment-day free-breathing 3D-CT and the tumor movement field in the reconstructed treatment-day 4D-CT is displayed by small surface and large surface, respectively.



(a) Free-breathing 3D-CT in treatment-day #1

(b) Reconstructed 4D-CT in the treatment-day #1

(c) Free-breathing 3D-CT in treatment-day #2

(d) Reconstructed 4D-CT in the treatment-day #2

(e) Free-breathing 3D-CT in treatment-day #3

(f) Reconstructed 4D-CT in the treatment-day #3

FIG. 12. Reconstructed 4D-CT from the single free-breathing 3D-CT in three treatment days. The tumors in the treatment-day free-breathing 3D-CT image are manually delineated by the inner contour. The estimated tumor movement fields in the treatment days are outlined by the outer contour.

inner and outer contours in the middle of Figs. 12(a), 12(c), and 12(e), respectively. The anatomy inside yellow box is enlarged. It can be seen that the estimated tumor movement fields can always cover the entire tumor, which provides very useful tumor motion information for radiation therapy on the treatment day.

III.C. Discussion

Although in the conventional method the bones between the free-breathing 3D-CT image of the planning day and the free-breathing 3D-CT images of the treatment days can be aligned automatically and also the soft tissues can be aligned manually, no tumor motion information on the treatment day is provided. Our method can automatically map the tumor movement from the planning day to the treatment days by also accommodating the new motion detected from the treatment day 3D-CT, and can thus precisely localize the moving tumor just prior to radiotherapy. In summary, our method will bring immediate improvement for lung tumor treatment in two aspects: (1) the reconstructed 4D-CT image will provide a complete description about lung/tumor motions on the treatment day, and thus the radiation fields can be more accurately aligned onto tumor for better treatment. (2) Our method is free of internal fiducial markers which could cause pneumothorax for many patients. With our method, the setup uncertainty used in the conventional PTV design can be potentially reduced.

Furthermore, we use a typical patient (given in Fig. 12) as an example to show that the margin in patient setup can be reduced by our method. In the planning day, it is straightforward to estimate tumor motion from the 4D-CT of the patient, which is about 5 mm. Based on this measurement, we usually expand clinical target volume (CTV) for 5 mm to form the internal target volume (ITV), in order to cover the whole tumor motion. In Fig. 12, each inner contour of the tumor outlines the CTV on the free-breathing 3D-CT of each treatment day. In the conventional patient setup based on only the free-breathing 3D-CT, we have to empirically expand ITV for another 7 mm in all directions to form the PTV for radiation treatment, in order to account for the setup uncertainty since it is difficult to obtain accurate tumor motion from the treatment day free-breathing 3D-CT. Thus, the overall dose margin is 12 mm (5 + 7 mm) for this particular patient. On the other hand, our method is able to provide the tumor movement field (i.e., the outer contours in Fig. 12) for each treatment day, which can be obtained by mapping the motion information encoded in the planning day 4D-CT to the treatment day by further considering the hidden motion information identified from the treatment day free-breathing 3D-CT. It is worth noting that the mean distance and standard deviation between the CTV and our estimated tumor movement

field in the treatment day is 3.73 ± 1.32 mm, which is close to the 5 mm motion estimated in the planning day by considering also the standard deviation here. Recall that, using our method, the maximum registration error between two extreme phases is 3.67 mm upon 300 landmarks of ten 4D-CT cases. Thus, instead of using the empirical 7 mm expansion from ITV to PTV, we could use 3.67 mm only to account for the setup uncertainty. Then, the overall dose margin can be reduced to 5 + 3.67 mm = 8.67 mm [here, we still use 5 mm expansion from CTV to ITV, instead of 3.73 mm estimated by our method, due to the consideration of standard deviation (1.32 mm)], which implies 27.8% of margin reduction for the patient. It is worth noting that this is the result for one patient. To make more general conclusion on the amount of margin reduction by our method, we will test it on more patient data in our future work.

Most IGRT methods assume that the respiratory motion difference between the planning day and the treatment days is small, so that the respiratory motion obtained from the planning day 4D-CT can be directly used for the treatment dose management.[31] In our method, we estimate the patient-specific respiratory motion on the treatment day, since our method is able to extract the motions hidden in the free-breathing 3D-CT. Note that the free-breathing 3D-CT does convey motion information, although it is mixed and incomplete. We have applied our method to ten cases and obtained significant improvement in patient setup over the conventional methods that generally use only the free-breathing 3D-CT, without guidance from the 4D-CT model built on the planning day. In our future work, we will further evaluate the performance of our method with respect to various factors, e.g., tumor size, tumor location, and patient breathing pattern.

IV. CONCLUSIONS

We have proposed a novel two-step method to reconstruct a new 4D-CT from a single free-breathing 3D-CT on the treatment day, for possible improvement of lung cancer image-guided radiation therapy, which is currently often done by registration of two free-breathing 3D-CT images for patient setup. In our method, the 4D-CT model is first built on the planning day by a robust spatiotemporal registration method. Then, with the guidance of this 4D-CT model and the inclusion of newly detected motion from the free-breathing 3D-CT on the treatment day, a 4D-CT on the treatment day can be reconstructed for more accurate patient setup. We have extensively evaluated each step of our method with both real and simulated datasets, and obtained very promising results in building 4D-CT model and reconstructing new 4D-CT on the treatment day.

ACKNOWLEDGMENTS

This work was supported in part by NIH grant CA140413, by National Science Foundation of China under Grant No. 61075010, and also by the National Basic Research Program of China (973 Program) Grant No. 2010CB732505.

[a]Electronic mail: grwu@med.unc.edu
[b]Authors to whom correspondence should be addressed. Electronic mail: dgshen@med.unc.edu and jun_lian@med.unc.edu

[1]H. Wagner, "Radiation therapy in the management of limited small cell lung cancer: When, where, and how much?," Chest **113**, 92S–100S (1998).
[2]C. Le Pechoux, "Role of postoperative radiotherapy in resected non-small cell lung cancer: A reassessment based on new data," Oncologist **16**, 672–681 (2011).
[3]P. Keall, G. Mageras, J. Balter, R. Emery, K. Forster, S. Jiang, J. Kapatoes, D. Low, M. Murphy, B. Murray, C. Ramsey, M. Van Herk, S. Vedam, J. Wong, and E. Yorke, "The management of respiratory motion in radiation oncology report of AAPM Task Group 76," Med. Phys. **33**, 3874–3900 (2006).
[4]E. Rietzel, G. T. Y. Chen, N. C. Choi, and C. G. Willet, "Four-dimensional image-based treatment planning: Target volume segmentation and dose calculation in the presence of respiratory motion," Int. J. Radiat. Oncol., Biol., Phys. **61**, 1535–1550 (2005).
[5]T. Zhang, N. P. Orton, and W. A. Tomé, "On the automated definition of mobile target volumes from 4D-CT images for stereotactic body radiotherapy," Med. Phys. **32**, 3493–3503 (2005).
[6]X. Zhang, K.-L. Zhao, T. M. Guerrero, S. E. McGuire, B. Yaremko, R. Komaki, J. D. Cox, Z. Hui, Y. Li, W. D. Newhauser, R. Mohan, and Z. Liao, "4D CT-based treatment planning for intensity-modulated radiation therapy and proton therapy for distal esophagus cancer," Int. J. Radiat. Oncol., Biol., Phys. **72**, 278–287 (2006).
[7]E. Rietzel, T. Pan, and G. T. Y. Chen, "Four-dimensional computed tomography: Image formation and clinical protocol," Med. Phys. **32**, 874–889 (2005).
[8]E. Johnston, M. Diehn, J. D. Murphy, J. Billy, W. Loo, and P. G. Maxim, "Reducing 4D CT artifacts using optimized sorting based on anatomic similarity," Med. Phys. **38**, 2424–2429 (2011).
[9]L. Xing, B. Thorndyke, E. Schreibmann, Y. Yang, T.-F. Li, G.-Y. Kim, G. Luxton, and A. Koong, "Overview of image-guided radiation therapy," Med. Dosim. **31**, 91–112 (2006).
[10]S. Gaede, J. Olsthoorn, A. V. Louie, D. Palma, E. Yu, B. Yaremko, B. Ahmad, J. Chen, K. Bzdusek, and G. Rodrigues, "An evaluation of an automated 4D-CT contour propagation tool to define an internal gross tumour volume for lung cancer radiotherapy," Radiother. Oncol. **101**, 322—328 (2011).
[11]F. Khan, G. Bell, J. Antony, M. Palmer, P. Balter, K. Bucci, and M. J. Chapman, "The use of 4DCT to reduce lung dose: A dosimetric analysis," Med. Dosim. **34**, 273–278 (2009).
[12]T. Guerrero, G. Zhang, W. Segars, T.-C. Huang, S. Bilton, G. Ibbott, L. Dong, K. Forster, and K. P. Lin, "Elastic image mapping for 4-D dose estimation in thoracic radiotherapy," Radiat. Prot. Dosim. **115**, 497–502 (2005).
[13]A. Trofimov, E. Rietzel, H.-M. Lu, B. Martin, S. Jiang, G. T. Y. Chen, and T. Bortfeld, "Temporo-spatial IMRT optimization: Concepts, implementation and initial results," Phys. Med. Biol. **50**, 2779–2798 (2005).
[14]G. T. Y. Chen, J. H. Kung, and K. P. Beaudette, "Artifacts in computed tomography scanning of moving objects," Semin. Radiat. Oncol. **14**, 19–26 (2004).
[15]S. S. Vedam, P. J. Keall, V. R. Kini, H. Mostafavi, H. P. Shukla, and R. Mohan, "Acquiring a four-dimensional computed tomography dataset using an external respiratory signal," Phys. Med. Biol. **48**, 45–62 (2003).
[16]G. Wu, Q. Wang, J. Lian, and D. Shen, "Estimating the 4D respiratory lung motion by spatiotemporal registration and building super-resolution image," in *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Toronto, 2011.
[17]R. Castillo, E. Castillo, R. Guerra, V. E. Johnson, T. McPhail, A. K. Garg, and T. Guerrero, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," Phys. Med. Biol. **54**, 1849–1870 (2009).
[18]T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," Neuroimage **45**, S61–S72 (2009).
[19]B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," Med. Image Anal. **12**, 26–41 (2008).
[20]A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren,

R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," Neuroimage **46**, 786–802 (2009).
[21]T. Yamamoto, U. Langner, B. W. Loo, Jr., J. Shen, and P. J. Keall, "Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients," Int. J. Radiat. Oncol., Biol., Phys. **72**, 1250–1258 (2008).
[22]T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Symmetric log-domain diffeomorphic registration: A demons-based approach," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, (Springer, NY, 2008), pp. 754–761.
[23]J. Ashburner, "A fast diffeomorphic image registration algorithm," Neuroimage **38**, 95–113 (2007).
[24]M. Donald, "Geometric modeling using octree encoding," Comput. Graph. Image Process. **19**, 129–147 (1982).
[25]Q. Wang, L. Chen, P. T. Yap, G. Wu, and D. Shen, "Groupwise registration based on hierarchical image clustering and atlas synthesis," Hum. Brain Mapp. **31**, 1128–1140 (2010).
[26]H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," Comput. Vis. Image Underst. **89**, 114–141 (2003).
[27]G. Wu, P.-T. Yap, M. Kim, and D. Shen, "TPS-HAMMER: Improving HAMMER registration algorithm by soft correspondence matching and thin-plate splines based deformation interpolation," Neuroimage **49**, 2225–2233 (2010).
[28]G. Wu, Q. Wang, H. Jia, and D. Shen, "Feature-based groupwise registration by hierarchical anatomical correspondence detection," Hum. Brain Mapp. **33**, 253–271 (2012).
[29]F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," IEEE Trans. Pattern Anal. Mach. Intell. **11**, 567–585 (1989).
[30]M. Jenkinson and S. M. Smith, "A global optimisation method for robust affine registration of brain images," Med. Image Anal. **5**, 143–156 (2001).
[31]Y. Chi, J. Liang, X. Qin, and D. Yan, "Respiratory motion sampling in 4DCT reconstruction for radiotherapy," Med. Phys. **39**, 1696–1704 (2012).
[32]Common space is a specific term in image processing. Group-mean image in the common space stands for a reference image to which all the subject images deform. Group-mean image is not any individual image, instead, it can be considered as a hidden image which will also be estimated during groupwise image registration.
[33]The program of diffeomorphic Demons is downloaded from www.insight-journal.org/browse/publications/154. In all of following experiments, we use the same parameters (-i 10 × 8 × 5 -s 2.2 -e).