Improved PET/CT Respiratory Motion Compensation by Incorporating Changes in Lung Density

Elise Emond, Alexandre Bousse, Ludovica Brusaferri, Brian F. Hutton, Kris Thielemans

Improved PET/CT Respiratory Motion Compensation by Incorporating Changes in Lung Density

Elise C. Emond, *Student Member, IEEE,* Alexandre Bousse, Ludovica Brusaferri, *Student Member, IEEE,* Brian F. Hutton, *Senior Member, IEEE* and Kris Thielemans, *Senior Member, IEEE*

***Abstract*—Positron Emission Tomography/Computed Tomography (PET/CT) lung imaging is highly sensitive to motion. Although several techniques exist to diminish motion artifacts, few account for both tissue displacement and changes in density due to the compression and dilation of the lungs, which cause quantification errors. This paper presents an experimental framework for joint activity image reconstruction and motion estimation in PET/CT, where the PET image and the motion are directly estimated from the raw data. Direct motion estimation methods for motion-compensated PET/CT are preferable as they require a single attenuation map only and result in optimal signal to noise ratio (SNR). Previous implementations, however, failed to address changes in density during respiration. We propose to account for such changes using the Jacobian determinant of the deformation fields. In a feasibility study, we demonstrate on a modified extended cardiac-torso (XCAT) phantom with breathing motion—where the lung density and activity vary—that our approach achieved better quantification in the lungs than conventional PET/CT joint activity image reconstruction and motion estimation that does not account for density changes. The proposed method resulted in lower bias and variance in the activity images, reduced mean relative activity error in the lung at the reference gate (-4.84% to -3.22%) and more realistic Jacobian determinant values.**

***Index Terms*—PET/CT, Respiratory motion compensation, lung, density**

I. Introduction

PET lung imaging suffers from patient respiratory motion, affecting the image resolution, localization, and quantification due to the mismatch of the acquired PET data with the attenuation map, typically computed from a single snapshot CT acquisition. A common approach to compensate for respiratory motion, known as motion-compensated image reconstruction (MCIR), is to use image registration to compute deformation fields between different respiratory states (determined using a respiratory signal [1]), in order to warp the attenuation map and match the PET data, while performing iterative reconstruction. This has shown promising results, particularly for improved tumor delineation in lung imaging.

Commonly used MCIR techniques rely on previous determination of the deformation fields from gated images, which are then incorporated into the reconstruction [2]. A standard option relies on registering gated PET images reconstructed without attenuation correction [1] or, more recently, with attenuation correction [3] and accurate time-of-flight (TOF) information [4], leading to a reduction in localized image artifacts. A second option exploits CINE-CT images to obtain CT-based deformation fields, which should be more robust because of better contrast and higher resolution. However, this is contingent on having corresponding CT and PET gates for good results (therefore requiring amplitude-based respiratory gating). Also, because of the longer CT acquisition, the radiation dose is increased, which cannot be justified in some circumstances.

Instead of deformation fields determined from image registration, the deformation can also be estimated directly from the acquired PET data. Joint reconstruction and motion estimation (JRM) [5]–[9] is a reconstruction strategy in which, from the entire PET acquisition (as well as a gated respiratory surrogate signal and a single CT image), it is possible to estimate an activity image and the deformation between (respiratory) states. The input single CT image does not need to correspond to one of the gates of the respiratory signal; the computed activity image corresponds to the same respiratory state as the attenuation map (obtained from the CT image). The estimated deformation operator can warp both activity and attenuation images to each gate of the respiratory signal.

The previous methods however usually do not account for the change of density in the lungs, due to respiration. This could be a source of error, especially as both activity concentration and attenuation are affected. In good approximation, the changes can be estimated from local volume changes [10] (due to overall lung compression and dilation), although the lung mass is not entirely preserved during the respiration. Some fluid exchanges indeed contribute to an additional change of total lung tissue mass [11], which can be up to 10% over the respiratory cycle [12], where the lowest mass corresponds approximately to end-expiration and the largest to end-inspiration (likely due to the distension of blood vessels following the dilation of the lungs). This would also contribute to the variation of the lung density and tracer concentration during respiration. Accounting for density

Elise C. Emond was supported by GlaxoSmithKline (BIDS3000030921). Alexandre Bousse was funded by GE Healthcare. Ludovica Brusaferri was funded by Siemens Healthineers and the EPSRC-funded UCL Centre for Doctoral Training in Medical Imaging (EP/L016478/1).

Elise C. Emond, Ludovica Brusaferri, Brian F. Hutton and Kris Thielemans are with the Institute of Nuclear Medicine, University College London, London NW1 2BU, UK (contact: elise.emond.16@ucl.ac.uk).

The Institute of Nuclear Medicine is partially supported by the NIHR University College London Hospitals Biomedical Research Centre.

Alexandre Bousse was with the Institute of Nuclear Medicine, University College London, London NW1 2BU, UK. He is now with LaTIM, INSERM, UMR 1101, *Université de Bretagne Occidentale*, Brest, France.

changes during respiration might be particularly relevant when studying lung diseases such as fibrosis or tumors where the diseased region of the lung is more rigid than other parts of the lung. Therefore, we want to investigate here a method which would comprise a mass-preserving deformation model, which could help with both reducing data inconsistencies and estimating motion more robustly. In this work, the volume changes will be estimated by computing the "Jacobian determinant" (sometimes also referred as "Jacobian") of the deformation field [13]. Mass-preserving image registration, via incorporation of the Jacobian of the transformation field in the process, has been utilized before [14]–[17], but the idea has not been brought directly into a joint estimation of PET activity image and motion so far.

This paper aims to demonstrate the feasibility of a reconstruction algorithm for respiratory-gated PET data that can take local density changes into account, when only a single CT image is available. We provide an evaluation based on simulated PET data of XCAT phantom images.

In a first section, the theory of mass-preserving registration will be discussed, as well as the modifications necessary to the existing PET JRM to incorporate mass preservation. The method introduced will then be validated using simulated XCAT data of lung acquisition, where density and activity concentration changes are included within the lung.

II. Theory

This section presents a method built upon [7]–[9] and therefore similar notations will be used.

A. Motion-free PET Image Reconstruction

The activity and attenuation maps are respectively denoted $\lambda \in \mathcal{C}^+$ and $\mu \in \mathcal{C}^+$, where $\mathcal{C}^+ = \mathcal{C}^0(\mathbb{R}^3, \mathbb{R}^+)$ is the set of non-negative continuous functions defined on $\mathbb{R}^3$.

The set of detected prompts $g_i \in \mathbb{N}$ at a detection bin $i \in [\![1, n_\mathrm{b}]\!]$, where n_b is the number of detection bins (either defined from the corresponding detector pair in non TOF case or from the detector pair and the time bin in TOF case), follows a Poisson distribution of expectation $\bar{g}_i(\lambda, \mu)$ such that:

$$g_i \sim \mathrm{Poisson}(\bar{g}_i(\lambda, \mu)) \,, \tag{1}$$

with

$$\bar{g}_i(\lambda, \mu) = \tau a_i(\mu) H_i \lambda + b_i \,, \tag{2}$$

where τ is the duration of the acquisition, $a_i(\mu)$ is the attenuation along the line of response L_i corresponding to the detection bin i, H_i and a_i are defined as

$$H_i \lambda = \int_{\mathbb{R}^3} \lambda(r) h_i(r) \, \mathrm{d}r \tag{3}$$

and

$$a_i(\mu) = \exp\left(- \int_{L_i} \mu(r) \, \mathrm{d}r \right), \tag{4}$$

where $h_i : \mathbb{R}^3 \to \mathbb{R}^+$ is the system response function corresponding to the detection bin i, which is compactly supported such that (3) is well-defined, and b_i is the number of background events (either scatter or randoms) at the bin i.

The log-likelihood L, omitting the terms independent of λ and μ, is given by:

$$L(\lambda, \mu) = \sum_{i=1}^{n_\mathrm{b}} g_i \log \bar{g}_i(\lambda, \mu) - \bar{g}_i(\lambda, \mu) \,, \tag{5}$$

where n_b denotes the number of detection bins and $\log(\cdot)$ the natural logarithm. Assuming the attenuation μ is known (e.g., from CT reconstruction), the activity image λ can be reconstructed by maximizing (5), for example with the expectation-maximization (EM) algorithm [18].

B. Standard Approach

PET acquisitions are in practice longer than cardiac and respiratory cycles. However, if we assume the acquired PET data can be binned into several n_g "gates" such that $\boldsymbol{g} = \{\boldsymbol{g}_\ell\}_{\ell=1}^{n_\mathrm{g}}$ where each $\boldsymbol{g}_\ell \in \mathbb{N}^{n_\mathrm{b}}$ is a vector regrouping the counts corresponding to gate ℓ at each bin—during which intra-gate motion can be considered negligible in comparison to PET resolution. The previous model from Section II-A can then be used separately at each gate ℓ,

$$\forall \ell \in [\![1, n_\mathrm{g}]\!], \quad [\boldsymbol{g}_\ell]_i = g_{i,\ell} \sim \mathrm{Poisson}(\bar{g}_{i,\ell}(\lambda_\ell, \mu_\ell)) \,, \tag{6}$$

which depends on the activity distribution λ_ℓ and the attenuation distribution μ_ℓ at gate ℓ through the expectation

$$\bar{g}_{i,\ell}(\lambda_\ell, \mu_\ell) = \tau_\ell \, a_i(\mu_\ell) H_i \lambda_\ell + b_{i,\ell} \,, \tag{7}$$

where τ_ℓ corresponds to the duration of the binned respiratory gate ℓ and $b_{i,\ell}$ is the background term for the ℓ-th gate at the detection bin i. The corresponding log-likelihood Λ_ℓ to be maximized is:

$$\Lambda_\ell(\lambda_\ell, \mu_\ell) = \sum_{i=1}^{n_\mathrm{b}} g_{i,\ell} \log \bar{g}_{i,\ell}(\lambda_\ell, \mu_\ell) - \bar{g}_{i,\ell}(\lambda_\ell, \mu_\ell) \,. \tag{8}$$

One approach to tackle motion blur is therefore to reconstruct a single image λ_{ℓ_0} from the gated PET data $\boldsymbol{g}_{\ell_0}$ (8). However this approach suffers from low SNR as only a fraction of the measured counts are used. To address this, each gate ℓ can be reconstructed individually, followed by image registration to a single reference gate and averaging, but this approach requires gated attenuation maps (Cine-CT) which is not always available in practice; and increases patient dose. Alternatively, the motion between each of the gates and the reference gate can be estimated (e.g., from CT or magnetic resonance imaging (MRI)) and incorporated in the reconstruction framework; it is referred to as MCIR.

C. Direct Approach

In this section we derive an extended version of the model proposed in [7]–[9].

We define the warping operator $\mathcal{W}_\varphi : \mathcal{C}^+ \to \mathcal{C}^+$ associated to a diffeomorphism $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ such that:

$$\mathcal{W}_\varphi f = f \circ \varphi, \quad \text{where } f \in \mathcal{C}^+ \,. \tag{9}$$

In addition, another warping operator can be introduced, using the approximation that the mass is preserved in the lung and both the changes in activity concentration and density

during the respiration are directly correlated to the change in volume [10]. A modified "mass-preserving" warping operator $\widetilde{\mathcal{W}}_\varphi$ can therefore be defined, incorporating into the previous $\mathcal{W}_\varphi$ the absolute value of the determinant of the Jacobian matrix of φ in $\mathbb{R}^3$, denoted $|\det \mathcal{J}_\varphi|$, such that $\forall f \in \mathcal{C}^+$

$$\widetilde{\mathcal{W}}_\varphi f = |\det \mathcal{J}_\varphi| \cdot \mathcal{W}_\varphi f \,, \tag{10}$$

where the Jacobian matrix,

$$\begin{aligned} \mathcal{J}_\varphi : \mathbb{R}^3 &\rightarrow \mathcal{M}_{3,3}(\mathbb{R}) \\ (x,y,z) &\mapsto \left(\frac{\partial \varphi}{\partial x}(r), \frac{\partial \varphi}{\partial y}(r), \frac{\partial \varphi}{\partial z}(r)\right), \end{aligned} \tag{11}$$

is differentiable.

The definition of this operator is motivated by the consideration of local mass preservation: if we consider a diffeomorphic deformation φ which transforms a non-negative continuous distribution f_1 into f_2 (i.e., $f_2 = f_1 \circ \varphi$), then:

$$\int_{\mathbb{R}^3} f_1(r)\,\mathrm{d}r = \int_{\mathbb{R}^3} |\det(\mathcal{J}_\varphi(r))|\, f_2(r)\,\mathrm{d}r \,, \tag{12}$$

where $\det(\mathcal{J}_\varphi(r))$ consequently reflects the volume changes.

At each gate $\ell \in [\![1, n_\mathrm{g}]\!]$, the corresponding motion is φ_ℓ. The mass-preserving motion-compensated maximum-likelihood is defined by replacing λ_ℓ and μ_ℓ by the deformed reference images λ and μ in (8) then summing over the gates:

$$L(\lambda, \{\varphi_\ell\}_{\ell=1}^{n_\mathrm{g}}, \mu) = \sum_{\ell=1}^{n_\mathrm{g}} \Lambda_\ell(\widetilde{\mathcal{W}}_{\varphi_l}\lambda, \widetilde{\mathcal{W}}_{\varphi_\ell}\mu)) \,. \tag{13}$$

The addition of mass-preserving conditions within a cost function (not necessarily using $|\mathcal{J}_\varphi|$) has been used by various authors, for CT [15], PET [16], [19], single photon emission computed tomography (SPECT) [20], [21] and MRI data [22], mostly for image registration. However, mass preservation has never been used, to the best of the authors' knowledge, directly within a motion-compensating PET image reconstruction scheme, using a log-likelihood cost function to estimate the motion between respiratory gates.

D. Discretization

In this work, the discretization is achieved following the scheme proposed in [5], [7]. Vectors derived from discretization appear in bold. The discretized images corresponding to λ and μ are $\boldsymbol{\lambda} \in \mathbb{R}^{n_\mathrm{v}}$ and $\boldsymbol{\mu} \in \mathbb{R}^{n_\mathrm{v}}$ respectively, where n_v is the number of voxels in the image, and $f_j = [\boldsymbol{f}]_j$ and $\mu_j = [\boldsymbol{\mu}]_j$ respectively correspond to the activity and attenuation at voxel $j \in [\![1, n_\mathrm{v}]\!]$.

We use B-spline functions for the deformation field [5], [7], [23]. The discrete motion is expressed with an n_c control point grid $\{r_n\}_{n=1}^{n_\mathrm{c}}$, $r_n \in \mathbb{R}^3$ and with a B-spline coefficient vector $\boldsymbol{\alpha} = \left(\boldsymbol{\alpha}^X = \left\{\alpha_n^X\right\}_{n=1}^{n_\mathrm{c}}, \boldsymbol{\alpha}^Y = \left\{\alpha_n^Y\right\}_{n=1}^{n_\mathrm{c}}, \boldsymbol{\alpha}^Z = \left\{\alpha_n^Z\right\}_{n=1}^{n_\mathrm{c}}\right)$, where $\alpha_n^C \in \mathbb{R}$ is the cubic B-spline coefficient in direction $C \in \{X, Y, Z\}$ and at control point r_n. The corresponding motion φ_α is defined at each location $r \in \mathbb{R}^3$ as on a subgrid of the original image grid comprising n_c control points of coordinates $\{r_n\}_{n=1}^{n_\mathrm{c}}$, such that:

$$\varphi_{\boldsymbol{\alpha}}(r) = r + \begin{pmatrix} \sum_{n=1}^{n_\mathrm{c}} \alpha_n^X \mathcal{B}\left(\frac{r-r_n}{D}\right) \\ \sum_{n=1}^{n_\mathrm{c}} \alpha_n^Y \mathcal{B}\left(\frac{r-r_n}{D}\right) \\ \sum_{n=1}^{n_\mathrm{c}} \alpha_n^Z \mathcal{B}\left(\frac{r-r_n}{D}\right) \end{pmatrix} \tag{14}$$

where D is the distance between the control points, r_n is the n-th control point and $\mathcal{B}(r) = b(x) \cdot b(y) \cdot b(z)$ where $r = (x, y, z)$ are the coordinates of a point in the grid and b is the (uniform) cubic B-spline function. We assume here that the control points are equally spaced in x, y and z directions but the implementation could be extended to uneven spacing.

The discrete version of the warping operator $\mathcal{W}_\varphi$ becomes a matrix $\boldsymbol{\mathcal{W}}_{\boldsymbol{\alpha}} \in \mathcal{M}_{n_\mathrm{v}, n_\mathrm{v}}(\mathbb{R})$. Additionally, we define the "discrete" determinant $\boldsymbol{\Delta}_{\boldsymbol{\alpha}} \in \mathbb{R}^{n_\mathrm{v}}$ by sampling $\det \mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)$ at each voxel.

We redefine the expected number of counts at a gate ℓ such as $\bar{g}_{i,\ell}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\mu}) = \tau_\ell\, a_i(\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\mu}) \left[\boldsymbol{H}\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\lambda}\right]_i + s_{i,\ell}$, where $\boldsymbol{H}$ is the discretized PET system matrix and $a_i(\boldsymbol{\mu})$ is the attenuation factor corresponding to the discretized attenuation map $\boldsymbol{\mu}$.

Finally, the discrete version of (13) is

$$L(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\mu}) = \sum_{\ell=1}^{n_\mathrm{g}} \Lambda_\ell(\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\lambda}, \widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\mu}) \,, \tag{15}$$

where $\boldsymbol{\theta} = \{\boldsymbol{\alpha}_\ell\}_{\ell=1}^{n_\mathrm{g}}$ and $\boldsymbol{\alpha}_\ell$ is the motion parameter at gate ℓ.

E. Joint Image Reconstruction and Motion Estimation

We propose to estimate the activity image $\boldsymbol{\lambda}$ and the motion parameter ensemble $\boldsymbol{\theta}$ by maximizing (15) with the addition of regularization terms on $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$

$$\Phi(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\mu}) \triangleq L(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\mu}) + \beta U(\boldsymbol{\lambda}) + \gamma V(\boldsymbol{\theta}) \,, \tag{16}$$

where β and γ are the penalty weights associated to the (quadratic) penalty terms $U(\boldsymbol{\lambda})$ and $V(\boldsymbol{\theta})$, respectively. Thus, the discussed JRM tries to solve the following optimization problem to estimate an activity image $\widehat{\boldsymbol{\lambda}}$ and a set of deformation parameters $\widehat{\boldsymbol{\theta}}$:

$$(\widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{\theta}}) \in \underset{\boldsymbol{\lambda},\, \boldsymbol{\theta}}{\arg\max}\, \Phi(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\mu}) \,. \tag{17}$$

The proposed method proceeds by alternating between $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$ updates, using block sequential regularized expectation-maximization (BSREM) algorithm proposed by Ahn *et al.* [24] for $\boldsymbol{\lambda}$—which will not be detailed here—and limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) [25] for $\boldsymbol{\theta}$—for which we propose a summary in Appendix B. When a mass-preserving model is used, it will be referred as "mass-preserving JRM", as compared to "standard JRM".

With the mass-preserving case, the derivation of the gradient of the cost function in the motion update needs to account for the additional dependence in $\boldsymbol{\theta}$ due to the addition of Jacobian determinant in the model. The gradient in $\boldsymbol{\theta}$ is somehow similar to that of [7] with the addition of the analytical derivatives of $|\boldsymbol{\Delta}_{\boldsymbol{\alpha}_\ell}|$ with respect to the B-spline coefficients $\boldsymbol{\alpha}_\ell$ which are obtained in a a similar way as in [15]. Details on derivation of the cost function are given in Appendix A and B.

F. Image & Motion Estimation Regularization

1) Activity Image regularization: In standard JRM [7], a quadratic prior was used to regularize image reconstruction. In this work, $U(\boldsymbol{\lambda})$ is the relative difference prior [26] and can be expressed as

$$U(\boldsymbol{\lambda}) = -\sum_{j=1}^{n_\mathrm{v}} \sum_{v \in \mathcal{N}_j} \omega_{j,v} \frac{(\lambda_j - \lambda_v)^2}{(\lambda_j + \lambda_v) + \gamma|\lambda_j - \lambda_v|}\,, \quad (18)$$

where $\omega_{j,v}$ in the inverse distance between the center of a voxel j and the center of a voxel v and $\mathcal{N}_j$ is the neighbourhood of voxel j.

2) Deformation regularization: $V(\boldsymbol{\theta})$ in (16) can be expressed as a sum of penalties over all the gates $\ell \in [\![1, n_\mathrm{g}]\!]$:

$$V(\boldsymbol{\theta}) = \sum_{\ell=1}^{n_\mathrm{g}} \mathcal{R}(\boldsymbol{\alpha}_\ell) \quad (19)$$

where $\mathcal{R}$ is a penalty on the coefficients $\{\boldsymbol{\alpha}_\ell\}_{\ell=1}^{n_\mathrm{g}}$. In this work, instead of only regularizing the deformation parameters, the regularization is also performed on the Jacobian determinant image. This should prevent the optimization in $\boldsymbol{\theta}$ to favor a mass-preserving motion that matches with the noise—as pointed out in [19] for mass-preserving registration—by constraining the Jacobian determinant image to be smooth. The smoothed total variation (STV) penalty [27] was chosen here:

$$\mathcal{R}_{\mathrm{STV}}(\boldsymbol{\alpha}_\ell) = -\sum_{j=1}^{n_\mathrm{v}} \sqrt{\|[\boldsymbol{F}|\,\boldsymbol{\Delta}_{\boldsymbol{\alpha}_\ell}\,|]_j\|_2^2 + \zeta^2}, \quad \zeta > 0 \quad (20)$$

where $\boldsymbol{F}$ is the finite forward difference operator used to approximate the gradient of an image and ζ (chosen as 0.3) is used as a smoothing factor, also enabling differentiation in 0. To ensure smooth deformation fields, a small regularization on the B-spline coefficients was also added within $V(\boldsymbol{\theta})$:

$$\mathcal{R}_{\mathrm{def}}(\boldsymbol{\alpha}_\ell) = -\frac{1}{2} \sum_{n=1}^{n_\mathrm{c}} \sum_{m \in \widetilde{\mathcal{N}}_n} \widetilde{\omega}_{n,m} \sum_{C \in \{X,Y,Z\}} \left(\alpha_{\ell,n}^C - \alpha_{\ell,m}^C\right)^2\,, \quad (21)$$

where $\widetilde{\omega}_{n,m}$ is the inverse distance between the control points n (i.e., 1, $\sqrt{2}$ or $\sqrt{3}$) and m, $\widetilde{\mathcal{N}}_n$ is the neighborhood of a control point n (within the n_c control points grid). The size of the neighborhoods $\widetilde{\mathcal{N}}_n$ was set to $3 \times 3 \times 3$. The values at the image borders were handled by adding padding to the image, using the nearest values.

More details are given in Appendix B.

III. Experiments

Five sets of lung XCAT images were generated, corresponding to five different respiratory gates ($n_\mathrm{g} = 5$). The corresponding activity and attenuation images are denoted $\{\boldsymbol{\lambda}_\ell\}_{\ell=1}^{n_\mathrm{g}}$ and $\{\boldsymbol{\mu}_\ell\}_{\ell=1}^{n_\mathrm{g}}$, respectively (where $\ell = 1$ is the end-inspiration and $\ell = 5$ is the end-expiration), mimicking displacement gating. The values in the activity images correspond to a standard FDG acquisition 60-min post-injection acquired in our imaging facility.

Additionally, to simulate density changes in lungs obeying mass preservation, the activity distribution and density values were varied uniformly in the lung, depending on the total volume change compared to mid-expiration, using XCAT gated lung masks. As a result, compared to the mid-expiration state (gate 3), the changes in density and activity concentration were respectively as follows: gate 1: -11%, gate 2: -6.9%, gate 3: 0%, gate 4: $+8.2\%$, gate 5: $+11.5\%$. It should however be noted that these changes do not correspond to the Jacobian determinant of the true deformation fields from XCAT, therefore the simulations include some model mismatch as would be the case for patient data. For healthy patients, however, we do not expect a uniform change in density and activity concentration throughout the lung, as some regions are expected to ventilate more than others as documented in [12], [28]. The resulting images are shown in Figures 1 and 2.

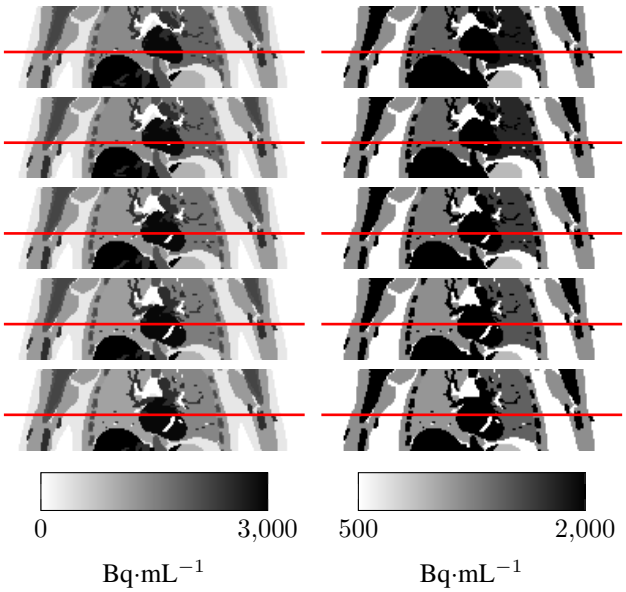


Fig. 1: Modified XCAT activity images at each respiratory gate used for the simulation: original scale (left) and reduced scale (right) for better visualization of the lung activity concentration change, from end-inspiration (top) to end-expiration (bottom); a red horizontal line was plotted to visualize respiratory motion.

After the activity and activity images were created, the gated projection data were obtained via forward projection of the activity images and separate attenuation correction factors were computed for each gate. Scatters and randoms were added with a level similar to a usual lung acquisition (about 60% of the total number of detected counts). Poisson noise was added to the projection data in the results presented (number of prompts $\approx 3 \cdot 10^7$), to simulate a standard FDG lung acquisition in our imaging facility. A GE Discovery 710 scanner geometry was used to simulate the data (using GE proprietary projectors in a MATLAB package). For all gates ℓ, the background sinogram $\{s_{i,\ell}\}_{i=1}^{n_\mathrm{b}}$ in (7) is considered to be known accurately during the entire reconstruction.

The input attenuation map corresponds to the end-expiration state $\boldsymbol{\mu}_5$, which means that the first gated activity image reconstructed in the algorithm is perfectly aligned with the attenuation map. The number of JRM iterations $\mathrm{numJRMIter}$

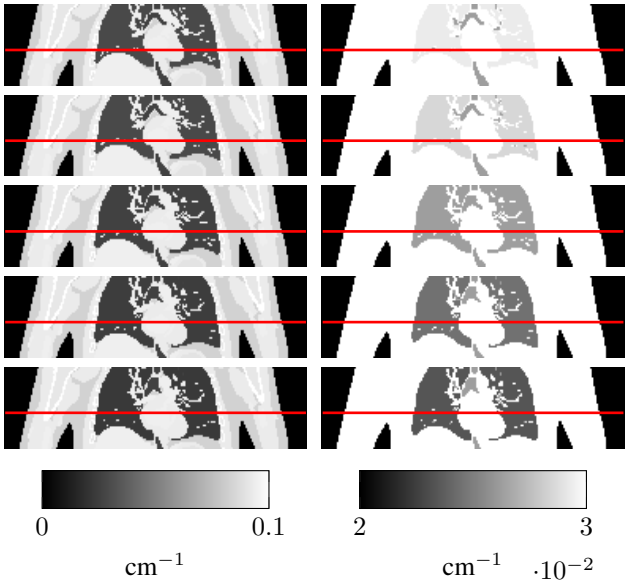


Fig. 2: Modified XCAT attenuation images at each respiratory gate used for the simulation: original scale (left) and reduced scale (right) for better visualization of the lung density change, from end-inspiration (top) to end-expiration (right); a red horizontal line was plotted to visualize respiratory motion.

was set to 3 (introduced in Appendix B A.).

In addition to the mass-preserving JRM presented here, the "standard" JRM (i.e., without mass preservation, as in [7]) was used to reconstruct the data. A range of regularization values γ for the motion update (16) was tested, from weak to strong visual smoothing of the Jacobian determinant images. The image reconstruction parameters were kept identical for all simulations (8 subsets). For each model and penalty weight, 30 different Poisson noise realizations were performed.

The image variance and squared bias in all attenuation images and only in the lung were studied to select the best regularization configuration for each model and to compare the two models. The attenuation image squared bias is given as:

$$\text{Bias}^2_{\boldsymbol{\mu}} = \frac{1}{n_\text{g}} \sum_{\ell=1}^{n_\text{g}} \frac{1}{|\mathcal{V}_\ell|} \sum_{j \in \mathcal{V}_\ell} \left(\frac{1}{K} \sum_{\kappa=1}^{K} \left[\boldsymbol{\mu}_\ell^{[\kappa]} \right]_j - [\boldsymbol{\mu}_\ell]_j \right)^2 , \tag{22}$$

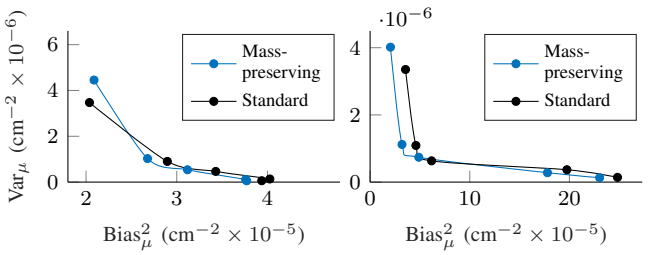
and the attenuation image variance as:

$$\text{Var}_{\boldsymbol{\mu}} = \frac{1}{n_\text{g}} \frac{1}{K-1} \sum_{\ell=1}^{n_\text{g}} \frac{1}{|\mathcal{V}_\ell|} \sum_{j \in \mathcal{V}_\ell} \sum_{\kappa=1}^{K} \left(\left[\boldsymbol{\mu}_\ell^{[\kappa]} \right]_j - m_{j,\ell}^{\text{atn}} \right)^2 , \tag{23}$$

where

$$m_{j,\ell}^{\text{atn}} = \frac{1}{K} \sum_{\kappa=1}^{K} \left[\boldsymbol{\mu}_\ell^{[\kappa]} \right]_j , \tag{24}$$

$K = 30$, $\boldsymbol{\mu}_\ell^{[\kappa]} = \dot{\boldsymbol{\mathcal{W}}}_{\hat{\boldsymbol{\alpha}}_\ell^{[\kappa]}} \boldsymbol{\mu}_1$ is the estimated μ image at a respiratory gate ℓ and at the κ-th noise realisation, $\dot{\boldsymbol{\mathcal{W}}}$ denoting either the mass-preserving JRM warping operator (i.e., $\dot{\boldsymbol{\mathcal{W}}} = \widetilde{\boldsymbol{\mathcal{W}}}$) or standard JRM warping operator (i.e., $\dot{\boldsymbol{\mathcal{W}}} = \boldsymbol{\mathcal{W}}$).



(a) Entire image (b) Lung only

Fig. 3: Tradeoff between variance and the (squared) bias in the μ images, in (a) the entire image and in (b) the lung only, between the two JRM models, using an edge-preserving JRM regularization.

Similarly, we define the image squared bias and variance in the activity images $\{\boldsymbol{\lambda}_\ell\}_{\ell=1}^{n_\text{g}}$ as follows:

$$\text{Bias}^2_{\boldsymbol{\lambda}} = \frac{1}{n_\text{g}} \sum_{\ell=1}^{n_\text{g}} \frac{1}{|\mathcal{V}_\ell|} \sum_{j \in \mathcal{V}_\ell} \left(\frac{1}{K} \sum_{\kappa=1}^{K} \left[\boldsymbol{\lambda}_\ell^{[\kappa]} \right]_j - [\boldsymbol{\lambda}_\ell]_j \right)^2 , \tag{25}$$

and

$$\text{Var}_{\boldsymbol{\lambda}} = \frac{1}{n_\text{g}} \frac{1}{K-1} \sum_{\ell=1}^{n_\text{g}} \frac{1}{|\mathcal{V}_\ell|} \sum_{j \in \mathcal{V}_\ell} \sum_{\kappa=1}^{K} \left(\left[\boldsymbol{\lambda}_\ell^{[\kappa]} \right]_j - m_{j,\ell}^{\text{act}} \right)^2 , \tag{26}$$

where

$$m_{j,\ell}^{\text{act}} = \frac{1}{K} \sum_{\kappa=1}^{K} \left[\boldsymbol{\lambda}_\ell^{[\kappa]} \right]_j , \tag{27}$$

$\boldsymbol{\lambda}_\ell^{[\kappa]} = \dot{\boldsymbol{\mathcal{W}}}_{\hat{\boldsymbol{\alpha}}_\ell^{[\kappa]}} \boldsymbol{\lambda}^{[\kappa]}$ is the estimated activity image at a respiratory gate ℓ and at the κ-th noise realization and $\boldsymbol{\lambda}^{[\kappa]}$ is the MCIR activity image.

These measures of variances and squared biases were used to assess the best regularization configuration (using a scatter plot), for each model.

We also define the mean relative error activity image $\mathbf{RD}_\ell \in \mathbb{R}^{n_\text{v}}$ at a gate ℓ such that at a voxel i:

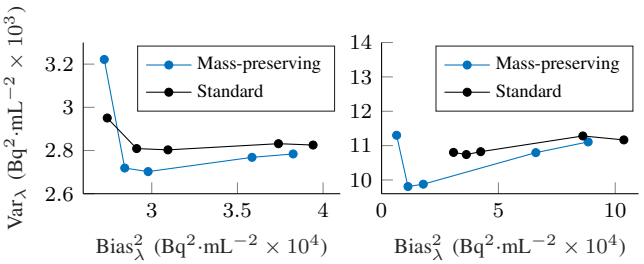
$$[\mathbf{RD}_\ell]_i = \frac{m_j(\boldsymbol{\lambda}_\ell) - [\boldsymbol{\lambda}_\ell]_i}{[\boldsymbol{\lambda}_\ell]_i} . \tag{28}$$

All measures in the lung were made using the respiratory-dependent masks from XCAT, which were eroded using a small kernel.

IV. Results

The scatter plots for the bias-variance study for the attenuation map are given in Figure 3 and for the activity image in Figure 4. The results show that using a mass-preserving model in JRM reduces the image squared biases, especially in the lung region of the activity image. The activity image variances are also decreased.

From the previous bias-variance analysis, we choose to compare the following results using a medium regularization. Reconstructed activity images corresponding to the two different models on $\boldsymbol{\theta}$, are shown alongside the true activity



(a) Entire image (b) Lung only

Fig. 4: Tradeoff between variance and the (squared) bias in the λ images, in (a) the entire image and in (b) the lung only, between the two JRM models, using an edge-preserving JRM regularization.

(a) $\boldsymbol{\lambda}_{\mathrm{Std}}$ (b) $\boldsymbol{\lambda}_{\mathrm{MP}}$ (c) $\boldsymbol{\lambda}^\star$

Fig. 5: Coronal views of (c) the true activity image $\boldsymbol{\lambda}^\star$ and the motion-compensated images using either (a) the standard model for $\boldsymbol{\lambda}_{\mathrm{Std}}$ or (b) the mass-preserving model for $\boldsymbol{\lambda}_{\mathrm{MP}}$.

image in Figure 5 (at the reference gate, i.e., end-expiration). Visually, the differences between the two reconstructed images are not conspicuous. The mean activity $\pm$ standard deviation in the lung (using a slightly eroded lung mask given by XCAT at end-expiration), averaged over all noise simulations, was found to be $1559.6 \pm 99.7\,\mathrm{Bq.mL^{-1}}$ in mass-preserving JRM and $1533.1 \pm 98.7\,\mathrm{Bq.mL^{-1}}$ in standard JRM, where the true activity concentration is of $1612.1\,\mathrm{Bq.mL^{-1}}$. The mean relative error in the lung mask at end-expiration is -3.22% for mass-preserving JRM and -4.84% for standard JRM. Mass-preserving JRM therefore performs better than standard JRM, although the difference is small.

When we compute the mean relative error images $\{\mathbf{RD}_\ell\}_{\ell=1}^{n_g}$ for both JRM versions, we can see that including mass preservation in the imaging model decreases the error in the lung and outside of the lung over the different gates. The images are shown in Figure 6.

Finally, we want to assess how closely the estimated Jacobian determinant images correspond to the simulated activity concentration and density change ratio images in Figure 7. The values in the liver and in the cardiac region are improved for the mass-preserving model,as they are closer to 1 (this value corresponds to the simulated density and activity concentration change ratio between all gates and end-expiration).

V. Discussion

The results presented in this work show that incorporating a mass-preserving condition within a reconstruction method, such as joint reconstruction and motion estimation (JRM), can reduce quantification errors, especially in the lungs. A direct estimation of the activity image and of the deformation fields between gates can also help with estimating motion more

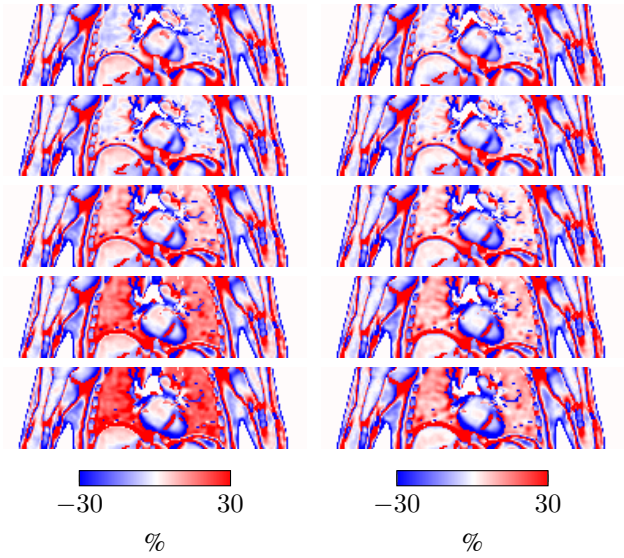


Fig. 6: Mean relative activity error image at each gate for standard JRM (left) and mass-preserving JRM (right), from end-expiration (top) to end-inspiration (bottom).

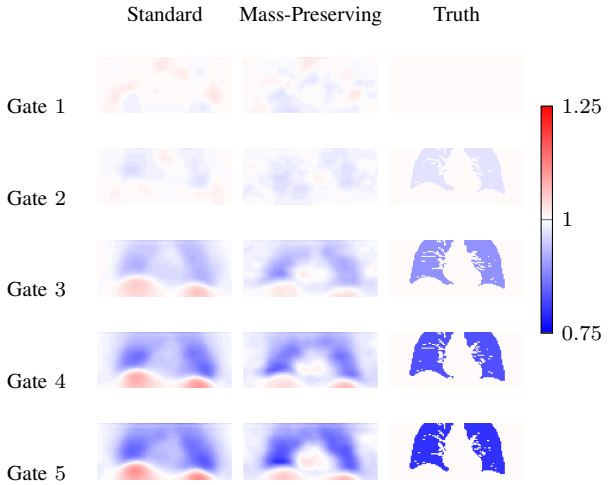


Fig. 7: Coronal views of the Jacobian determinant images (unitless), at each respiratory gate, computed from the deformation fields estimated using either the mass-preserving model (1st column) or the standard model (2nd column), compared to the true activity concentration and density change ratios (3rd column).

accurately, especially internally in the lung, while needing only one CT acquisition and respiratory-gated PET data.

A bias-variance study over all gates was first done, from a range of penalty weights for the regularization of the deformation parameters. The scatter plots showed that both bias and variance reduce in the gate-dependent activity images (i.e., after using the deformation operator applied to the reconstructed image) when the mass-preserving model is used— both inside and outside the lungs (e.g., in the cardiac region or in the liver). Additionally, the mean relative error images for all different gates were computed. The errors in the lung decreased, especially for the gates which were the most

different to the reference end-expiration gate. In particular, for medium regularization on the deformation parameters, the mean relative error in the lung of the reconstructed activity image at end-expiration decreased from -4.84% to -3.22%. The Jacobian determinant images obtained from the deformation fields were also more consistent with the simulated density and activity concentration changes, with values closer to 1 outside of the lungs.

Although we do not show large benefits when the mass-preserving method is applied to the XCAT phantom, our approach could have a larger impact for chest acquisitions of patient data. A possible domain of application could be data of patients suffering from diffuse lung diseases, such as idiopathic pulmonary fibrosis or chronic obstructive pulmonary disease; since the lung rigidity varies locally because of the pathology, lung internal motion could be more precisely estimated by considering mass preservation.

A drawback for joint reconstruction in PET is however linked to rather slow convergence and therefore increased computational time (in our settings, using an implementation based on CPU projectors without multi-threading, it took about 11 times longer to process the data with JRM as compared to standard MCIR, i.e., approximately two days), especially when the activity image used to initialize the algorithm has quantification issues (due to misalignment with the μ map). As suggested in [7], the initialization of the activity image at the start of the alternating strategy is important for both algorithm acceleration and convergence (as the likelihood is not concave, it is necessary to start as close as possible to the true activity image to ensure that it converges). Therefore, it is recommended to initialize the activity image using a gated reconstruction corresponding to the sinogram of the respiratory gate which is the closest to the gate of the input μ map. Additionally, estimating coarse deformation fields between all gates should be beneficial to accelerate the method presented and could also be used for a less noisy activity image initialization using MCIR. To do so, standard methods such as registering non attenuation-corrected gated images could be used [1]. Another alternative would be the use of a reconstruction of type Maximum-Likelihood reconstruction of Activity and Attenuation (MLAA) [29], for which the attenuation sinogram can be estimated up to a constant [30] in TOF PET (within the activity hull). When applied to respiratory-gated data, although MLAA μ maps are expected to be very noisy, recent work on gated data have shown that it could be possible to extract information on motion [31], when deep learning is used to denoise the images [32].

Moreover, a parameterization of the deformation field using cubic B-splines was used in this work. Although B-splines have the advantage of being simple for derivation and computational reasons, other deformation parameterization could be more adapted for lung registration, for example to handle sliding motion against the ribs. Possible approaches could be to extend supervoxel techniques [33], [34] or to use "fluid material" deformation fields, such as introduced by Christensen *et al.* [35].

Finally, the method presented here could be extended to estimate the values in the μ map jointly with the deformation fields and the activity image. Using the additional constraint linked to the modulation of the attenuation and the activity concentration changes with the Jacobian determinant, it may be possible to estimate the attenuation sinograms accurately, determining intrinsically the additive attenuation sinogram [30] linked to standard MLAA TOF reconstruction.

VI. Conclusion

This study shows the feasibility of a mass-preserving joint activity image reconstruction and motion estimation in PET/CT, which could be also extended to PET/MR imaging. Future work will involve validation on patient data.

Appendix A: Derivation of the Determinant of the Deformation Jacobian Matrix for B-Spline Representation

The element (k,l) of the 3×3 matrix $\mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)$ is given by:

$$[\mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)]_{k,l} = \delta_{k,l} + \sum_{n=1}^{n_c} \frac{\alpha_n^{C_k}}{D} \frac{\partial \mathcal{B}}{\partial x_l}\left(\frac{r-r_n}{D}\right) \tag{29}$$

where $\delta_{k,l}$ is the Kronecker delta for k,l and $\frac{\partial \mathcal{B}}{\partial x_l}$ is the partial derivative of $\mathcal{B}$ with respect to x_l, C_k corresponding to X, Y, Z for $k=1$, $k=2$, $k=3$, respectively, and x_l corresponding to x, y or z for $l=1$, $l=2$, $l=3$ respectively. The derivatives of a determinant of a matrix can be obtained via Jacobi's formula, with respect to a given $\alpha_n^{C_k}$:

$$\frac{\partial \det \mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)}{\partial \alpha_n^{C_k}}(\boldsymbol{\alpha}) = \operatorname{tr}\left(\operatorname{adj}\left(\mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)\right) \frac{\partial \mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)}{\partial \alpha_n^{C_k}}(\boldsymbol{\alpha})\right) \tag{30}$$

where $\operatorname{tr}(\cdot)$ is the trace of the matrix and $\operatorname{adj}(\cdot)$ is its adjugate. For a given r, the partial derivatives of $\mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)$ with respect to $\alpha_n^{C_k}$ are equal to:

$$\frac{\partial \mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)}{\partial \alpha_n^{C_k}}(\boldsymbol{\alpha}) = \frac{1}{D}\sum_{l=1}^{3} \frac{\partial \mathcal{B}}{\partial x_l}\left(\frac{r-r_n}{D}\right) . \tag{31}$$

By the chain rule, the partial derivative of $|\det \mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)|$ is therefore:

$$\frac{\partial\, |\det \mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)|}{\partial \alpha_n^{C_k}} = \operatorname{sgn}(\det \mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)) \cdot \operatorname{tr}\left(\operatorname{adj}\left(\mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)\right) \cdot \frac{\partial \mathcal{J}_{\varphi_{\boldsymbol{\alpha}}}(r)}{\partial \alpha_n^{C_k}}\right) , \tag{32}$$

where $\operatorname{sgn}(\cdot)$ is the sign function. The non-differentiability of $|\cdot|$ function in 0 is handled by adding a small number to the Jacobian determinant. The analytical derivatives of $|\det(\mathcal{J}_{\varphi_{\boldsymbol{\alpha}}})|$ were validated against finite differences in terms of absolute percentage error between the gradient and the finite differences, normalized with respect to the maximum absolute value of the gradient, for random B-spline coefficients and $\epsilon = 0.00001$ used in finite differences. The mean and maximum errors were found to be less than $3\cdot 10^{-11}$ and $4\cdot 10^{-10}$ respectively.

APPENDIX B: IMPLEMENTATION AND DERIVATION DETAILS

A. Joint Reconstruction Outline

Joint reconstruction of activity image/motion estimation consists of $\mathrm{numJRMIter}$ alternations between:

- *Motion Update*: maximizing the log-likelihood, from a given activity image, with respect to the B-spline coefficients $\boldsymbol{\theta}$ to estimate the deformation between the reference respiratory state (corresponding to the CT image used to compute the input attenuation map) and the PET respiratory gates (binned using a respiratory surrogate signal). The optimization is performed for each respiratory gate separately.
- *Regularized Image Reconstruction*: given the deformation fields estimated during the previous step, MCIR is used to obtain $\boldsymbol{\lambda}$, with a modified version of BSREM. Contrary to [7], the image is always reinitialized to $\mathbf{1}_{n_\mathrm{v}}$ before each reconstruction.

The algorithm is initialized from a gated regularized image reconstruction at end-expiration (according to the respiratory trace). The first motion estimation is chosen to have more iterations than following ones (here twice as many). A concise pseudo-code is given in Algorithm 1 and some additional details on the motion update and image update are provided in following subsections B and C. A more detailed summary of the implementation used in this work can be found in [8].

ALGORITHM 1
Joint Reconstruction of Motion and Activity

Input: μ, gated projection data $\boldsymbol{g}$
$\boldsymbol{\theta}^{(0)} \leftarrow \mathbf{0}$
$\boldsymbol{\lambda}^{(0)} \leftarrow$ BSREM($\boldsymbol{g}_1, \boldsymbol{\mu}, \beta$)
for $r = 1, \ldots, \mathrm{numJRMIter}$ **do**
  $\boldsymbol{\theta}^{(r)} \leftarrow$ Motion estimation from $(\boldsymbol{g}, \mu, \boldsymbol{\theta}^{(r-1)}, \boldsymbol{\lambda}^{(r-1)}, \gamma)$
  $\boldsymbol{\lambda}^{(r)} \leftarrow$ MC-BSREM($\boldsymbol{g}, \mu, \boldsymbol{\theta}^{(r)}, \beta$)
end for
Output: estimated PET image $\widehat{\boldsymbol{\lambda}}$, B-spline coefficients $\widehat{\boldsymbol{\theta}}$

B. Motion Update

The penalized log-likelihood $\Phi(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\mu})$ is maximized with respect to $\boldsymbol{\theta}$ during the motion update. Let $\boldsymbol{\theta}^{(k)}$ be a current estimate of $\boldsymbol{\theta}$ at iteration k. $\boldsymbol{\theta}^{(k+1)}$ is obtained by performing a maximization along a gradient ascent search direction $\boldsymbol{t}^{(k)}$:

$$\boldsymbol{\alpha}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \delta^{(k)} \boldsymbol{t}^{(k)} \tag{33}$$

where

$$\delta^{(k)} = \underset{\delta \geq 0}{\arg\max}\, \Phi(\boldsymbol{\lambda}, \boldsymbol{\theta}^{(k)} + \delta \boldsymbol{t}^{(k)}, \boldsymbol{\mu}) \tag{34}$$

and $\boldsymbol{t}^{(k)}$ is a gradient ascent direction that incorporates the inverse Hessian of Φ, computed with a L-BFGS quasi-Newton line-search algorithm, handling bound constraints [25]; although any line-search method could be used. A Fortran implementation [37] was used to compute $\boldsymbol{t}$ and $\delta^\star$ at each motion estimation iteration. The step length (34) satisfies the Wolfe conditions [38], i.e., guarantees sufficient increase of (16).

To simplify the notation, dependencies on $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ will be omitted in this appendix. For each gate ℓ, the expected number of counts at bin i $\bar{g}_{i,\ell}(\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\lambda}, \widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\mu})$, are regrouped in a vector $\bar{\boldsymbol{g}}_\ell(\boldsymbol{\alpha}_\ell) \in \mathbb{R}^{n_\mathrm{b}}$. The likelihood $L(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\mu})$ will be denoted $L(\boldsymbol{\theta})$ and $\boldsymbol{J}(\cdot)$ will represent the Jacobian matrix in $\boldsymbol{\alpha}_\ell$.

Using the chain rule, the gradient of the log-likelihood in $\boldsymbol{\alpha}_\ell$ is:

$$\nabla_{\boldsymbol{\alpha}_\ell} L(\boldsymbol{\theta}) = \boldsymbol{J}(\bar{\boldsymbol{g}}_\ell(\boldsymbol{\alpha}_\ell))^\top \nabla_\ell(\bar{\boldsymbol{g}}_\ell(\boldsymbol{\alpha}_\ell)) \tag{35}$$

where $\nabla_\ell(\bar{\boldsymbol{g}}_\ell(\boldsymbol{\alpha}_\ell)) = \boldsymbol{g}_\ell / \bar{\boldsymbol{g}}_\ell(\boldsymbol{\alpha}_\ell) - \mathbf{1}_{\mathbb{R}^{n_\mathrm{c}}}$.

The latter is modified with the incorporation of the mass-preserving warping operator $\widetilde{\boldsymbol{\mathcal{W}}}$ instead of the regular warping operator $\boldsymbol{\mathcal{W}}$ used in [7] (cf Appendix A):

$$\begin{aligned}\boldsymbol{J}(\bar{\boldsymbol{g}}_\ell(\boldsymbol{\alpha}_\ell)) = &-\tau_\ell \,\mathrm{diag}\{\boldsymbol{H}_\mathrm{a}(\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\mu})\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\lambda}\}\boldsymbol{\mathcal{L}}\boldsymbol{J}(\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\mu}) \\ &+ \tau_\ell \boldsymbol{H}_\mathrm{a}(\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\mu})\boldsymbol{J}(\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{\lambda})\end{aligned}$$

where $\boldsymbol{\mathcal{L}} \in \mathcal{M}_{n_\mathrm{b}, n_\mathrm{v}}(\mathbb{R})$ represents the line integral matrix where an element (i, j) is the length of intersection of the line connecting the two detectors corresponding to the bin i with voxel j, $\boldsymbol{H}_\mathrm{a}(\boldsymbol{\mu}) = \mathrm{diag}\left\{\exp\left(-\boldsymbol{\mathcal{L}}\boldsymbol{\mu}\right)\right\}\boldsymbol{H}$ and $\forall \boldsymbol{f} \in \{\boldsymbol{\mu}, \boldsymbol{\lambda}\}$, $\boldsymbol{J}(\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{f})$ is the Jacobian matrix associated to $\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{f}$ with respect to $\boldsymbol{\alpha}_\ell$. We have the following formula:

$$\begin{aligned}\boldsymbol{J}(\widetilde{\boldsymbol{\mathcal{W}}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{f}) = &\,\mathrm{diag}\{|\,\boldsymbol{\Delta}_{\boldsymbol{\alpha}_\ell}|\}\boldsymbol{J}(\boldsymbol{\mathcal{W}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{f}) \\ &+ \mathrm{diag}\{\boldsymbol{\mathcal{W}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{f}\}\boldsymbol{J}(|\,\boldsymbol{\Delta}_{\boldsymbol{\alpha}_\ell}|)\end{aligned}$$

where $\boldsymbol{J}(\boldsymbol{\mathcal{W}}_{\boldsymbol{\alpha}_\ell}\boldsymbol{f})$ is obtained as in [7] and $J(|\,\boldsymbol{\Delta}_{\boldsymbol{\alpha}_\ell}|)$ is obtained by discretizing (32).

The gradient $\nabla\Phi$, corresponding to the regularized motion estimation step, is equal to $\nabla L + \nabla V$, where ∇L is the concatenation of all gate-dependent gradients $\nabla_{\boldsymbol{\alpha}_\ell} L$ and ∇V is the gradient of V.

C. Regularized Image Reconstruction

The image update maximizes the motion-dependent penalized log-likelihood (16), using BSREM [24], using the relative difference prior introduced in [26]. This algorithm was also modified as motion-compensated BSREM (MC-BSREM), using the warping operators introduced in this work. The step size for gated reconstructions (to initialize the joint reconstruction) at a given iteration k was chosen as $\frac{1}{\frac{k}{15}+1}$ and for MC-BSREM reconstruction as $\frac{1}{n_\mathrm{g}}\frac{1}{\frac{k}{15}+1}$, in order to diminish the problem of using too large step sizes at the edges of the field of view when parts of the images are only present in some gates. The parameters of the update are the number of subsets $\mathrm{numSubs}$ and the penalty weighting factor β; the image update stops either when the median relative change between two iterations is lower than 0.1% inside the XCAT phantom or after $300/\mathrm{numSubs}$ iterations.

REFERENCES

[1] S. A. Nehmeh and Y. E. Erdi, "Respiratory motion in positron emission tomography/computed tomography: A review," *Seminars in Nuclear Medicine*, vol. 38, no. 3, pp. 167–176, 2008.

[2] F. Qiao, T. Pan, *et al.*, "A motion-incorporated reconstruction method for gated PET studies," *Physics in Medicine and Biology*, vol. 51, no. 15, pp. 3769–3783, 2006. DOI: 10.1088/0031-9155/51/15/012.

[3] Y. Lu, K. Fontaine, *et al.*, "Respiratory motion compensation for pet/ct with motion information derived from matched attenuation-corrected gated pet data," *Journal of Nuclear Medicine*, vol. 59, no. 9, pp. 1480–1486, 2018. DOI: 10.2967/jnumed.117.203000.

[4] E. C. Emond, A. Bousse, *et al.*, "Effect of attenuation mismatches in time of flight PET reconstruction," *Physics in Medicine and Biology*, vol. 65, no. 8, p. 085 009, 2020. DOI: 10.1088/1361-6560/ab7a6f.

[5] M. W. Jacobson and J. A. Fessler, "Joint estimation of image and deformation parameters in motion-corrected PET," in *2003 IEEE Nuclear Science Symposium. Conference Record (IEEE Cat. No.03CH37515)*, 2003, 3290–3294 vol.5. DOI: 10.1109/NSSMIC.2003.1352599.

[6] A. Rezaei, C. Michel, *et al.*, "Simultaneous reconstruction of the activity image and registration of the CT image in TOF-PET," *Physics in Medicine and Biology*, vol. 61, no. 4, pp. 1852–1874, 2016, ISSN: 13616560. DOI: 10.1088/0031-9155/61/4/1852.

[7] A. Bousse, O. Bertolli, *et al.*, "Maximum-likelihood joint image reconstruction/motion estimation in attenuation-corrected respiratory gated PET/CT using a single attenuation map," *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 217–228, 2016. DOI: 10.1109/TMI.2015.2464156.

[8] A. Bousse, O. Bertolli, *et al.*, "Maximum-likelihood joint image reconstruction and motion estimation with misaligned attenuation in TOF-PET/CT," *Physics in Medicine and Biology*, vol. 61, no. 3, pp. L11–L19, 2016. DOI: 10.1088/0031-9155/61/3/l11.

[9] A. Bousse, R. Manber, *et al.*, "Evaluation of a direct motion estimation/correction method in respiratory-gated PET/MRI with motion-adjusted attenuation," *Medical Physics*, vol. 44, no. 6, pp. 2379–2390, 2017. DOI: 10.1002/mp.12253.

[10] V. Cuplov, B. F. Holman, *et al.*, "Issues in quantification of registered respiratory gated PET/CT in the lung," *Physics in Medicine and Biology*, vol. 63, no. 1, p. 015 007, 2017. DOI: 10.1088/1361-6560/aa950b.

[11] S. El-Chemaly, S. J. Levine, and J. Moss, "Lymphatics in lung disease," *Annals of the New York Academy of Sciences*, vol. 1131, no. 1, pp. 195–202, 2008. DOI: 10.1196/annals.1413.017.

[12] T. Guerrero, K. Sanders, *et al.*, "Dynamic ventilation imaging from four-dimensional computed tomography," *Physics in Medicine and Biology*, vol. 51, no. 4, pp. 777–791, 2006. DOI: 10.1088/0031-9155/51/4/002.

[13] E. Haber and J. Modersitzki, "Numerical methods for volume preserving image registration," *Inverse Problems*, vol. 20, no. 5, pp. 1621–1638, 2004. DOI: 10.1088/0266-5611/20/5/018.

[14] J. M. Reinhardt, K. Ding, *et al.*, "Registration-based estimates of local lung tissue expansion compared to xenon ct measures of specific ventilation," *Medical Image Analysis*, vol. 12, no. 6, pp. 752–763, 2008. DOI: 0.1016/j.media.2008.03.007.

[15] Y. Yin, E. A. Hoffman, and C.-L. Lin, "Mass preserving nonrigid registration of CT lung images using cubic B-spline," *Medical physics*, vol. 36, no. 9, pp. 4213–4222, 2009. DOI: 10.1118/1.3193526.

[16] F. Gigengack, L. Ruthotto, *et al.*, "Motion correction in dual gated cardiac PET using mass-preserving image registration," *IEEE Transactions on Medical Imaging*, vol. 31, no. 3, pp. 698–712, 2012. DOI: 10.1109/TMI.2011.2175402.

[17] V. Gorbunova, J. Sporrin, *et al.*, "Mass preserving image registration for lung CT," *Medical Image Analysis*, vol. 16, no. 4, pp. 786–795, 2012, ISSN: 13618415. DOI: 10.1016/j.media.2011.11.001.

[18] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Transactions on Medical Imaging*, vol. 1, no. 2, pp. 113–122, 1982. DOI: 0.1109/TMI.1982.4307558.

[19] K. Thielemans, E. Asma, and R. M. Manjeshwar, "Mass-preserving image registration using free-form deformation fields," in *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, 2009, pp. 2490–2495. DOI: 10.1109/NSSMIC.2009.5402070.

[20] H. Ue, H. Haneishi, *et al.*, "Nonlinear motion correction of respiratory-gated lung SPECT images," *IEEE Transactions on Medical Imaging*, vol. 25, no. 4, pp. 486–495, 2006. DOI: 10.1109/TMI.2006.871546.

[21] T. Marin and J. G. Brankov, "Deformable left-ventricle mesh model for motion-compensated filtering in cardiac gated SPECT," *Medical Physics*, vol. 37, no. 10, pp. 5471–5481, 2010. DOI: 10.1118/1.3483098.

[22] H. Chang and J. M. Fitzpatrick, "A Technique for Accurate Magnetic Resonance Imaging in the Presence of Field Inhomogeneities," *IEEE Transactions on Medical Imaging*, vol. 11, no. 3, pp. 319–329, 1992. DOI: 10.1109/42.158935.

[23] D. Rueckert, L. I. Sonoda, *et al.*, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999. DOI: 10.1109/42.796284.

[24] S. Ahn and J. A. Fessler, "Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms," *IEEE Transactions on Medical Imaging*, vol. 22, no. 5, pp. 613–626, 2003. DOI: 10.1109/TMI.2003.812251.

[25] J. Nocedal, "Updating quasi-newton matrices with limited storage," *Mathematics of computation*, vol. 35, no. 151, pp. 773–782, 1980. DOI: 10.2307/2006193.

[26] J. Nuyts, D. Bequé, *et al.*, "A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography," *IEEE Transactions on Nuclear Science*, vol. 49, no. 1, pp. 56–60, 2002. DOI: 10.1109/TNS.2002.998681.

[27] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992. DOI: 10.1016/0167-2789(92)90242-F.

[28] K. Cao, K. Ding, *et al.*, "Improving intensity-based lung CT registration accuracy utilizing vascular information," *International Journal of Biomedical Imaging*, vol. 2012, 2012, ISSN: 16874188. DOI: 10.1155/2012/285136.

[29] J. Nuyts, P. Dupont, *et al.*, "Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms," *IEEE Transactions on Medical Imaging*, vol. 18, no. 5, pp. 393–403, 1999. DOI: 10.1109/42.774167.

[30] M. Defrise, A. Rezaei, and J. Nuyts, "Time-of-flight PET data determine the attenuation sinogram up to a constant," *Physics in Medicine and Biology*, vol. 57, pp. 885–899, 2012. DOI: 10.1088/0031-9155/57/4/885.

[31] D. Hwang *et al.*, "Data-driven respiratory phase-matched attenuation correction without CT," in *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC) (in press)*, 2019.

[32] D. Hwang, S. K. Kang, *et al.*, " Generation of PET Attenuation Map for Whole-Body Time-of-Flight 18 F-FDG PET/MRI Using a Deep Neural Network Trained with Simultaneously Reconstructed Activity and Attenuation Maps," *Journal of Nuclear Medicine*, vol. 60, no. 8, pp. 1183–1189, 2019, ISSN: 0161-5505. DOI: 10.2967/jnumed.118.219493.

[33] M. P. Heinrich, I. J. A. Simpson, *et al.*, "Deformable image registration by combining uncertainty estimates from supervoxel belief propagation," *Medical Image Analysis*, vol. 27, pp. 57–71, 2016. DOI: 10.1016/j.media.2015.09.005.

[34] L. Lafitte, C. Zachiu, *et al.*, "Accelerating multi-modal image registration using a supervoxel-based variational framework," *Physics in Medicine and Biology*, vol. 63, no. 23, p. 235 009, 2018. DOI: 10.1088/1361-6560/aaebc2.

[35] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996. DOI: 10.1109/83.536892.

[37] C. Zhu, R. H. Byrd, *et al.*, "Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization," *ACM Transaction on Mathematical Software*, vol. 23, no. 4, 550—560, 1997. DOI: 10.1145/279232.279236.

[38] J. J. Moré and D. J. Thuente, "Line Search Algorithms with Guaranteed Sufficient Decrease," *ACM Transaction on Mathematical Software*, vol. 20, no. 3, pp. 286–307, 1994. DOI: 10.1145/192115.192132.