IPEM Institute of Physics and Engineering in Medicine

ACCEPTED MANUSCRIPT • OPEN ACCESS

Consistent and invertible deformation vector fields for a breathing anthropomorphic phantom: a post-processing framework for the XCAT phantom

View the article online for updates and enhancements.

Consistent and invertible deformation vector fields for a breathing anthropomorphic phantom: a post-processing framework for the XCAT phantom

Björn Eiben[*,1], Jenny Bertholet[*,2], Martin J. Menten[2,3],
Simeon Nill[2], Uwe Oelfke[2], Jamie R. McClelland[1]

[1]Centre for Medical Image Computing, Radiotherapy Image Computing Group, Department of Medical Physics and Biomedical Engineering, University College London, United Kingdom
[2]Joint Department of Physics, Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, United Kingdom
[3]Biomedical Image Analysis Group, Department of Computing, Imperial College London, United Kingdom
[*]Authors contributed equally

email: b.eiben@ucl.ac.uk (BE) and jenny.bertholet@icr.ac.uk (JB)

Abstract

Breathing motion is challenging for radiotherapy planning and delivery. This requires advanced four-dimensional (4D) imaging and motion mitigation strategies and associated validation tools with known deformations. Numerical phantoms such as the XCAT provide reproducible and realistic data for simulation-based validation. However, the XCAT generates partially inconsistent and non-invertible deformations where tumours remain rigid and structures can move through each other. We address these limitations by post-processing the XCAT deformation vector fields (DVF) to generate a breathing phantom with realistic motion and quantifiable deformation.

An open-source post-processing framework was developed that corrects and inverts the XCAT-DVFs while preserving sliding motion between organs. Those post-processed DVFs are used to warp the first XCAT-generated image to consecutive time points providing a 4D phantom with a tumour that moves consistently with the anatomy, the ability to scale lung density as well as consistent and invertible DVFs. For a regularly breathing case, the inverse consistency of the DVFs was verified and the tumour motion was compared to the original XCAT. The generated phantom and DVFs were used to validate a motion-including dose reconstruction (MIDR) method using isocenter shifts to emulate rigid motion. Differences between the reconstructed doses with and without lung density scaling were evaluated.

The post-processing framework produced DVFs with a maximum 95th-percentile inverse-consistency error of 0.02 mm. The generated phantom preserved the dominant sliding motion between the chest wall and inner organs. The tumour of the original XCAT phantom preserved its trajectory while deforming consistently with the underlying tissue. The MIDR was compared to the ground truth dose reconstruction illustrating its limitations. MIDR with and without lung density scaling resulted in

i

small dose differences up to 1 Gy (prescription 54 Gy).
The proposed open-source post-processing framework overcomes important limitations
of the original XCAT phantom and makes it applicable to a wider range of validation
applications within radiotherapy.

ii

I.   Introduction

Respiratory motion is a challenge for accurate radiation therapy (RT) planning and delivery (Keall et al., 2006). Pre-treatment 4D-computed tomography (CT) imaging is commonly used to calculate margins for passive motion mitigation (van Herk, 2004; Guckenberger et al., 2009) while active motion mitigation (e.g. gating, tracking) (Keall et al., 2006) is increasingly being used but requires real-time motion monitoring (Bertholet et al., 2019b). In addition, motion including dose reconstruction (MIDR) can estimate the actually delivered dose (Poulsen et al., 2012; Kamerling et al., 2016); but it is currently limited by the simplicity of motion models and the accuracy of deformable image registration (DIR) for dose accumulation (Brock et al., 2017; Paganelli et al., 2018b). There is therefore a need for adequate validation tools for 4D imaging, motion monitoring, DIR or 4D dose reconstruction and accumulation methods. In particular, the lack of ground truth deformation vector fields (DVF) for 4D anthropomorphic data is a major limitation for the validation of DIR-based methods (Chetty and Rosu-Bubulac, 2019).

Commercial and custom-built physical 4D phantoms with varying levels of anthropomorphic details can be used for end-to-end tests (Bertholet et al., 2019b). However, dosimetric measurements in physical phantoms commonly lack volumetric information (Ehrbar et al., 2016) and there is always a trade-off between the anthropomorphism, complexity of motion and reproducibility of motion (Bertholet et al., 2019b). In particular, there is a need to be able to simulate patient-specific inter- and intra-cycle variations of respiratory motion which are known challenges for sorting-based 4D imaging techniques, DIR or motion modelling (Chetty and Rosu-Bubulac, 2019; McClelland et al., 2013).

Digital 4D phantoms (Segars et al., 2010; Williams et al., 2013; Mishra et al., 2012; Paganelli et al., 2017) provide reproducible data for simulation-based validation. The extended cardiac torso (XCAT) phantom proposed by Segars et al. (2010) is based on the visible male and female anatomical datasets from the National Library of Medicine and provides a time-series with cardiac and/or respiratory motion. Organ surfaces are defined by nonuniform rational B-splines (NURBS) whose motion is obtained by displacing the NURBS' control points. Respiratory motion is user-defined via chest and diaphragm displacement traces. This allows to simulate intra- and inter-cycle variation, for instance by using displacement traces measured from patients or volunteers. The outputs of the XCAT include a time-series

of attenuation volumes (convertible to Hounsfield Unit (HU)) and the DVFs from the first ($i = 0$) to each following frame ($i = [1, \ldots, N]$). Rigid tumours can be generated in a separate output file and set to move such that their centre-of-mass follows the motion of any location in the body. The tumours are then added onto the phantom attenuation volumes by overriding the grey value at their set location.

Owing to its realism and flexibility the XCAT phantom has been used for a wide range of applications (see e.g. Segars et al. (2018) and references therein). However, there are also some important limitations.

The NURBS approach allows to handle the sliding of the lungs along the chest wall in principle. Nevertheless, various organ boundaries can be observed to move through each other in the attenuation volumes, possibly due in part to the voxelization of the output. Therefore, the sliding motion at the lung/chest wall interface is not strictly respected in the output attenuation volumes and DVFs. Similarly when a rigid tumour is superimposed onto deforming anatomical structures, it results in disappearing and reappearing structures at the tumour boundaries causing an inconsistency not only between the output XCAT image and the DVF but also between the tumour and the anatomy, since no one-to-one mapping exists in such a scenario. This presents a substantial limitation for RT applications – in particular SBRT – where the aim is to deliver a high dose to the tumour with a sharp gradient at the tumour boundary. Mathematically, the inconsistencies in the XCAT DVFs can manifest as folding which makes them not invertible. Folding is defined as two points in image A being mapped to the same point in image B and as a result an inverse transformation (or mapping) does not exist. DVFs in the direction inverse to the one output by the XCAT simulator are for instance required to resample images from the first to consecutive time points or in general to establish a mapping between arbitrary time points.

Another potential limitation of the XCAT phantom for dose calculation is that lung tissue density is simulated to be constant and uniform whereas, in reality, lung volume changes with respiration while approximately maintaining the same mass resulting in tissue density changes and non-uniformity during respiration (Guerrero et al., 2006). In other words, a constant density of a region undergoing volume change would violate mass-preservation.

Williams et al. (2013) proposed a modified mass-conserving XCAT phantom with a deforming tumour and showed a substantial difference in delivered dose when compared to

the original XCAT. The authors introduced density changes to the lung tissue on the basis of the Jacobian determinant of the XCAT DVFs. However, when non-physical folding occurs, this manifests as negative values of the Jacobian determinant. To avoid scaling with negative values, Williams et al. (2013) applied a median filter to the volume change measurements. While their work addressed the intensity changes required to generate a mass-conserving XCAT phantom and the inconsistencies between a tumour and the surrounding tissue by deforming the tumour with the underlying anatomy, they did not correct any inconsistencies in the XCAT DVFs directly and did not address the sliding interface issue.

Here we propose a post-processing framework for the XCAT allowing to generate a time-resolved phantom with deforming lung tumour, realistic respiration-induced lung density variations as well as known, consistent DVFs and their inverse preserving the sliding at the lung/chest wall interfrace. The post processing framework contains three important steps described in sections II.A. to II.C. First, only the initial XCAT volume with a tumour is generated while the DVFs for the entire time-series are generated. Second, the DVFs are processed for consistency and inverted (see section II.B.). Third, the post-processed XCAT time-series is obtained by deforming the reference XCAT including the tumour using the generated DVFs. At the same time lung densities are scaled according to measured volume changes (see section II.C.). The final data-set contains the post-processed XCAT time-series and the DVFs in both directions. We validate the implemented method in section II.D. and demonstrate its applicability by using it as an MIDR validation tool in section II.E. The results are presented and discussed alongside other possible applications in sections III. and IV. respectively.

II. Material and methods

The proposed post-processing framework is illustrated in figure 1 and described in detail below. The source-code can be accessed here: https://github.com/UCL/cid-X.

II.A. XCAT phantom

The XCAT simulator is used to generate the first anatomical image of a time-series $i = 0$ and all DVFs for the consecutive time points $i = [1, \ldots, N]$ according to input displacement
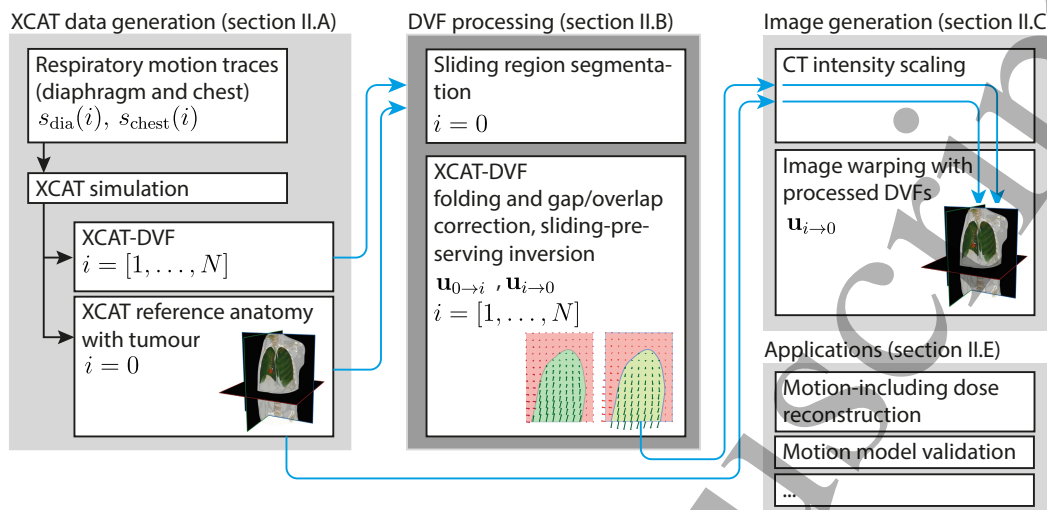
Figure 1: XCAT post-processing steps. Sections describing the different steps are indicated.

traces for diaphragm and chest, i.e. $s_{\mathrm{dia}}(i)$ and $s_{\mathrm{chest}}(i)$ (see figure 1, 'XCAT data generation'). This information is then used by the proposed post-processing framework as described below (see figure 1, 'DVF processing' and 'image generation').

II.B. DVF post-processing

Overview   Internal organs may slide within limits along each other. This sliding motion however is in general small except the one occurring along the chest wall due to respiration. Hence, here we focus on the correct handling of the more significant sliding along the chest wall. We assume that regions that slide along each other can deform independently, but deformations are constrained at the sliding interface such that gaps and overlaps must not occur. Hence we generate a separate DVF for each sliding region and apply a no-gap-overlap constraint at the sliding interface. Gaps and overlaps occur if, after the independent transformations, the sliding interfaces no longer coincide. This is addressed by locally correcting the DVFs. A DVF that maps from time a to b is denoted as $\mathbf{u}_{a \to b}(\mathbf{x})$, where each vector originates in a and ends in b. Hence, to deform an image $I_a(\mathbf{x})$ from a to b we require $\mathbf{u}_{b \to a}(\mathbf{x})$, such that $I_b(\mathbf{x}) = I_a(\mathbf{u}_{b \to a}(\mathbf{x}))$. This resampling operation is sometimes referred to as "pull" configuration.
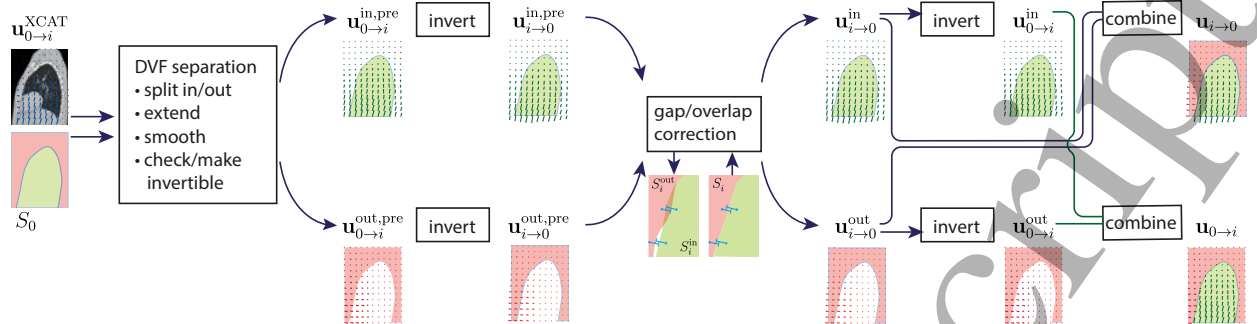
Figure 2: Workflow used to generate the sliding-preserving forward and backward DVF of the XCAT-generated transformation. 'In' and 'out' regions as defined by the signed distance map S are coloured in green and red respectively. Where processing is performed separately for each region, the complementary colour is set to white for illustration purposes. Processing steps for each box are described in detail in the text.

Sliding region segmentation To determine the location of a smooth and plausible sliding interface, we follow an approach adapted from Vandemeulebroucke et al. (2012), where a motion mask is generated on the basis of anatomical features derived from a CT image, namely a body mask, the bony anatomy, and the lungs. From these features an initial levelset contour and a corresponding speed function is derived. The evolution of the levelset results in the sliding interface coinciding with the zero-crossing of the final levelset function. In contrast to Vandemeulebroucke et al. (2012) we do not need to extract the features from an image, but can access the segmentations from the XCAT output directly and thus utilise the levelset evolution to obtain a smooth and plausible sliding interface. The sliding interface segmentation only needs to be generated for the first time point in a DVF-time series and is saved as a signed distance function S_0. Negative values of S_0 indicate the internal organs such as lungs, heart, liver, whereas positive values indicate the chest wall including ribs, muscles, skin etc.; DVFs associated with the inner and outer region are indicated by 'in' and 'out' respectively.

In order to produce consistent forward and backward DVFs, we follow the procedure outlined in figure 2 which was developed to allow plausible motion along the sliding interface. The following paragraphs correspond to the processing steps depicted as boxes in the figure.

DVF separation The initial XCAT DVF is split into the two main sliding regions as indicated by the sign of S_0. To obtain two DVFs for the whole field of view without a sharp fall-off to zero at the sliding interface, a nearest neighbour interpolation is used to fill the

II.B.   DVF post-processing

area at the other side of the sliding interface with a reasonable estimate. A smooth fall-off is introduced by linearly scaling the DVF back to zero after a distance of $5\,\mathrm{mm}$ measured from the sliding interface. Thereafter, both extended DVFs are smoothed with a Gaussian kernel ($\sigma = 4\,\mathrm{mm}$[1]) and checked for folding by evaluating the determinant of the Jacobian matrix based on finite forward differences. If required, local smoothing with a small Gaussian kernel is applied until no more folding is detected.

Inversion The previous step results in the two separate DVFs $\mathbf{u}_{0\to i}^{\mathrm{in,pre}}$ and $\mathbf{u}_{0\to i}^{\mathrm{out,pre}}$. These two preliminary DVFs, $\mathbf{u}_{0\to i}^{\mathrm{in,pre}}$ and $\mathbf{u}_{0\to i}^{\mathrm{out,pre}}$ are inverted to obtain $\mathbf{u}_{i\to 0}^{\mathrm{in,pre}}$ and $\mathbf{u}_{i\to 0}^{\mathrm{out,pre}}$. All inversions are performed using the open-source nifty-reg software package (Modat et al., 2010) as described by Veiga et al. (2015).

Gap/overlap correction Due to independent processing of both DVFs, gaps and overlaps may have been introduced at the sliding interface – or could have been existent in the original simulation. Gaps and overlaps between two sliding regions can be detected by the sign of the product of the signed distance function that is transformed twice, i.e. $S_i^{\mathrm{in}}(\mathbf{x}) = S_0(\mathbf{u}_{i\to 0}^{\mathrm{in,pre}}(\mathbf{x}))$ and $S_i^{\mathrm{out}}(\mathbf{x}) = S_0(\mathbf{u}_{i\to 0}^{\mathrm{out,pre}}(\mathbf{x}))$, i.e. gaps and overlaps occur at those positions where $S_i^{\mathrm{in}}(\mathbf{x}) \cdot S_i^{\mathrm{out}}(\mathbf{x}) < 0$ (Eiben et al., 2018). At these locations, the DVFs needs to be corrected such that the zero crossing of the transformed signed distance maps coincides at a new, 'best compromise' sliding interface position at time point i. We find this new sliding interface by identifying the zero-crossing of $S_i^{\mathrm{in}}(\mathbf{x}) + S_i^{\mathrm{out}}(\mathbf{x})$ and generate a corresponding signed distance function, $S_i(\mathbf{x})$ with the same zero-crossing. Now an appropriate correction needs to be applied to the preliminary inside and outside DVFs $\mathbf{u}_{i\to 0}^{\mathrm{in,pre}}$ and $\mathbf{u}_{i\to 0}^{\mathrm{out,pre}}$ such that the zero-crossing of the transformed version of S_0 coincides with the one of S_i. The length of these corrections can be directly measured by computing $L_i^{\mathrm{in}} = S_i^{\mathrm{in}} - S_i$ and $L_i^{\mathrm{out}} = S_i^{\mathrm{out}} - S_i$. The direction of these corrections should be orthogonal to the sliding surface at the point where the preliminary 'in' or 'out' DVF is pointing to, i.e. $\mathbf{D}_i^{\mathrm{in}}(\mathbf{x}) = \nabla S_0(\mathbf{u}_{i\to 0}^{\mathrm{in,pre}}(\mathbf{x}))$. This correction should be applied fully only where a gap and overlap exists and otherwise should smoothly fall off to zero. To achieve this an initial weight image is generated with values of 1 in the gap/overlap area, i.e. $S_i^{\mathrm{in}}(\mathbf{x}) \cdot S_i^{\mathrm{out}}(\mathbf{x}) < 0$ and zero otherwise. Then we apply a

---

[1]The smoothing kernel width was determined empirically to balance a global DVF smoothing to be low and at the same time keep the number of local smoothing operations at reasonable processing time.

II.B. DVF post-processing

linear diffusion process (see e.g. Weickert (1998)) modified such that the gap/overlap area takes the role of a constant source - resulting in full gap/overlap correction where it occurred and a smooth transition to the surrounding area. With the final weight image $W_i(\mathbf{x})$ the 'inside' gap/overlap corrected DVF is calculated $\mathbf{u}_{i\rightarrow0}^{\mathrm{in}} = \mathbf{u}_{i\rightarrow0}^{\mathrm{in,pre}}(\mathbf{x}) + W_i L_i^{\mathrm{in}}\mathbf{D}_i^{\mathrm{in}}(\mathbf{x})$ with $\mathbf{u}_{i\rightarrow0}^{\mathrm{out}}$ following accordingly.

Combining inside and outside DVFs The corrected DVFs are once more inverted to obtain the DVFs in the original direction resulting in updated DVFs from time point 0 to i which are finally combined into a single DVF using the signed distance map S_0.

$$\mathbf{u}_{0\rightarrow i}(\mathbf{x}) = \begin{cases} \mathbf{u}_{0\rightarrow i}^{\mathrm{in}}(\mathbf{x}) & \text{if } S_0(\mathbf{x}) < 0 \\ \mathbf{u}_{0\rightarrow i}^{\mathrm{out}}(\mathbf{x}) & \text{otherwise} \end{cases}, \tag{1}$$

Combinations of DVFs for time points i follow accordingly.

II.C. Intensity scaling and time-series XCAT generation

Mathematically, the determinant of the Jacobian of the DVF represents the local volume change and can therefore be used to directly quantify the density change in a mass-conserving case. Lung volume and density changes are mainly due to air-flow and can therefore be scaled using the determinant of the Jacobian of the DVF provided that the intensity is proportional to mass-density and has a value of zero for air. The reference XCAT, X_0, was obtained in HU which are proportional to electron density (and therefore, in good approximation, proportional to mass-density) but where the intensity of air is -1000 HU.

Therefore the HU intensity within the lungs of X_0 was first scaled for each time instance i according to

$$X_0^{\mathrm{sc},i}(\mathbf{x}) = \begin{cases} \frac{X_0(\mathbf{x})+1000}{\det \mathrm{Jac}(\mathbf{u}_{0\rightarrow i}(\mathbf{x}))} - 1000 & \forall \mathbf{x} \in \Omega_{\mathrm{lungs},0} \\ X_0(\mathbf{x}) & \text{otherwise} \end{cases}, \tag{2}$$

where $\det \mathrm{Jac}(\mathbf{u}_{0\rightarrow i})$ is the determinant of the Jacobian of the DVF (always positive following the post-processing) and $\Omega_{\mathrm{lungs},0}$ defines the domain of the lungs at the time instance $i = 0$.

Thereafter, the final post-processed anatomical image – including the tumour – for each time point, X_i^{post}, was obtained by warping the result of (2) with the post-processed DVF:

$$X_i^{\mathrm{post}}(\mathbf{x}) = X_0^{\mathrm{sc},i}(\mathbf{u}_{i\rightarrow0}(\mathbf{x})). \tag{3}$$

Because X^{post} are obtained from a reference XCAT with uniform lung density of $-742\,\mathrm{HU}$, different starting points for the XCAT-DVF generation will result in different lung density distributions over the time-series. We compared the lung density for the X_i^{post} with a starting point at the end-exhale ($\phi = \pi/2$), mid-exhale ($\phi = \pi/6$) and end-inhale ($\phi = \pi$) in equation 4 to those observed in a real patient (Guerrero et al., 2006).

Transforming the first image containing the tumour to all other time points by the post-processed DVFs ensures that the tumour moves and deforms according to its location in the lungs with the tumour boundaries following the underlying deformation (Williams et al., 2013).

II.D. Validation experiments

II.D.1. XCAT data generation

The XCAT phantom was generated with an isotropic voxel size of $2\,\mathrm{mm}$ with a $35\,\mathrm{mm}$-diameter spherical tumour in the lower right lung. Data were generated at $f = 25\,\mathrm{Hz}$ for regular motion following the displacement curves:

$$
\begin{aligned}
s_{\mathrm{dia}}(i) &= A_{\mathrm{dia}} & \sin^4(ki + & 0.5\pi & +\phi), \\
s_{\mathrm{chest}}(i) &= A_{\mathrm{chest}} & \sin^4(ki + & 0.45\pi & +\phi),
\end{aligned}
\tag{4}
$$

where i is the time-series instance, $k = \pi/N$, ϕ determines the starting phase, and N is the number of instances per simulated breath cycle. Chest and diaphragm amplitudes were selected as $A_{\mathrm{chest}} = 10\,\mathrm{mm}$ and $A_{\mathrm{dia}} = -30\,\mathrm{mm}$ and time instances $i = [0, \dots, N]$, with $N = 79$ were generated. The $\sin^4$ pattern and the amplitude were chosen such that the tumour moves with an SI amplitude of about $2\,\mathrm{cm}$, spends more time at exhale and the phase difference between *diaphragm* and *chest* was introduced to cause hysteresis in the tumour trajectory as observed in real patients (Seppenwoolde et al., 2002). No cardiac motion was applied.

II.D.2. DVF inconsistencies

In order to motivate our post-processing framework we selected adjacent, exemplary structures and followed their displacement firstly according to the original XCAT-DVF and secondly according to the post-processed DVF. The voxel classes were selected by their corre-

sponding label in the XCAT DVF text output file. Furthermore, to quantify the extent of the folding occurrences in the original XCAT and the improvement provided by our post-processing framework, the number of negative Jacobian determinant values for the data generated as described in the previous section before and after the post-processing was measured. Gradient calculations for computing the Jacobian determinant were calculated using a finite forward difference.

II.D.3.   Geometric fidelity

To quantify the consistency of the forward and backward DVFs produced by our framework, we computed the symmetric inverse consistency by composing $\mathbf{u}_{i \to 0}(\mathbf{u}_{0 \to i}(\mathbf{x}))$ and $\mathbf{u}_{0 \to i}(\mathbf{u}_{i \to 0}(\mathbf{x}))$ at each voxel within the simulated anatomy. The two measurements of residual Euclidean displacement lengths from both compositions were combined and statistical evaluation was performed on the combined set.

To quantify the deviation of the tumour motion from the original XCAT, the X_i^{post} series was compared to the conventional XCAT series, X_i^{conv} obtained from the XCAT software, with non-deforming tumour and constant and uniform lung density. Volumes were overlaid and compared visually. Tumour centroid trajectories and tumour volumes were compared.

II.E.   Application: Motion including dose reconstruction

We demonstrate the use of the post-processed XCAT for MIDR validation. For this, the subset of X_i^{post} representing the ten phases of a 4DCT were selected and the mid-ventilation instance was chosen for treatment planning using a 9-beam step-and-shoot intensity modulated radiotherapy (IMRT) technique following the RTOG 1021 protocol to deliver 54 Gy to 95% of the planning target volume (PTV) in 3 fractions. Planning was performed for an Elekta linac with Agility MLC (Elekta AB, Stockholm, Sweden) using the Elekta treatment planning system (TPS) Monaco (research version 5.40.00) with a Monte Carlo dose engine (2% uncertainty per calculation). Note that a single X^{post} instance can represent the mid-ventilation phase of the 4DCT given that the motion is regular and would results in perfect sorting.

Treatment delivery was simulated in our in-house emulator (Fast et al., 2014). The

MIDR used in this study was described by (Menten et al., 2020) and further extended to account for variable anatomy. In short, the total treatment fluence was discretized into sub-beams and the isocenter of each sub-beam was shifted to emulate rigid motion of the entire anatomy similar to Poulsen et al. (2012). We used binning sizes of $1\,\mathrm{mm}$ for the isocenter positions and $0.5\,\mathrm{mm}$ difference in a single leaf position for the different aperture shapes. The final dose of the motion encoded plan on the planning CT was calculated in the TPS.

MIDR using the isocenter shift method was validated using X^{post} as a ground truth moving anatomy. In this case, each sub-beam is associated with the XCAT instance at which it was delivered to provide the actual beam/anatomy configuration. For each XCAT instance, the dose from all associated sub-beams was calculated in the TPS and exported for accumulation onto the planning CT in MatlabR2018a (The Mathworks, Natick, MA) using the known DVFs and direct dose mapping. The difference between Shift-MIDR and ground truth MIDR represents the error on Shift-MIDR due to the simplicity of the motion model.

In order to quantify the effect of lung density variations, ground-truth MIDR was performed for X_i^{post} as well as for the post-processed XCAT but without lung density scaling $X_i^{\mathrm{densConst}} = X_0(\mathbf{u}_{i\to0}(\mathbf{x}))$. Note that the tumour is deforming in both cases and the same DVFs are used for dose accumulation, however $X_i^{\mathrm{densConst}}$ has a constant and uniform lung density of -742 HU.

MIDR using the conventional XCAT was not feasible because no one-to-one mapping exists in this case and the DVFs are available only in one direction, i.e. the dose cannot be warped onto the planning CT. Dose from any instance of the conventional XCAT can nonetheless be warped onto the first instance (different from the planning CT) using the conventional XCAT DVFs. We simulated delivery of the whole treatment plan at the end-exhale and at the end-inhale phase using X^{post}, $X^{\mathrm{densConst}}$ and the conventional XCAT, X^{conv}. The dose was then warped back onto the first instance using either the post-processed DVFs (for X^{post}, $X^{\mathrm{densConst}}$) or the conventional XCAT-DVFs (for X^{conv}). The warped doses were compared to illustrate the effect of tumour deformation alone (difference between X^{conv} and $X^{\mathrm{densConst}}$).

II.E.   Application: Motion including dose reconstruction

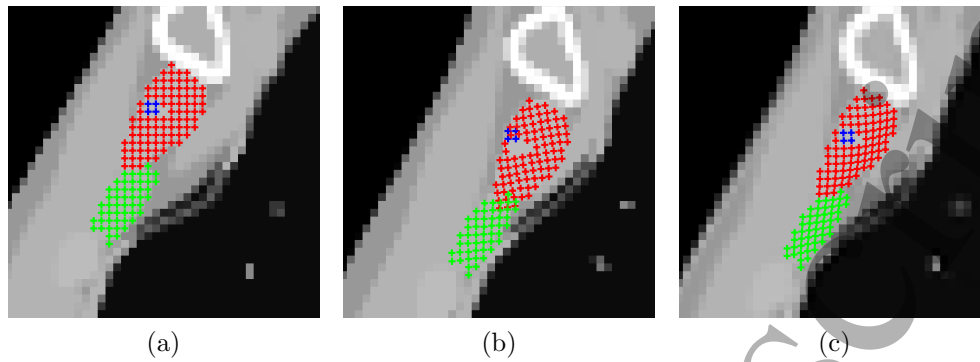|		
(a)	(b)	(c)

Figure 3: *Phantom anatomy and three selected structures, i.e. rib muscle (red), left rib (green), and sternum (blue) at the first time point, mid-exhale (a) and end exhale (b,c). (a) The XCAT image and structures at the first time point, (b) XCAT output image and the points from the first time point pushed as defined by the XCAT-DVF, and (c) the same points pushed according to the post-processed DVF with the corresponding image produced by the post-processing framework.*

III. Results

III.A. DVF inconsistencies

Figure 3 shows a zoomed-in field of view of the upper thorax in sagittal orientation. The grey-scale images show the anatomy, whereas the crosses mark three different structures as defined in the XCAT output in different colours. Using the original XCAT-DVF to transform the structural points from the first time point (3a, mid-exhale) to time point $i = 19$ (exhale) results in structures moving through each other (3b). The transformation according to the post-processed DVF eliminates this implausible motion (3c). We also refer the reader to the supplementary animation for a visualisation of the full breathing cycle.

Figure 4 shows the folding occurrences – i.e. negative Jacobian determinant values – in a sagittal slice for the original (red, 4a) and the post-processed (green, 4b) DVFs at time point $i = 19$. Furthermore, the total number of folding occurrences are shown in figure 4c. Note, that for the post-processed DVF, each region (i.e. inside and outside) does *not* contain any negative Jacobian determinant values. However, when calculating the Jacobian determinant numerically on the final, combined DVF with a finite difference operator, the gradient is measured across region boundaries and thus appears to contain folding.
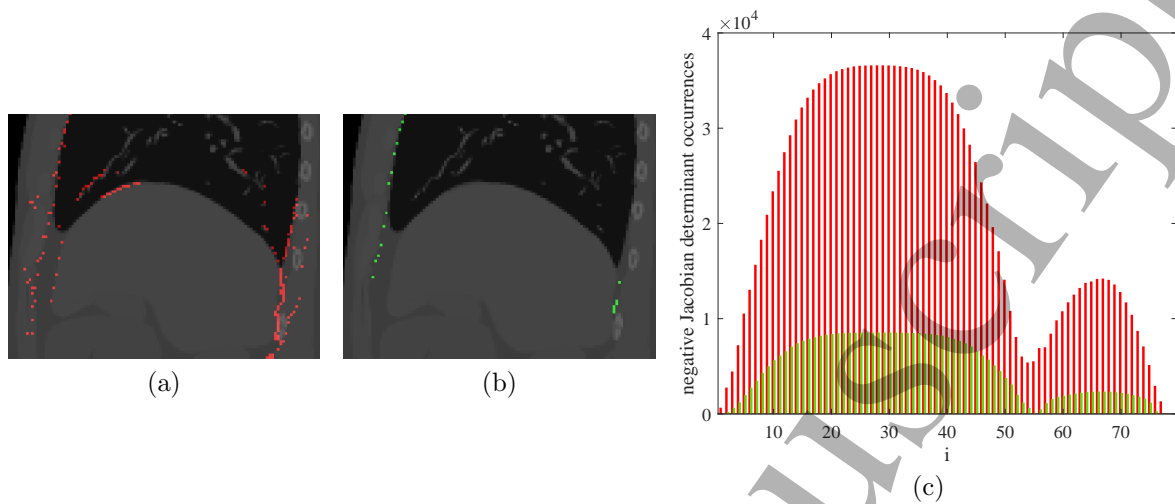
Figure 4: Negative Jacobian determinant occurrences for the original XCAT DVF (a) and the post-processed DVF (b) at time point $i = 19$. (c) The total number of negative Jacobian determinant values for both DVFs for all simulated time points. Folding in the post-processed DVF is detected only at the sliding boundary, where the gradient is calculated across the region boundary.

III.B.   Geometric fidelity

The maximum 95th percentile of the inverse consistency distance over all time points was 0.02 mm whereas figure 5a shows the 99th percentile with a maximum inverse consistency error of 0.5 mm. Larger inverse-consistency errors (≥ 1 mm) were confined to the sliding interface which is shown in figure 5b for time-instance $i = 19$ which has the largest motion amplitude and inverse consistency error. The localisation of the largest inverse consistency errors to this region can be attributed to method used to compose the DVFs. For composing DVFs, vector fields need to be evaluated at inter-grid locations and thus depend on an interpolation method. Our composition method uses a tri-linear interpolation. While this interpolation assumption is accurate for a continuous DVF it leads to errors close to the sliding boundary where there is a clear discontinuity.

X^{conv} and X^{post} were compared visually (see supplemental video) showing that the post-processing did not change the motion pattern substantially. As illustrated in figure 6, the post-processing did not alter the tumour centroid trajectory. However, tumour volume was constant at 24.3 cc in X^{conv} but varied between 19.4 and 27.5 cc in X_i^{post} correlating with the tumour trajectory.
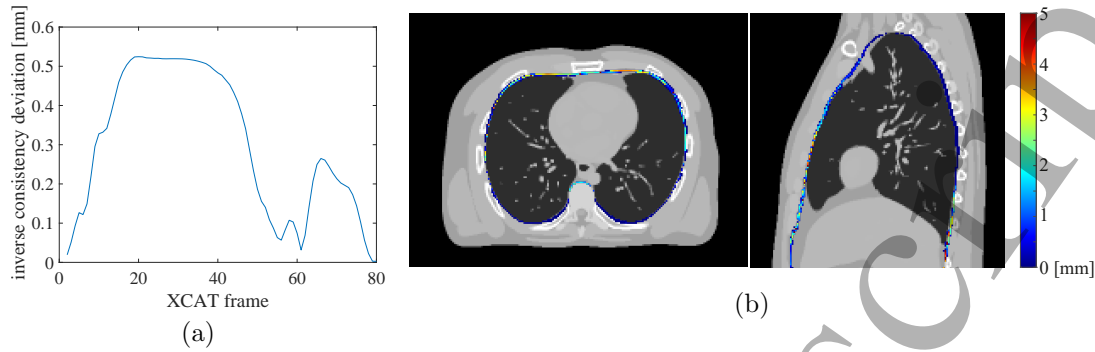
III.B.   Geometric fidelity

Figure 5: (a) 99th percentile of the inverse consistency deviation from the identity transformation measured in mm inside the simulated anatomy after composition of forward and backward transformation for each time point. (b) Colour overlay of the inverse consistency error overlayed onto the simulated anatomy for time point 19. Only errors larger than 0.01 mm are shown.

III.C.   Intensity scaling and time-series XCAT generation

The lung density varied with respiration and depended on the starting point of the original XCAT simulation because the following volumes were deformed (and lung density was scaled) from X_0 with uniform lung density (Figure 7). The median value and distribution of lung density for a starting point at mid-exhale were the most realistic compared to those observed in a patient by Guerrero et al. (2006). For the remainder of this paper, all results are shown for a starting point at mid-exhale.

III.D.   Motion including dose reconstruction

Figure 8 shows the dose volume histogram (DVH) for the simulated delivery of the SBRT plan to a static anatomy (reference, full lines) as well as for the reconstructed dose to the moving anatomy using the isocenter shift method on the planning CT, (Shift MIDR, dotted lines) and for the ground truth MIDR (GT MIDR) using X^{post} (dashed lines). The dose to 95% of the GTV volume (D95%) was 60.6 Gy in the static delivery but only 50.5 Gy in GT MIDR and 53.7 Gy in Shift MIDR (red line). The dose to the spinal cord (blue line) was very similar in the static delivery and GT MIDR (D2% was 12.9 Gy in GT MIDR and 13 Gy in the static deliver) but the Shift MIDR DVH differs substantially (D2% was 12.5 Gy). For the oesophagus (yellow line), D2% was 8.8 Gy in the static case, 8.5 Gy for GT MIDR and 8.2 Gy for Shfit MIDR while for the heart (magenta line), V20Gy was 5.3% for the static
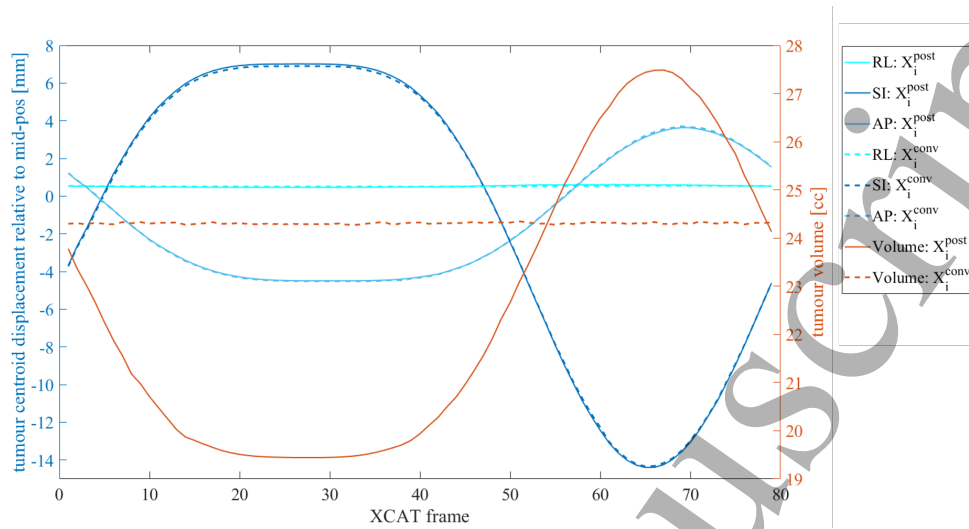
Figure 6: Tumour centroid trajectory (blue) and volume (red) for X_i^{post} (full line) and X_i^{conv} (dashed line).

case, 4.5% for GT MIDR and 4.3% for Shift MIDR.

The difference for GT MIDR on X_i^{post} and MIDR on $X_i^{\text{densConst}}$ is shown in Figure 9. Absolute dose differences in the order of 1 Gy were observed mainly in the lung tissue surrounding the tumour which resulted in negligible differences in DVH parameters.

The full plan dose was delivered on the end-exhale and end-inhale instances of the phantom for X^{post} (Figure 10) and $X^{\text{densConst}}$ using the post-processed DVFs and X^{conv} using the conventional XCAT DVFs. Because of the sliding motion at the chest-wall interface, the dose distribution is sharp in this region, particularly for the inhale case (Figure 10, red arrow). The difference between $X^{\text{densConst}}$ and X^{conv} illustrates the effect of the tumour rigidity in the conventional phantom. The difference between X^{post} and X^{conv} accounts, in addition, for the difference in lung tissue density at the two extremes of the breathing cycle.
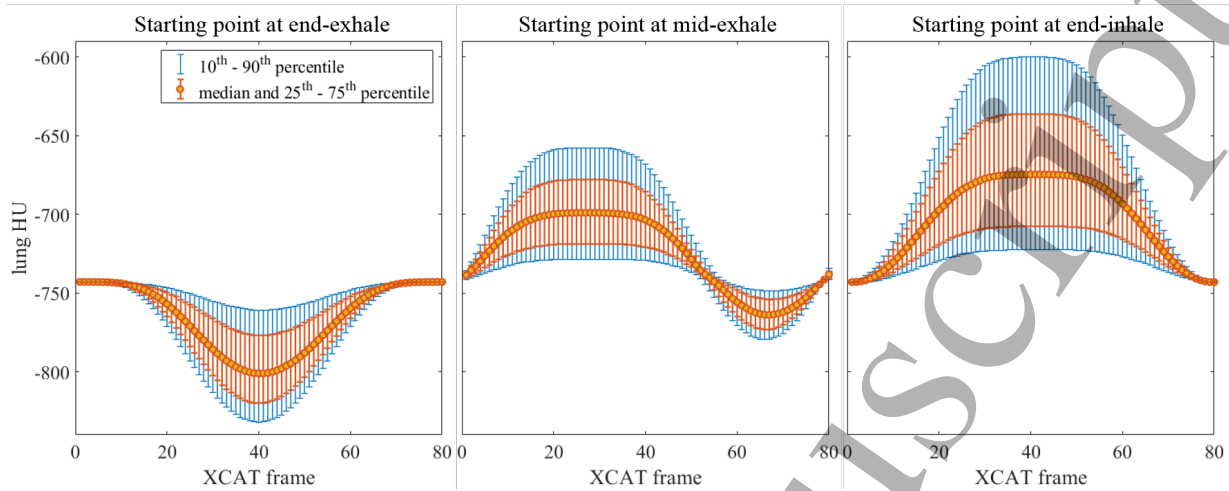
III.D.   Motion including dose reconstruction

Figure 7: Lung density distribution for X_i^{post} series with reference volume at the end-exhale (left), mid-exhale (centre) and end-inhale (right). The marker shows the median density, the red and blue errorbars show the 25-75th and 10-90th respectively.
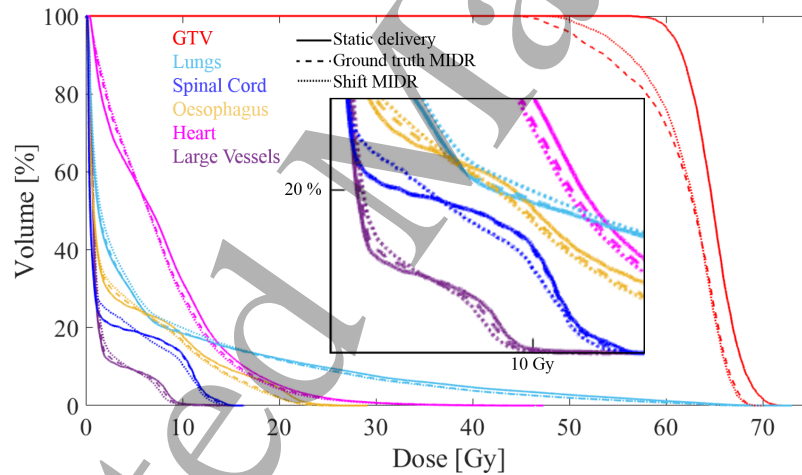

Figure 8: DVHs for the static delivery (reference, full lines), for MIDR using the isocenter shift method (dotted lines) and using the ground truth moving anatomy X^{post} (dashed lines).
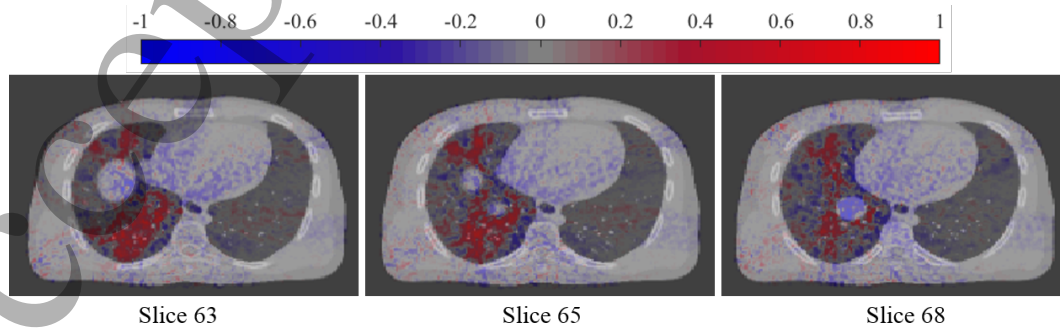

Slice 63   Slice 65   Slice 68

Figure 9: MIDR dose difference $(X_i^{post} - X_i^{densConst})$ in Gy for three different axial slices.

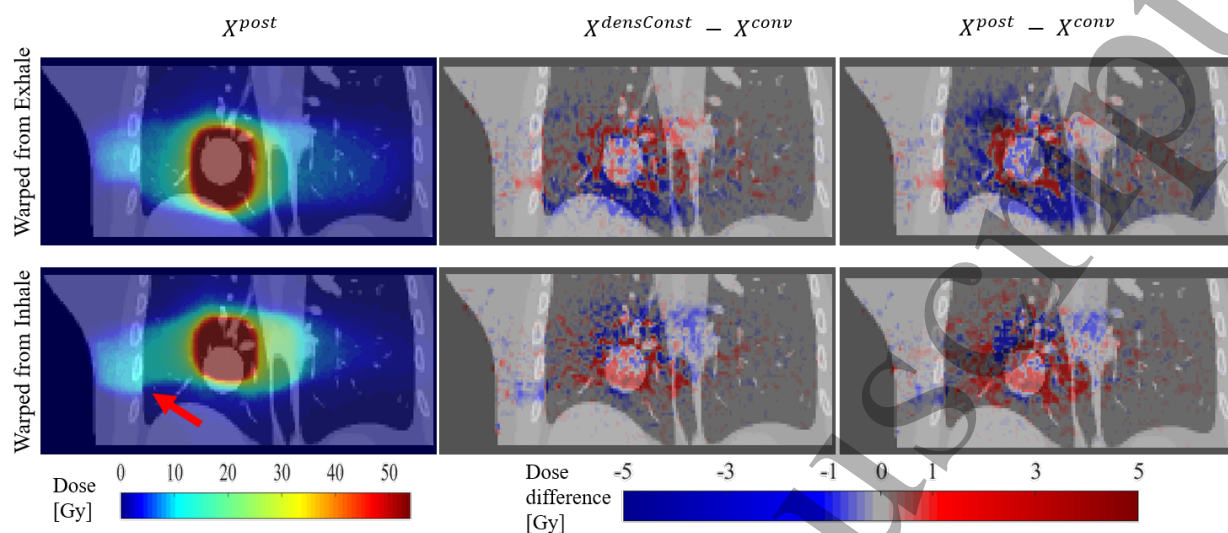III.D.   Motion including dose reconstruction

Figure 10: *Left: Dose delivered on exhale (top) or inhale (bottom) of X^{post} warped onto the first volume of the series using the post-processed DVFs in a coronal view. Note the sharpness of the dose distribution at the chest-wall interface because of the sliding motion. Middle: Difference between the dose warped from $X^{densConst}$ and X^{conv}. Right: Difference between the dose warped from X^{post} and X^{conv}.*

IV.   Discussion

In this study, we introduced a post-processing framework for the digital 4D XCAT phantom which generates a 4D anthropomorphic phantom moving with respiration. Tumours included in the lungs move consistently according to the motion of the surrounding lung tissue because a single DVF is used to describe the motion and deformation of the lungs and tumour. At the same time lung density changes depending on the local deformation are simulated and the sliding motion at the lung/chest wall interface is preserved. The forward and backward DVFs from the first (reference) volume to the following volumes are also generated and were used for dose accumulation between different respiratory states.

The post-processing addresses some shortcomings of the original XCAT phantom such as the inconsistency between XCAT image and original DVF which was caused by a rigid tumour superimposed on a deforming anatomy. The availability of forward and backward DVFs facilitates – by the means of composition – a mapping between arbitrary time points and also makes this phantom more readily usable to validate applications that do not produce DVFs in the same direction as the original XCAT simulator, for instance motion models (McClelland et al., 2013). The DVF inversions were performed using the open-source nifty-reg

registration package (Modat et al., 2010; Veiga et al., 2015) but this could easily be replaced with another implementation if required. We demonstrated the consistency of the forward and backward DVFs which link different time points by the means of inverse consistency measurements for the whole simulated anatomy. A tumour – including its boundaries – placed anywhere within will have the same consistency because it follows the motion as defined by the DVFs. The no-gap-overlap criterion used for the post-processing of the DVFs can correct potential organ overlap at the sliding interface, however, in any case care should be taken when choosing the input motion traces such that visible overlap is minimised since the correction will impact the XCAT simulated deformations locally.

Williams et al. (2013) had previously proposed a post-processing method for the XCAT phantom that implements density changes and a deforming tumour. Although they require deformations in the forward and backward direction for mapping the volume changes from one time point to the other, they did not address the inconsistencies in the XCAT DVFs explicitly. Instead they used a DVF inversion method that usually converges to a pseudo-inverse (Chen et al., 2008) when folding occurs and thus mathematically no inverse exists. In the present study, we proposed to generate a smooth, non-folding DVF per region while preserving the sliding interface between chest wall and inner organs. This allows an inverse DVF to be calculated numerically for each region without relying on an algorithm to converge to a pseudo-inverse. Transformations that contain folding and their pseudo-inverses have unavoidable inverse consistency errors, whereas using a non-folding transformation and the numerical estimate of its inverse can result in arbitrarily small inverse consistency errors (except at the sliding boundaries), limited only by the accuracy of the inversion algorithm.

While our method preserves the most prominent sliding occurrence, i.e. sliding between the chest wall and the inner organs, it does not explicitly handle sliding between inner organs such as the heart and the lungs. Cardiac motion was not introduced in this study since the focus was on respiratory motion. Further work is therefore needed to extend the current method to also handle cardiac motion.

The determinant of the Jacobian of the DVF represents the local volume change and was used to scale lung density but the distribution of density values within the lung depends on the starting point of the XCAT simulation (figure 7). Lung density was more uniform close to the reference position with greater variations away from the reference. A reference

volume with non-uniform lung density would allow to obtain more realistic values. However, as observed in figure 9, this would likely have a negligible impact on dosimetric applications.

The phantom was used to validate MIDR for a regularly breathing phantom using the known DVFs for dose accumulation on a reference volume. In Figure 8, the DVHs calculated using an isocenter shift method to model motion during delivery can be compared to the ground truth using X^{post}. Note that the isocenter shift method is not intended to accurately reconstruct the dose to the OARs (Poulsen et al., 2012). However, in this case, GTV D95% was also overestimated by Shift MIDR. Even though the relative position of the sub-beams with respect to the target positions is correct in Shift MIDR, the beam attenuation may differ due to motion of the surrounding anatomy which results in errors on the reconstructed target dose. With the possibility to calculate the actually delivered dose to the target and to the OARs, X^{post} can be used to validate other motion models for dose reconstruction such as the use of a 4DCT (Kamerling et al., 2016) or more sophisticated deformable motion models (McClelland et al., 2013) with irregular breathing patterns.

The impact of local lung density variations on deposited dose was limited to about 1 Gy locally, resulting in negligible differences in DVH endpoints. The effect of lung density changes alone is therefore expected to be small for MIDR. In the extreme cases of full dose delivery at end-exhale or end-inhale, a comparison with the original XCAT (with a rigid tumour) (Figure 10) illustrates the relative importance of tumour deformation alone ($X^{\text{densConst}} - X^{\text{conv}}$) compared to tumour deformation and lung density variations combined ($X^{\text{post}} - X^{\text{conv}}$). In the present study the tumour volume variations were large (range: 19.4–27.5 cc) and no scaling of the tumour density was applied. Real-life tumour volume and density changes should be investigated to ensure realistic variations in the phantom. As discussed in Williams et al. (2013), tumours tend to rigidify the local lung tissue. This could be addressed in several ways. A dedicated tumour NURBS object could be added to the XCAT phantom to offer more realistic tumour deformation. Due to the closed-source nature of the XCAT phantom however, this approach can only be followed by the XCAT inventors and developers. Alternatively, in the future the post-processing framework could be extended with a local incompressibility constraint similar to approaches used in image registration (Rohlfing et al., 2003). Along the same lines, further local DVF modifications could be implemented in the future to simulate effects such as tumour baseline shifts or infiltration of the tumour into the chest wall with corresponding local motion restrictions.

The proposed phantom with corresponding forward and backwards DVFs can be used as a ground truth to validate the geometric (Paganelli et al., 2019; Eiben et al., 2019) and dosimetric (Bertholet et al., 2019a) accuracy of different types of motion models such as rigid shifts (Poulsen et al., 2012; Kamerling et al., 2017), 4D imaging (Kamerling et al., 2016; Paganelli et al., 2018a) or deformable motion models (McClelland et al., 2013; Paganelli et al., 2019). Other applications include the validation of DIR-based image processing for motion mitigation strategies (van de Lindt et al., 2019) and 4DCT-based planning methods comparison (Wolthaus et al., 2008).

V.   Conclusion

The implemented open-source post-processing framework adds important functionality to the original XCAT phantom by generating sliding-preserving forward and backward ground-truth deformation vector fields which in turn can be used to more realistically adapt the previously constant lung CT intensity values. The framework warps the anatomy and tumour together which efficiently eliminates previously existing inconsistencies at the tumour boundary. This will make the phantom applicable to a wider range of radiotherapy related applications such validation of motion-including dose reconstruction methodologies or deformable image registration and motion modelling techniques.

Acknowledgements

Appendix

A   Animated, post-processed XCAT phantom

The supplemental video shows the time series for X^{post} as well as the histogram for the lung density changes and the overlay of the original XCAT and the post-processed XCAT.

References

J. Bertholet, B. Eiben, M. J. Menten, E. H. Tran, D. J. Hawkes, A. Wetscherek, S. Nill, J. R. McClelland, and U. Oelfke. In silico validation of motion-including dose reconstruction for MR-guided lung SBRT using a patient specific motion model. In *19th International Conference on the use of Computers in Radiation Therapy, Montreal, CA*, 2019a.

J. Bertholet, A.-C. Knopf, B. Eiben, J. R. McClelland, A. Grimwood, E. Harris, M. J. Menten, P. R. Poulsen, D. T. Nguyen, P. J. Keall, and U. Oelfke. Real-time intrafraction motion monitoring in external beam radiotherapy. *Physics in Medicine and Biology*, 64 (15):TR01, 2019b. doi: 10.1088/1361-6560/ab2ba8.

K. K. Brock, S. Mutic, T. R. McNutt, H. Li, and M. L. Kessler. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132: Report. *Medical Physics*, 44(7):e43–e76, 2017. doi: 10.1002/mp.12256.

M. Chen, W. Lu, Q. Chen, K. J. Ruchala, and G. H. Olivera. A simple fixed-point approach to invert a deformation field. *Medical Physics*, 35(1):81–88, 2008.

I. J. Chetty and M. Rosu-Bubulac. Deformable Registration for Dose Accumulation. *Seminars in Radiation Oncology*, 29(3):198–208, 2019. doi: 10.1016/j.semradonc.2019.02.002.

S. Ehrbar, R. Perrin, M. Peroni, K. Bernatowicz, T. Parkel, I. Pytko, S. Klöck, M. Guckenberger, S. Tanadini-Lang, D. C. Weber, and A. Lomax. Respiratory motion-management in stereotactic body radiation therapy for lung cancer - A dosimetric comparison in an anthropomorphic lung phantom (LuCa). *Radiotherapy and Oncology*, 121(2):328–334, 2016. doi: 10.1016/j.radonc.2016.10.011.

B. Eiben, E. H. Tran, M. J. Menten, U. Oelfke, D. J. Hawkes, and J. R. McClelland. Statistical motion mask and sliding registration. In S. Klein, M. Staring, S. Durrleman, and S. Sommer, editors, *Biomedical Image Registration*, pages 13–23, Cham, 2018. Springer International Publishing. ISBN 978-3-319-92258-4.

B. Eiben, J. Bertholet, E. H. Tran, A. Wetscherek, U. Oelfke, and J. R. McClelland. Geometric accuracy of surrogate-driven respiratory motion models for MR-guided lung radiotherapy. In *19th International Conference on the use of Computers in Radiation Therapy, Montreal, CA*, 2019.

M. F. Fast, S. Nill, J. L. Bedford, and U. Oelfke. Dynamic tumor tracking using the Elekta Agility MLC. *Medical Physics*, 41(11):111719, 2014. doi: 10.1118/1.4899175.

M. Guckenberger, T. Krieger, A. Richter, K. Baier, J. Wilbert, R. A. Sweeney, and M. Flentje. Potential of image-guidance, gating and real-time tracking to improve accuracy in pulmonary stereotactic body radiotherapy. *Radiotherapy and Oncology*, 91(3):288–295, 2009. doi: 10.1016/j.radonc.2008.08.010.

T. Guerrero, K. Sanders, E. Castillo, Y. Zhang, L. Bidaut, T. Pan, and R. Komaki. Dynamic ventilation imaging from four-dimensional computed tomography. *Physics in Medicine and Biology*, 51(4):777–791, 2006. doi: 10.1088/0031-9155/51/4/002.

C. P. Kamerling, M. F. Fast, P. Ziegenhein, M. J. Menten, S. Nill, and U. Oelfke. Real-time 4D dose reconstruction for tracked dynamic MLC deliveries for lung SBRT. *Medical Physics*, 43(11):6072–6081, 2016. doi: 10.1118/1.4965045.

Consistent invertible XCAT DVFs   page 22

C. P. Kamerling, M. F. Fast, P. Ziegenhein, M. J. Menten, S. Nill, and U. Oelfke. Online dose reconstruction for tracked volumetric arc therapy: Real-time implementation and offline quality assurance for prostate SBRT. *Medical Physics*, 44(11):5997–007, 2017. doi: 10.1002/mp.12522.

P. J. Keall, G. S. Mageras, J. M. Balter, R. S. Emery, K. M. Forster, S. B. Jiang, J. M. Kapatoes, D. A. Low, M. J. Murphy, B. R. Murray, C. R. Ramsey, M. B. Van Herk, S. S. Vedam, J. W. Wong, and E. Yorke. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. *Medical Physics*, 33:3874–3900, 2006. doi: 10.1118/1.2349696.

J. McClelland, D. Hawkes, T. Schaeffter, and A. King. Respiratory motion models: A review. *Medical Image Analysis*, 17(1):19–42, 2013. doi: 10.1016/j.media.2012.09.005.

M. J. Menten, J. K. Mohajer, R. Nilawar, J. Bertholet, A. Dunlop, A. U. Pathmanathan, M. Moreau, S. Marshall, A. Wetscherek, S. Nill, A. C. Tree, and U. Oelfke. Automatic reconstruction of the delivered dose of the day using MR-linac treatment log files and online MR imaging. *Radiotherapy and Oncology*, 145:88–94, apr 2020. ISSN 01678140. doi: 10.1016/j.radonc.2019.12.010.

P. Mishra, S. St. James, W. P. Segars, R. I. Berbeco, and J. H. Lewis. Adaptation and applications of a realistic digital phantom based on patient lung tumor trajectories. *Physics in Medicine and Biology*, 57(11):3597–3608, 2012. doi: 10.1088/0031-9155/57/11/3597.

M. Modat, G. R. Ridgway, Z. A. Taylor, M. Lehmann, J. Barnes, D. J. Hawkes, N. C. Fox, and S. Ourselin. Fast free-form deformation using graphics processing units. *Computer Methods and Programs in Biomedicine*, 98(3):278 – 284, 2010. doi: https://doi.org/10.1016/j.cmpb.2009.09.002. HP-MICCAI 2008.

National Library of Medicine. Visible Human Male and Female datasets, http://www.nlm.nih.gov/ research/visible/visible_human.html.

C. Paganelli, P. Summers, C. Gianoli, M. Bellomi, G. Baroni, and M. Riboldi. A tool for validating MRI-guided strategies: a digital breathing CT/MRI phantom of the abdominal site. *Medical and Biological Engineering and Computing*, 55(11):2001–2014, 2017. doi: 10.1007/s11517-017-1646-6.

C. Paganelli, J. Kipritidis, D. Lee, G. Baroni, P. Keall, and M. Riboldi. Image-based retrospective 4D MRI in external beam radiotherapy: A comparative study with a digital phantom. *Medical Physics*, 45(7):3161–3172, 2018a. doi: 10.1002/mp.12965.

C. Paganelli, G. Meschini, S. Molinelli, M. Riboldi, and G. Baroni. patient-specific validation of deformable image registration in radiation therapy: Overview and caveats. *Medical Physics*, 45(10):e908–e922, 2018b. doi: 10.1002/mp.13162.

C. Paganelli, S. Portoso, N. Garau, G. Meschini, R. Via, G. Buizza, P. J. Keall, M. Riboldi, and G. Baroni. Time-resolved volumetric MRI in MRI-guided radiotherapy: an in-silico comparative analysis. *Physics in Medicine and Biology*, 2019. doi: 10.1088/1361-6560/ab33e5.

P. Poulsen, M. Schmidt, P. Keall, E. Worm, W. Fledelius, and L. Hoffmann. A method of dose reconstruction for moving targets compatible with dynamic treatments. *Medical Physics*, 39(10):6237–6246, 2012.

T. Rohlfing, C. R. Maurer, D. A. Bluemke, and M. A. Jacobs. Volume-preserving nonrigid registration of mr breast images using free-form deformation with an incompressibility constraint. *IEEE Transactions on Medical Imaging*, 22(6):730–741, 2003. doi: 10.1109/TMI.2003.814791.

W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. Tsui. 4D XCAT phantom for multimodality imaging research. *Medical Physics*, 37(9):4902–4915, 2010. doi: 10.1118/1.3480985.

W. P. Segars, B. M. W. Tsui, J. Cai, F. Yin, G. S. K. Fung, and E. Samei. Application of the 4-D XCAT phantoms in biomedical imaging and beyond. *IEEE Transactions on Medical Imaging*, 37(3):680–692, March 2018. ISSN 1558-254X. doi: 10.1109/TMI.2017.2738448.

Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. Van Herk, J. V. Lebesque, and K. Miyasaka. Precise and Real-Time Measurement of 3D Tumor Motion in Lung due to Breathing and Heartbeat, Measured during Radiotherapy. *International Journal of Radiation Oncology*Biology*Physics*, 53(4):822–834, 2002.

T. N. van de Lindt, M. F. Fast, S. R. van Kranen, M. E. Nowee, E. M. Jansen, U. A. van der Heide, and J. J. Sonke. MRI-guided mid-position liver radiotherapy: Validation of image

processing and registration steps. *Radiotherapy and Oncology*, 138:132–140, 2019. doi: 10.1016/j.radonc.2019.06.007.

M. van Herk. Errors and margins in radiotherapy. *Seminars in radiation oncology*, 14(1): 52–64, 2004. doi: 10.1053/j.semradonc.2003.10.003.

J. Vandemeulebroucke, O. Bernard, S. Rit, J. Kybic, P. Clarysse, and D. Sarrut. Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT. *Medical physics*, 39(2):1006–1015, 2012.

C. Veiga, A. M. Loureno, S. Mouinuddin, M. van Herk, M. Modat, S. Ourselin, G. Royle, and J. R. McClelland. Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm. *Medical Physics*, 42(2):760–769, 2015. doi: 10.1118/1.4905050.

J. Weickert. *Anisotropic diffusion in image processing*, volume 1. Teubner Stuttgart, 1998.

C. L. Williams, P. Mishra, J. Seco, S. S. James, R. H. Mak, R. I. Berbeco, and J. H. Lewis. A mass-conserving 4D XCAT phantom for dose calculation and accumulation. *Medical Physics*, 2013. doi: 10.1118/1.4811102.

J. W. Wolthaus, J. J. Sonke, M. van Herk, J. S. Belderbos, M. M. Rossi, J. V. Lebesque, and E. M. Damen. Comparison of Different Strategies to Use Four-Dimensional Computed Tomography in Treatment Planning for Lung Cancer Patients. *International Journal of Radiation Oncology Biology Physics*, 70(4):1229–1238, 2008. doi: 10.1016/j.ijrobp.2007.11.042.