# Motion Compensation for Dynamic PET with Continuous Motion Blur

László Szirmay-Kalos, Dóra Varnyú, and Milán Magdics

Abstract—Dynamic positron emission tomography reconstructs the space-time dependence of radiotracer concentration in a human or animal body. The self-motion of the examined subject often cannot be avoided. Therefore, the motion should be measured and this information should be incorporated into the reconstruction process. The ML-EM reconstruction procedure [1] can be generalized to consider subject movement by concatenating a geometric transformation to the system matrix multiplication. However, a transformation is valid only for a single point in time, thus time frames should be short enough to allow the assumption that motion is constant within them. Fast movements, which occur in the study of neurological disorders, would require a very high number of frames, leading to unacceptable reconstruction times. To attack this problem, we re-formulate the ML-EM scheme to incorporate continuous motion during frames. Instead of a geometric transformation, we apply 3D motion blur. For efficient implementation, the paths of object-space voxels are approximated by polylines and antialiasing line drawing methods are adapted to calculate the blurred result. Having executed motion blur on the voxel array before forward projection and twice during back projection, all other components of the simulation can be reused from the dynamic reconstruction system developed without motion compensation.

## I. INTRODUCTION

N dynamic positron emission tomography (PET), we analyze the dynamic nature of biological processes [2], [3], [4], [5]. The measurement time is divided into time frames according to the tracer dynamics. Direct dynamic tomography algorithms usually iterate two steps: first, an ML-EM [1] static forward projection and back projection are executed for each frame, then the parameters of a time function of predefined algebraic form is fit in each voxel [6]. Thus, the number of projections is proportional to the number of frames, which should be kept under control to avoid unacceptably high reconstruction times and memory requirements.

The self-motion of the examined patient often cannot be avoided, for example when studying neurological disorders

D. Varnyú is with the Budapest University of Technology and Economics, Department of Control Engineering and Information Technology, Hungary (email: varnyu.dora@gmail.com)

M. Magdics is with the Budapest University of Technology and Economics, Department of Control Engineering and Information Technology, Hungary (e-mail: magdics@iit.bme.hu)

like epilepsy, thus the motion should be measured and this information should be incorporated into the reconstruction process [7]. Accurate reconstruction requires projections that simulate all relevant phenomena including positron range, gamma-photon scattering and absorption both in the measured object and in detector crystals, which are all affected by the motion of the body. Forward and back projections are executed for each frame in every iteration, which can be separated from the motion by assuming that inside a frame, subject movement is negligible and a single geometric transformation describes its position. However, this is usually not the case when neurological disorders are examined, since subject motion has generally much quicker dynamics than the tracer diffusion.

Gated approaches [8], [9] keep hits close to discrete time samples to solve the problem of motion inside the frame, but they ignore significant information and are thus less reliable for low-statistics or low-dose experiments.

The other option is to re-bin the list-mode data and find the event list that would have been measured if the subject had not moved, then execute a conventional reconstruction for the modified sinogram [10], [11], [12]. Such event-space aka sinogram filtering techniques are simple to implement and efficient, but they are also rather limited. Event-based methods cannot be applied for effects happening in the detectors as they either add noise or result in filtered measured data that are not of Poission distribution. They also suffer from the missing or lost data problem, which comes from the fact that events are not measured outside the field of view of the scanner, thus they cannot be transformed in to compensate for the events that are transformed out due to subject motion, making the reconstruction distorted. Furthermore, event-based methods operate on LORs, so they transform lines to lines, thus they are not applicable for non-linear deformations.

In this paper, a model-based motion compensation technique is proposed, which incorporates motion into the forward projection and back projection steps. Unlike previous methods, we consider continuous motion inside frames, which allows the significant reduction of the number of frames in case of fast motion.

The structure of this paper is the following. In Section II, we generalize the ML-EM scheme for continuous motion. Section III presents the results obtained first with a 2D phantom, then during a fully 3D reconstruction. Finally, we close the paper with conclusions.

## **II. RECONSTRUCTION ALGORITHM**

This section revisits the ML-EM scheme and generalizes it for continuous object motion. Movement can be arbitrary, not

Manuscript received December 21, 2019. The research has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications), and by ÚNKP-17-4-I. New National Excellence Program of the Ministry of Human Capacities.

L. Szirmay-Kalos is with the Budapest University of Technology and Economics, Department of Control Engineering and Information Technology, Hungary (e-mail: szirmay@iit.bme.hu)



Fig. 1: Overview of the generalized ML-EM reconstruction for continuously moving or deforming objects

necessarily rigid body motion.

Radiotracer isotopes randomly decay generating events in the detectors of the tomograph. An event consists of the identification of the involved detectors, also called *Line Of Response* or *LOR*, and the time of the detection. To get an efficient numerical method, the measured volume is divided into voxels and the measurement time is discretized by *ticks*  $t_1, t_2, \ldots, t_{N_F}$ , then events are binned into *frames*  $(t_F, t_{F+1})$ . The number of voxels and frames are critical, because the computational complexity of the reconstruction is proportional to their product.

Radiotracer concentration is described by a kinetic model  $\mathcal{K}(\mathbf{p}_V, t)$ , which represents each voxel V with a parameter vector  $\mathbf{p}_V = (\mathbf{p}_{V,1}, \dots, \mathbf{p}_{V,N_P})$ . The expected number of decays in voxel V in time interval  $[t_F, t_{F+1}]$  is

$$\tilde{x}_F(\mathbf{p}_V) = \int_{t_F}^{t_{F+1}} \mathcal{K}(\mathbf{p}_V, t) e^{-\lambda t} \mathrm{d}t, \qquad (1)$$

where  $\lambda$  is the decay rate of the tracer. (Expectations are generally indicated by a tilde in this paper.)

The correspondence between decays and detected events is established by the system matrix  $\mathbf{A}$ , which expresses the probability of detecting an event in LOR L given that a decay happened in voxel V. If the measured object is moving, then the system matrix is also time-dependent, since a point of its body will be at different places at different times. To handle this, we introduce two voxel arrays, an *object-space* array that follows the measured object, and a *tomograph-space* array that is fixed to the tomograph. To make the distinction clear, object-space voxel activities are denoted by lower case letters and indexed with V or V', while tomograph-space voxel activities with upper-case letters and indexed with W. The system matrix describing the correspondence between tomograph-space voxels and LORs is not time dependent, but tomograph-space voxels have dynamic activity due to both the dynamic radiotracer concentration and the time-dependent transformation between object-space and tomograph-space.

Knowing the transformation between object-space and tomograph-space, we can obtain an *overlap factor*  $O_{V',W}(t)$  between object-space voxel V' and tomograph-space voxel W, which expresses the ratio of the volume in voxel W that actually comes from object-space voxel V'.

The expected number of decays in tomograph-space voxel W in time interval  $[t_F, t_{F+1}]$  is

$$\tilde{X}_{W,F} = \sum_{V'} \int_{t_F}^{t_{F+1}} \mathcal{K}(\mathbf{p}_{V'}, t) e^{-\lambda t} O_{V',W}(t) dt$$
$$= \sum_{V'} \tilde{x}_F(\mathbf{p}_{V'}) o_{V',W,F}, \qquad (2)$$

where  $o_{V',W,F}$  is the *average overlap* of object-space voxel V and tomograph-space voxel W during frame F:

$$o_{V',W,F} = \frac{\int\limits_{t_F}^{t_{F+1}} \mathcal{K}(\mathbf{p}_{V'}, t) e^{-\lambda t} O_{V',W}(t) \mathrm{d}t}{\int\limits_{t_F}^{t_{F+1}} \mathcal{K}(\mathbf{p}_{V'}, t) e^{-\lambda t} \mathrm{d}t}.$$
 (3)

From now on, we assume that this average overlap is independent of parameter vector  $\mathbf{p}_V$ , which is the case when either  $\mathcal{K}(\mathbf{p}_V, t)e^{-\lambda t}$  or the overlap factor are constant in time. Otherwise, it is a good approximation. Taking advantage of w this, we get

$$o_{V',W,F} \approx \frac{1}{t_{F+1} - t_F} \int_{t_F}^{t_{F+1}} O_{V',W}(t) \mathrm{d}t.$$
 (4)

The expected number of detection events  $\tilde{y}_{L,F}$  in LOR L in frame F is the sum of the contributions of all voxels:

$$\tilde{y}_{L,F} = \sum_{W} \mathbf{A}_{L,W} \tilde{X}_{W,F},\tag{5}$$

which is a conventional forward projection.

The measured number of hits in LOR L in frame F follows a Poisson distribution of expectation  $\tilde{y}_{L,F}$ . Because of the statistical independence of different LORs and different frames, the combined probability considering all LORs and frames is the product of the elementary probabilities. According to the concept of *maximum-likelihood reconstruction*, unknown parameters are searched to maximize the following log-likelihood:

$$\log \mathcal{L} = \sum_{L} \sum_{F} \left( y_{L,F} \log \tilde{y}_{L,F} - \tilde{y}_{L,F} \right), \tag{6}$$

where  $y_{L,F}$  is the measured number of detector hits in LOR L in frame F.

The likelihood has an extremum where all partial derivatives are zero:

$$\sum_{F} \sum_{W} \frac{\partial \bar{X}_{W,F}}{\partial \mathbf{p}_{V,P}} \sum_{L} \left( \mathbf{A}_{L,W} \frac{y_{L,F}}{\tilde{y}_{L,F}} - \mathbf{A}_{L,W} \right) = 0, \quad (7)$$

for all object-space voxels  $V = 1, 2, ..., N_V$  and parameters  $P = 1, ..., N_P$ .

Using Eq. 2 and also exploiting our simplifying assumption that the average overlap is independent of the parameter vector, we get

$$\frac{\partial X_{W,F}}{\partial \mathbf{p}_{V,P}} = \sum_{V'} \frac{\partial \tilde{x}_F(\mathbf{p}_{V'})}{\partial \mathbf{p}_{V,P}} o_{V',W,F} = \frac{\partial \tilde{x}_{V,F}}{\partial \mathbf{p}_{V,P}} o_{V,W,F}.$$
 (8)

since the derivative at the right hand side is non-zero only if V = V'. Thus, Eq. 7 simplifies to

$$\sum_{F} \frac{\partial \tilde{x}_{F}(\mathbf{p}_{V})}{\partial \mathbf{p}_{V,P}} \sum_{L} \sum_{W} o_{V,W,F} \left( \mathbf{A}_{L,W} \frac{y_{L,F}}{\tilde{y}_{L,F}} - \mathbf{A}_{L,W} \right) = 0.$$
(9)

The computation of the derivatives of the log-likelihood requires a forward projection and an update operation in each frame F. Indeed, in frame F, the expected number of radioactive decays in voxel V is  $\tilde{x}_F(\mathbf{p}_V)$ , which is forward projected via all overlapped tomograph-space voxels to obtain  $\tilde{y}_{L,F}$  according to Eq. 5. The update operation is called the *back projection* and obtains a new estimate of the activity as

$$\begin{aligned} x_{V,F} &= \tilde{x}_F(\mathbf{p}_V) \cdot \frac{\sum_W o_{V,W,F} \sum_L \mathbf{A}_{L,W} \frac{y_{L,F}}{\tilde{y}_{L,F}}}{\sum_W o_{V,W,F} \sum_L \mathbf{A}_{L,W}} \\ &= \tilde{x}_F(\mathbf{p}_V) \cdot \frac{\sum_W o_{V,W,F} R_{W,F}}{\sum_W o_{V,W,F} S_{W,F}} \\ &= \tilde{x}_F(\mathbf{p}_V) \cdot \frac{\hat{R}_{V,F}}{\hat{S}_{V,F}}, \end{aligned}$$
(10)

where

$$\hat{R}_{V,F} = \sum_{W} o_{V,W,F} R_{W,F} \tag{11}$$

is the *motion blurring* of the numerator of the classical back projection

$$R_{W,F} = \sum_{L} \mathbf{A}_{L,W} \frac{y_{L,F}}{\tilde{y}_{L,F}},\tag{12}$$

and

$$\hat{S}_{V,F} = \sum_{W} o_{V,W,F} S_{W,F} \tag{13}$$

is the motion blurring of the denominator of the classical back projection

$$S_{W,F} = \sum_{L} \mathbf{A}_{L,W}.$$
 (14)

Substituting these into Eq. 7, we obtain:

$$\sum_{F} S_{V,F} \frac{\partial \tilde{x}_{F}(\mathbf{p}_{V})}{\partial \mathbf{p}_{V,P}} \left( \frac{x_{V,F}}{\tilde{x}_{F}(\mathbf{p}_{V})} - 1 \right) = 0.$$
(15)

In this equation,  $\tilde{x}_F(\mathbf{p}_V)$  depends only on the unknown parameter vector  $\mathbf{p}_V$  of voxel V, whereas the calculation of  $x_{V,F}$  consists of a forward and a back projection, therefore its value depends on the parameter vectors of all voxels. So if  $x_{V,F}$  was known, then the computation could be decoupled for different voxels, creating a system of equations with  $N_P$  unknowns. In this way, forward/backward projection is separated from the calculation of the parameter values, thus the complexity of the algorithm will be the sum of the complexities of the two steps and not their product [6]. Using a one-step-late strategy, i.e. taking  $x_{V,F}$  from the previous iteration, the non-linear equation is solved, which can also be imagined as a *curve-fitting* process.

In summary, an iteration of the motion-compensated reconstruction consists of the following steps (Fig. 1):

- 1) For each frame, we compute  $x_{V,F}$ , which in turn includes the following steps:
  - a) Evaluation of the integral of the radiotracer concentration for each object-space voxel V and for this frame to get  $x_F(\mathbf{p}_V)$  (Eq. 1).
  - b) Motion blur to compute the activity of tomographspace voxel W, i.e.  $\tilde{x}_{W,F}$  (Eq. 2).
  - c) Forward projection of the tomograph-space voxels to obtain expected LOR hits  $\tilde{y}_{L,F}$  (Eq. 5).
  - d) Computation of numerator  $R_{W,F}$  and denominator  $S_{W,F}$  of the classical back projection (Eq. 12 and Eq. 14).
  - e) Motion blur of the tomograph-space numerator  $R_{W,F}$  and denominator  $S_{W,F}$  voxel arrays into object-space voxel arrays  $\hat{R}_{V,F}$  and  $\hat{S}_{V,F}$ , respectively (Eq. 11 and Eq. 13).
  - f) Multiplying the object-space activity by the ratio of  $\hat{R}_{V,F}$  and  $\hat{S}_{V,F}$  to evaluate  $x_{V,F}$  (Eq. 10).
- 2) Fitting the parameters on the resulting  $x_{V,F}$  values of each object-space voxel V and frame F by solving Eq. 15.

It is often advantageous to use a simplified system matrix in back projection and simulate accurate particle transport only in the forward projector [13], [14], [15]. For the suggested motion



Fig. 2: Two-dimensional brain phantom used in the experiments.



Fig. 3: Deforming the activity field with time dependent swirl.

compensation technique, simplification means that instead of considering continuous motion, we take a single sample of the movement at the middle of the frame. That is, instead of blurring activities inside a frame, the voxel array is transformed to the object state at the middle of the frame. Note that if blurring is replaced by transformation in both forward and back projection, we get the classical motion compensation approach back. Thus, during the demonstration of the result, we compare three methods: a) transformation in both forward and back projection, which is the classical approach; b) blurring in forward direction and transformation in the back projection, which is the unmatched version of our proposed technique; c) blurring in both forward and back projection, which is the matched version of the proposed approach.

#### **III. RESULTS**

## A. 2D brain phantom

To examine the proposed methods, we first used a 2D brain phantom consisting of four regions: gray matter, white matter, blood, and air (Fig. 2) [16].

The object-spare coordinates are in [-2, 2] units and the measurement time is 10 seconds. The applied motion is a non-linear swirl type deformation, which transforms (x, y) to (x', y') based on the current time t (Fig. 3):

$$x' = x \cos\left(\frac{te^{-\sqrt{x^2 + y^2}}}{2}\right) + y \sin\left(\frac{te^{-\sqrt{x^2 + y^2}}}{2}\right),$$
  
$$y' = -x \sin\left(\frac{te^{-\sqrt{x^2 + y^2}}}{2}\right) + y \cos\left(\frac{te^{-\sqrt{x^2 + y^2}}}{2}\right).$$
  
(16)



Fig. 4:  $L_2$  error curves of the 2D brain phantom as a function of the number if ML-EM iterations, when the measurement time is decomposed into 1, 2, and 3 frames

We executed 50 generalized ML-EM iterations setting the number of frames to 1, 2, and 3. The compared cases were the following. In "No motion", the object is not moving and therefore there is no need for motion compensation. This case can be considered as limit that can be approached with perfect motion compensation. The case of "Motion, no compensation" is the other extreme. Swirling is executed during the simulation of the measurement, but no motion compensation is used during reconstruction. "Motion, transformation in forward and back projection" is the classical approach for motion compensation. The two versions of the new method include the



Fig. 5: Reconstructed activity images of the 2D brain phantom

unmatched "Motion, blurring in forward and transformation in back projection" and the matched "Motion, blurring in both forward and back projection". The  $L_2$  error curves as a function of the iteration number are shown by Fig. 4 and the reconstructed activities by Fig. 5.

Motion compensation resulted in errors between those of the unmoving and the moving but not compensated cases. The matched version was better than the unmatched one, which in turn was better than the classical approach, but the difference diminished as we took more frames. When the frame number was low (only one in this experiment), the unmatched back projector caused a divergence of the error after initial convergence.

#### B. 3D Derenzo phantom

The suggested motion compensation techniques were analyzed in 3D using a Derenzo phantom [17] (Fig. 6) simulation assuming the Tera-tomo nanoScan detector geometry [18], [19]. The phantom moved 8 voxels in the positive x-axis direction in 300 seconds at a constant speed, then moved back to its initial position in another 300 seconds. (No rotational motion was applied.) The measurement time was divided into 2 frames. The examined techniques were the following: a) No compensation, b) Event-based compensation (transforming events to a reference point in time at the beginning of the reconstruction), c) Model-based, transformation in both forward and back projection, d) Model-based, blurring in forward and transformation in back projection, and e) Model-based, blurring in both forward and back projection.

Fig. 7 presents the CC error curves as a function of the iteration number and Fig. 8 displays the reconstructed activity images. Results show that the event-based compensation outperforms the model-based algorithms in terms of accuracy. However, it should be noted that this comparison is not fair



Fig. 6: Slices of the 3D Derenzo phantom

since event-based compensation transforms all events according to their actual time, while other approaches should bin the events into two frames. Moreover, event-based compensation is limited to rigid body motion.

Among the model-based solutions, the methods proposed in this paper resulted in lower errors than the classical approach of transforming activities in both forward and back projection. Although the difference was slight, the matched version performed better than the unmatched one, as was the case in the 2D examinations as well.

### **IV. CONCLUSIONS**

By adding three motion blur executions to the MLEM reconstruction pipeline, it is possible to compensate larger and even non-rigid-body motions without dramatically increasing the number of frames. The blurring operation can be well supported by the GPU and its extra time is negligible compared to the other steps of the reconstruction. Two versions of the proposed approach were presented in the paper, one with a matched and one with an unmatched forward and back projector. Simulation results show that both of them outperform the classical motion compensation method, but the matched version achieves slightly better accuracy.



Fig. 7: CC error curves of the moving Derenzo phantom as a function of the number of ML-EM iterations



(a) No compensation



(b) Event-based compensation (c) Model-based, transformation in both forward and back projection





(d) Model-based, blurring in forward and transformation in back projection (e) Model-based, blurring in both forward and back projection

Fig. 8: Reconstructed activity images of the 3D Derenzo phantom

#### REFERENCES

- L. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imaging*, vol. 1, pp. 113–122, 1982.
- [2] A. J. Reader and H. Zaidi, "Advances in PET image reconstruction," *PET Clinics*, vol. 2, no. 2, pp. 173–190, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1556859807000193
- [3] A. J. Reader and J. Verhaeghe, "4d image reconstruction for emission tomography," *Physics in Medicine and Biology*, vol. 59, no. 22, p. R371, 2014. [Online]. Available: http://stacks.iop.org/0031-9155/59/i= 22/a=R371
- [4] L. Szirmay-Kalos and A. Kacsó, "Regularizing direct parametric reconstruction for dynamic pet with the method of sieves," in *Molecular Imaging Conference*, ser. MIC '16, 2016, pp. M16D–1.
- [5] L. Szirmay-Kalos, A. Kacsó, M. Magdics, and B. Tóth, "Dynamic pet reconstruction on the gpu," *Periodica Polytechnica Electrical Engineering and Computer Science*, vol. 62, no. 4, pp. 134–143, 2018. [Online]. Available: https://pp.bme.hu/eecs/article/view/11739
- [6] G. Wang and J. Qi, "An optimization transfer algorithm for nonlinear parametric image reconstruction from dynamic pet data," *IEEE Trans Med Imaging*, vol. 31, no. 10, pp. 1977–1988, 2012.
- [7] J. Jiao, A. Bousse, K. Thielemans, N. Burgos, P. S. J. Weston, J. M. Schott, D. Atkinson, S. R. Arridge, B. F. Hutton, P. Markiewicz, and S. Ourselin, "Direct parametric reconstruction with joint motion estimation/correction for dynamic brain pet data," *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 203–213, Jan 2017.
- [8] M. Toussaint, J. P. Dussault, and R. Lecomte, "Revisiting motion compensation models in pet image reconstruction," in 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), April 2016, pp. 90–94.
- [9] J. Cui, J. Yang, E. Graves, and C. S. Levin, "Gpu-enabled pet motion compensation using sparse and low-rank decomposition," in 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC), Oct 2012, pp. 3367–3370.
- [10] M. Menke, M. S. Atkins, and K. R. Buckley, "Compensation methods for head motion detected during pet imaging," *IEEE Transactions on Nuclear Science*, vol. 43, no. 1, pp. 310–317, Feb 1996.
- [11] P. Buhler, U. Just, E. Will, J. Kotzerke, and J. van den Hoff, "An accurate method for correction of head movement in pet," *IEEE Transactions on Medical Imaging*, vol. 23, no. 9, pp. 1176–1185, Sept 2004.
- [12] C. Chan, X. Jin, E. K. Fung, M. Naganawa, T. Mulnix, R. E. Carson, and C. Liu, "Event-by-event respiratory motion correction for pet with 3dimensional internal-external motion correlation," in 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC), Oct 2012, pp. 2117–2122.
- [13] G. Zeng and G. Gullberg, "Unmatched projector/backprojector pairs in an iterative reconstruction algorithm," *IEEE Transactions on Medical Imaging*, vol. 19, no. 5, pp. 548–555, 2000.
- [14] N.-Y. Lee and Y. Choi, "Theoretical investigation on an unmatchedf backprojector for iterative reconstruction in emission computed tomography," *Journal of the Korean Physical Society*, vol. 59, no. 2, pp. 367–375, 2011.
- [15] M. Magdics, L. Szirmay-Kalos, B. Tóth, and A. Penzov, "Analysis and control of the accuracy and convergence of the ML-EM iteration," *LECTURE NOTES IN COMPUTER SCIENCE*, vol. 8353, pp. 147–154, 2014.
- [16] L. Szirmay-Kalos, M. Magdics, B. Tóth, and T. Bükki, "Averaging and Metropolis iterations for positron emission tomography," *IEEE Trans Med Imaging*, vol. 32, no. 3, pp. 589–600, 2013.
- [17] S. E. Derenzo, "Mathematical removal of positron range blurring in high resolution tomography," *IEEE Trans. Nucl. Sci.*, vol. 33, pp. 546–549, 1986.
- [18] M. Magdics et al., "TeraTomo project: a fully 3D GPU based reconstruction code for exploiting the imaging capability of the NanoPET/CT system," in *World Molecular Imaging Congress*, 2010.
- [19] M. Magdics et al., "Performance Evaluation of Scatter Modeling of the GPU-based "Tera-Tomo" 3D PET Reconstruction," in *IEEE Nuclear Science Symposium and Medical Imaging*, 2011, pp. 4086–4088.