# JOINT ESTIMATION OF RESPIRATORY MOTION AND ACTIVITY
# IN 4D PET USING CT SIDE INFORMATION

*Matthew W. Jacobson and Jeffrey A. Fessler*

The Department of Electrical Engineering and Computer Science,
The University of Michigan

## ABSTRACT

In previous work, we proposed a Poisson statistical model for gated PET data in which the distribution was parametrized in terms of both image intensity and motion parameters. The motion parameters related the activity image in each gate to that of a base image in some fixed gate. By doing maximum loglikelihood (ML) estimation of all parameters simultaneously, one obtains an estimate of the base gate image that exploits the full set of measured sinogram data. Previously, this joint ML approach was compared, in a highly simplified single-slice setting, to more conventional methods. Performance was measured in terms of the recovery of tracer uptake in a synthetic lung nodule. This paper reports the extension to 3D with much more realistic simulated motion. Furthermore, in addition to pure ML estimation, we consider the use of side information from a breath-hold CT scan to facilitate regularization, while preserving hot lesions of the kind seen in FDG oncology studies.

## 1. INTRODUCTION

In past years, the conventional practice in PET has been to reconstruct based on ungated sinogram data and to ignore the effects of anatomical motion in the patient. The consequences of this practice are motion artefacts in the reconstructed image and the blur of possibly cancerous lesions that one wishes to quantify. In more recent years, there has been much interest (e.g., [1, 2, 3]) in overcoming these drawbacks by reconstructing from gated sinograms in ways that account for patient motion between gates.

The intuitive extension of conventional reconstruction methods to gated data is to reconstruct an image from each sinogram gate (or frame), align these images by image registration methods, and finally to fuse them together in some sort of consolidation step (e.g., by averaging them). Variants of this approach, which we call Frame-Wise reconstruction with Post-Registration (FWPR) have been proposed [1, 2, 3]. In these types of approaches, the registration step is done post-reconstruction and hence, typically, does not

make full use of the sinogram measurement statistics. Moreover, each image registration step occurs between noisy images reconstructed from low-count, single-gate data, and can be expected to result in poor alignment.

This motivated us in earlier work [4] to pose a parametric Poisson model for the gated measurements involving a single unknown activity image and a set of deformation variables describing how that image deforms from gate to gate due to patient motion. By maximizing the loglikelihood for this model, a method we call Joint Estimation by Deformation Modeling (JEDM), one determines both image and deformation parameter estimates jointly from the full set of sinogram measurements. We previously assessed JEDM for a highly simplified simulated thorax scan, involving a single-slice image, a 2-gate acquisition and a 1-parameter deformation model. Preliminary results showed JEDM to outperform other methods in terms of both lung lesion tracer uptake recovery and motion estimation.

In this article, we continue this work, but look at much more realistic simulations. Reconstruction is done with a 3D B-spline motion model based on 2300 parameters per gate. Moreover, the simulated phantom is derived from CT scans of actual thorax anatomy. Thirdly, whereas before we tested only pure maximum likelihood estimation, here we add a roughness penalty term to accomplish regularization. In designing the penalty, we exploit side information – that is assumed to be available from a breath hold CT scan – to avoid smoothing over hot lesions in a region of interest.

### 1.1. Parametric Statistical Model for Gated PET Data

In the statistical PET reconstruction literature, it is common to model a vector of ungated measured projections $\boldsymbol{y}$ as Poisson with mean

$$\bar{\boldsymbol{y}}(\boldsymbol{\lambda}) = \tau(\boldsymbol{P}\boldsymbol{\lambda} + \boldsymbol{r}) \qquad (1)$$

where $\boldsymbol{\lambda}$ is an unknown vector of activity image samples, i.e., each component $\boldsymbol{\lambda}_j$, $j = \{1, \ldots, J\}$ of $\boldsymbol{\lambda}$ represents a sample of a continuous activity distribution at location $(x_j, y_j, z_j)$. We also think of $\boldsymbol{\lambda}$ as the image volume that we wish to reconstruct. The remaining quantities (all assumed

known) are the duration of the scan $\tau$, a forward projection matrix $\boldsymbol{P}$, and a mean vector of background count rates $\boldsymbol{r}$.

When considering a concatenation $\boldsymbol{y} = (\boldsymbol{y}^0, \ldots, \boldsymbol{y}^{T-1})$ of gated sinograms, where $T$ is the number of gates, a natural extension of (1) is

$$\bar{\boldsymbol{y}}^0(\boldsymbol{\lambda}) = \tau_0(\boldsymbol{P\lambda} + \boldsymbol{r}^0) \tag{2}$$

$$\bar{\boldsymbol{y}}^t(\boldsymbol{\lambda}, \boldsymbol{\alpha}^t) = \tau_t(\boldsymbol{P}\mathcal{T}(\boldsymbol{\alpha}^t)\boldsymbol{\lambda} + \boldsymbol{r}^t), \quad 1 \leq t \leq T-1. \tag{3}$$

That is, the mean of the sinogram data in gate 0 are now based on projections of $\boldsymbol{\lambda}$, the activity image samples in that gate. In each subsequent gate, however, the mean is based on projections of $\mathcal{T}(\boldsymbol{\alpha}^t)\boldsymbol{\lambda}$, a transformation of $\boldsymbol{\lambda}$. The transformation operator $\mathcal{T}(\boldsymbol{\alpha}^t)$ is parametrized by an unknown deformation parameter vector $\boldsymbol{\alpha}^t$.

We shall restrict our attention to a class of transformations $\mathcal{T}$ commonly considered in non-rigid, intensity-based image registration,

$$[\mathcal{T}(\boldsymbol{\alpha}^t)\boldsymbol{\lambda}]_j = \sum_k \boldsymbol{\lambda}_k h_k(d(x_j, y_j, z_j, \boldsymbol{\alpha}^t)), \quad t > 0.$$

Here $h_k(x, y, z)$ are activity basis functions used to interpolate the voxel values $\boldsymbol{\lambda}_k$ in between voxel locations and $d(x, y, z, \boldsymbol{\alpha})$ is a $\boldsymbol{\alpha}$-dependent coordinate transformation of the form

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \stackrel{d}{\mapsto} \begin{bmatrix} x + \sum_{k=1}^{K_x} \alpha_{X,k} b_k^X(x, y, z) \\ y + \sum_{k=1}^{K_y} \alpha_{Y,k} b_k^Y(x, y, z) \\ z + \sum_{k=1}^{K_z} \alpha_{Z,k} b_k^Z(x, y, z) \end{bmatrix}.$$

The $\{b_k^C(x, y, z)\}_{C=\{X,Y,Z\}}$ are deformation basis functions. In what follows, we shall use, for the $b_k^C$, cubic B-splines centered at the nodes of a control point grid. This choice has been found to provide an effective motion model for various imaging modalities [5, 6] . For the image domain interpolators $h_k$, we use cubic cardinal splines situated at each pixel location.

## 1.2. Proposed Reconstruction Methods

In this work, we shall compare three methods, described below, for reconstructing $\boldsymbol{\lambda}$. Two of these (JEDM and FWPR-PA) were considered in our previous work [4].

**1. Joint Estimation with Deformation Modeling (JEDM).** In the JEDM method, we consider the joint penalized loglikelihood function based on (2) and (3)

$$\Phi_{\text{JEDM}}(\boldsymbol{\lambda}, \{\boldsymbol{\alpha}^t\}) = \text{KL}(\boldsymbol{y}^0, \bar{\boldsymbol{y}}^0(\boldsymbol{\lambda}))$$
$$+ \sum_{t>0} \text{KL}(\boldsymbol{y}^t, \bar{\boldsymbol{y}}^t(\boldsymbol{\lambda}, \boldsymbol{\alpha}^t)) + \beta R_{\text{act}}(\boldsymbol{\lambda}). \tag{4}$$

where $\text{KL}()$ is the Kullback-Leibler distance,[1] $R_{\text{act}}$ is an activity roughness penalty function, and $\beta \geq 0$ is a regularization parameter. We then reconstruct an image $\hat{\boldsymbol{\lambda}}$ according to

$$\hat{\boldsymbol{\lambda}} = \underset{\boldsymbol{\lambda}}{\operatorname{argmin}} \left\{ \min_{\{\boldsymbol{\alpha}^t\}} \Phi_{\text{JEDM}}(\hat{\boldsymbol{\lambda}}, \{\boldsymbol{\alpha}^t\}) \right\}$$

We introduced this approach in [4], but tested there only the unregularized case (i.e., $\beta = 0$), which corresponds to pure maximum likelihood estimation.

**2. Frame-Wise reconstruction with Post-Registration (FWPR).** An FWPR method is our generic terminology for a method in which images $\{\hat{\boldsymbol{\lambda}}^t\}$, each reconstructed separately from the corresponding sinogram gate $\boldsymbol{y}^t$, are post-registered and consolidated in some way to produce a final image $\hat{\boldsymbol{\lambda}}$. There are two varieties that we consider here.

(a) **FWPR with Post-Averaging (FWPR-PA).** In the post-registration step, each $\hat{\boldsymbol{\lambda}}^t$, $t > 0$, is registered to the common target image $\hat{\boldsymbol{\lambda}}^0$. This yields deformation parameter estimates $\{\hat{\boldsymbol{\alpha}}^t\}$ satisfying, to some degree of accuracy

$$\mathcal{T}(\hat{\boldsymbol{\alpha}}^t)\hat{\boldsymbol{\lambda}}^t \approx \hat{\boldsymbol{\lambda}}^0. \tag{5}$$

Consolidation is then accomplished by taking the weighted average (according to gate duration) of $\hat{\boldsymbol{\lambda}}^0$ and the $\{\mathcal{T}(\hat{\boldsymbol{\alpha}}^t)\hat{\boldsymbol{\lambda}}^t\}$,

$$\hat{\boldsymbol{\lambda}} = \frac{\tau_0\hat{\boldsymbol{\lambda}}^0 + \sum_{t=1}^{T-1} \tau_t\mathcal{T}(\hat{\boldsymbol{\alpha}}^t)\hat{\boldsymbol{\lambda}}^t}{\sum_{t=0}^{T-1} \tau_t}.$$

This approach is an intuitive and natural one, and seems to have occured independently to various investigators (e.g., [1, 2, 4]). However, unlike JEDM, neither the determination of $\{\hat{\boldsymbol{\alpha}}^t\}$, nor the final consolidation of the $\{\hat{\boldsymbol{\lambda}}^t\}$ are based on a model for the measurement statistics.

(b) **FWPR with Penalized-Likelihood Consolidation (FWPR-PLC).** In the post-registration step, each $\hat{\boldsymbol{\lambda}}^t$, $t > 0$, serves this time as a target image to which $\hat{\boldsymbol{\lambda}}^0$ is registered. This yields deformation parameter estimates $\{\hat{\boldsymbol{\alpha}}^t\}$ satisfying, to some degree of accuracy

$$\mathcal{T}(\hat{\boldsymbol{\alpha}}^t)\hat{\boldsymbol{\lambda}}^0 \approx \hat{\boldsymbol{\lambda}}^t. \tag{6}$$

Consolidation is then accomplished by substituting these $\{\hat{\boldsymbol{\alpha}}^t\}$ into (4) as knowns and carrying out penalized likelihood minimization with respect to $\boldsymbol{\lambda}$,

$$\hat{\boldsymbol{\lambda}} = \underset{\boldsymbol{\lambda}}{\operatorname{argmin}} \left\{ \Phi_{\text{JEDM}}(\boldsymbol{\lambda}, \{\hat{\boldsymbol{\alpha}}^t\}) \right\}.$$

---

[1]The KL distance is the negative of a loglikelihood function for Poisson measurements. Minimizing it corresponds to maximum likelihood estimation.

This variation is slightly more statistically principled than FWPR-PA, in that the consolidation step exploits the loglikelihood of the measurements. A similar approach has been considered recently in [3], but where the $\{\hat{\boldsymbol{\alpha}}^t\}$ are derived from the registration of gated CT scans. Here, we assume that only a breath-hold CT scan is available, and so we must derive the deformation parameter estimates from the gated emission data.

### 1.3. Lesion Preserving Penalty Design Strategy

One of the aims of motion correction in thorax scans (and the one of central interest to us here) is to prevent the blur of lung lesions that one wishes to quantify. If $R_{\text{act}}(\boldsymbol{\lambda})$ penalizes roughness indiscriminately throughout the image $\boldsymbol{\lambda}$, we would blur such lesions and perhaps detract from the purpose of the motion-correction. We therefore employ a scheme whereby CT side-information is used to exclude these lesions from penalization.

Since a breath hold CT scan is available, we know the approximate location of the lesion in the PET image near full inspiration. In addition, we know the approximate amplitude of typical respiratory motion. These two pieces of information can be combined to localize the lesion to some generously sized ROI, whose dimensions are on the order of a few centimeters. We can adjust $R_{\text{act}}(\boldsymbol{\lambda})$ to ignore voxels in this region. However, since many voxels outside of this ROI remain subject to the roughness penalty, we can still hope that the regularization will greatly reduce underdetermined behavior in the reconstruction.

### 2. EXPERIMENTS

We derived synthetic thorax PET images at 5 levels of respiration from real thorax CT images. The CT images were downsampled to voxels of size $4 \times 4 \times 2$ millimeters, cropped to a grid size of $81 \times 105 \times 17$, and mapped to appropriate PET image intensity values. These 5 images were co-registered using using an $11 \times 14 \times 5 \times 3$ control point grid of cubic B-spline deformation basis functions using the Sum of Squared Differences (SSD) criterion. The resulting parameters $\{\boldsymbol{\alpha}_{\text{true}}^t\}$ were taken as ground truth. A synthetic ellipsoidal lung lesion of axial radius 2 mm and transaxial radius 4 mm was inserted into the gate 0 image $\boldsymbol{\lambda}_{\text{true}}$. The four subsequent gates were then replaced by $\mathcal{T}(\boldsymbol{\alpha}_{\text{true}}^t)\boldsymbol{\lambda}$. A slice of $\boldsymbol{\lambda}_{\text{true}}$ containing the lung lesion is shown in Figure 1(a). A sense of the lesion blur that would result from an ungated reconstruction is given in Figure 1(b) where the superposition of the image from all gates is shown.

Poisson sinogram measurements $\{\boldsymbol{y}^t\}$ described by (2) and (3) were simulated with 0.5 million total counts per slice, 10% of which were background. Each slice of the

5 gates was forward projected into a 105 radial by 128 angle system of sinogram bins. In Figures 2(a)-(c), a common slice of images reconstructed using the 3 motion correction strategies of Section 1.2 are shown, along with the percent lesion uptake recovered. Similarly, in Figure 2(d), we reconstructed from the superposition of the $\{\boldsymbol{y}^t\}$ (to imitate an ungated reconstruction). All frame-wise reconstructions (in FWPR schemes) and the ungated reconstruction used quadratically penalized likelihood estimation. Moreover, all penalized likelihood steps used to obtain these images excluded an approximately $3 \times 3 \times 1.5$ cm ROI around the lesion (as discussed in Section 1.3). We also tested JEDM without excluding any ROI from the roughness penalty (see Figure 2(e)). Conversely, Figure 2(f) shows the result of pure maximum likelihood estimation, in which the motion parameters were known and no roughness penalty was applied (i.e., $\Phi_{\text{JEDM}}(\cdot, \{\boldsymbol{\alpha}_{\text{true}}^t\})$ was minimized with $\beta = 0$).

### 3. CONCLUSIONS

In these preliminary results, the JEDM approach greatly outperformed ungated, unmotion-corrected reconstruction and significantly outperformed FWPR motion-correction strategies in terms of tracer uptake estimation. However, the semi-statistical FWPR-PLC approach does not lag too far behind. This may motivate a hybrid approach in which an FWPR-PLC reconstruction is used to initialize a JEDM reconstruction, and thereby refined. We shall consider this further in future work. Comparing Figures 2(a) and (e), one sees that indiscriminate roughness penalties can greatly detract from lesion recovery. The lesion-preserving strategy of Section 1.3 appears to achieve the best of both worlds in terms of recovery and noise reduction. In Figure 2(a), the technique endowed JEDM with nearly the same uptake recovery as the fully unsmoothed, known-motion case (f) but with the lower background noise of (e).

### 4. REFERENCES

[1] G. J. Klein, B. W. Reutter, and R. H. Huesman, "Four-dimensional affine registration models for respiratory-gated PET," *IEEE Trans. Nuc. Sci.*, vol. 48, no. 3, pp. 756–60, June 2001.

[2] B. Thorndyke, E. Schreibmann, P. Maxim, B. Loo, A. Boyer, A. Koong, and L. Xing, "Enhancing 4D PET through retrospective stacking," *Proc. Amer. Assoc. Phys. Med.*, vol. 32, no. 6, pp. 2096, June 2005.

[3] F. Qiao, T. Pan, J. W. Clark, and O. R. Mawlawi, "Compensating respiratory motion in PET image reconstruction using 4D PET/CT," in *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, 2005.
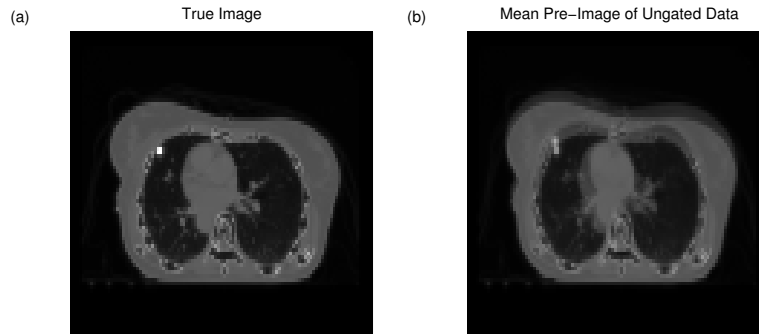
**Fig. 1**. (a) The true activity (slice #8) in gate 0. (b) The superposition of the true activity (slice #8) from all gates.
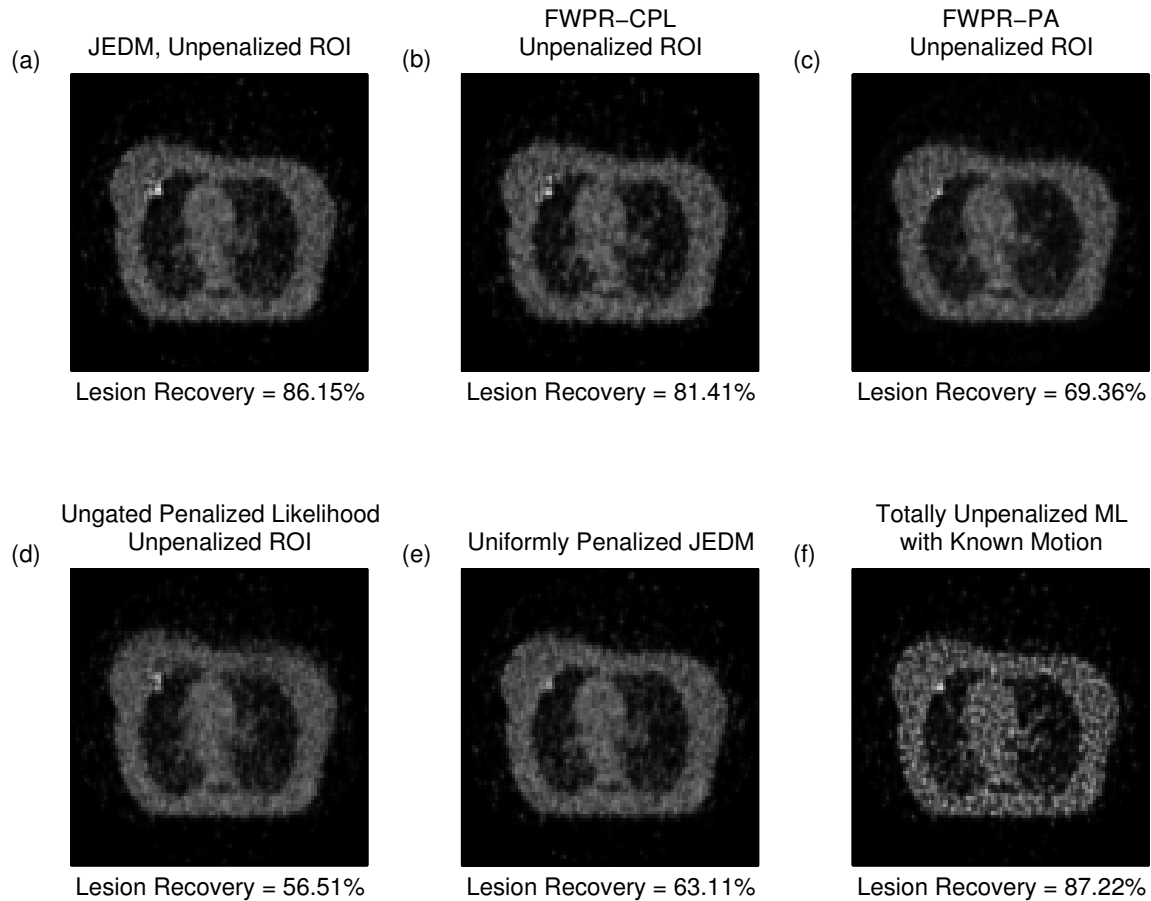


**Fig. 2**. Slice #8 of the image reconstructed by various methods, and with the percent tracer uptake recovery shown for each.

[4] M. W. Jacobson and J. A. Fessler, "Joint estimation of image and deformation parameters in motion-corrected PET," in *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, 2003, vol. 5, pp. 3290–4.

[5] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "PET-CT image registration in the chest using free-form deformations," *IEEE Trans. Med. Imag.*, vol. 22, no. 1, pp. 120–8, Jan. 2003.

[6] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 712–21, Aug. 1999.