

**Revisiting motion compensation models
in PET image reconstruction**

M. Toussaint, J.-P. Dusseault,
R. Lecomte

G-2016-25

April 2016

Cette version est mise à votre disposition conformément à la politique de libre accès aux publications des organismes subventionnaires canadiens et québécois.

Avant de citer ce rapport, veuillez visiter notre site Web (<https://www.gerad.ca/fr/papers/G-20016-25>) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

This version is available to you under the open access policy of Canadian and Quebec funding agencies.

Before citing this report, please visit our website (<https://www.gerad.ca/en/papers/G-20016-25>) to update your reference data, if it has been published in a scientific journal.

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2016
– Bibliothèque et Archives Canada, 2016

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*.

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2016
– Library and Archives Canada, 2016

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Télec. : 514 340-5665
info@gerad.ca
www.gerad.ca

Revisiting motion compensation models in PET image reconstruction

Maxime Toussaint^a

Jean-Pierre Dussault^a

Roger Lecomte^b

^a *Department of Computer Science, Université de Sherbrooke, Sherbrooke (Québec) Canada, J1K 2R1*

^b *Sherbrooke Molecular Imaging Center, CRCHUS, and Department of Nuclear Medicine and Radiobiology, Université de Sherbrooke, Sherbrooke (Québec) Canada J1H 5N4*

`maxime.toussaint@usherbrooke.ca`

`jean-pierre.dussault@usherbrooke.ca`

`roger.lecomte@usherbrooke.ca`

April 2016

Les Cahiers du GERAD

G–2016–25

Copyright © 2016 GERAD

Abstract : Positron emission tomography (PET) image reconstruction in the presence of periodic motion, such as heartbeat and breathing, has been actively investigated since the introduction of gated data acquisition. While each gated sinogram can be reconstructed independently, better images can be obtained by incorporating all gated sinograms, which are considered statistically independent, into an iterative reconstruction scheme. However, this gain in image quality is achieved at the expense of substantial extra computational resources to handle all gated sinograms simultaneously. We provide a comparative study with a model that is built upon the interdependence of the parameters defining the stochastic processes behind each gated sinogram to yield a single merged sinogram. We show, using simulated and experimental datasets, that images of similar quality than with the classical model can be produced, while requiring much less computing resources.

Key Words : Positron emission tomography, Motion compensation, Image reconstruction, Poisson process

Résumé : La reconstruction d'image par tomographie par émission de positrons (TEP) en présence de mouvement périodique, tel que les battements du coeur et la respiration, a été étudiée intensivement depuis l'introduction de l'acquisition par fenêtres synchronisées. Bien que chacun des sinogrammes synchronisés peut être reconstruit indépendamment, de meilleures images peuvent être obtenues en incorporant tous les sinogrammes synchronisés, qui sont considérés statistiquement indépendants, dans un modèle de reconstruction itératif. Cependant, ce gain en qualité d'image est obtenu en échange d'une augmentation importante des ressources informatiques, causé par la manipulation simultanée de tous les sinogrammes synchronisés.

Nous comparons cette approche à un modèle qui est construit à partir de l'interdépendance des paramètres définissant le processus stochastique sous-jacent chacun des sinogrammes synchronisés afin de manipuler un unique sinogramme fusionné. Nous montrons, à partir de données simulées et expérimentales, que des images de qualité semblable à ceux du modèle classique peuvent être produites, tout en exploitant beaucoup moins de ressources informatiques.

1 Introduction

PET is an imaging technique that reconstructs the *in vivo* radiotracer distribution in a subject. Since several minutes of data acquisition are often needed to produce usable images, PET is especially prone to motion artefacts, which can lead to false diagnoses [1]. Periodic motions, such as heartbeat and breathing, represent a special case. Due to their cyclic nature, the motion they induce through time can be defined as a loop of typically 8 to 16 motion states. Thus, motion-free datasets can be created by distributing each acquired event according to which motion state it originates from. Independent three-dimensional (3D) reconstruction of each of those datasets is, *de facto*, the motion compensation approach used in clinical applications. However, the injected dose of radiotracer or the total acquisition time usually needs to be increased in order for each dataset, called gated sinogram, to preserve a statistical accuracy equivalent to an ungated acquisition.

After reaching kinetic equilibrium, the radiotracer distribution in the tissues can be assumed to be mostly unaltered throughout the acquisition duration. However, the subject's motion will result in spatial displacement of tissues. For periodic motion, the displacement functions that define the motion from any one gate to the other gates can be assumed to be identical for each cycle. Thus, the radiotracer distribution as a function of time can be implicitly defined from its distribution for one gate, named the reference gate. One of the most studied motion compensation approach pre-computes these displacement functions to model the link between the gated sinogram frames and the radiotracer distribution of the reference gate [2]. We dubbed this approach 3.5D since it explicitly computes a 3D image by using a model that takes advantage of data acquired in 4D (three spatial coordinates + one motion cycle). A key aspect of this model is to consider that the random processes defining the data acquisition are statistically independent along the motion axis. In the classical 3.5D model, this hypothesis is used to define the quality of an estimate as the average of the correspondence of each implicitly defined gated estimate to its gated sinogram.

Here, we propose to investigate a modified version of the 3.5D approach that is built from the previously stated hypothesis and the property of the independent sum of Poisson processes. By using them together, the quality of a reference gate estimate can be characterized as the correspondence between the sum of the implicitly defined gated estimates and the sum of gated sinograms. It will be shown, using qualitative and quantitative comparisons of images obtained by the two models, that images of comparable quality can be reconstructed, at a reduced computational cost, by using this model. An equivalent model was used in [3]. Our contributions reside in the numerical and theoretical demonstration of its validity.

2 Motion compensation models

The following motion compensation models are presented in their basic formulation in order to highlight their differences.

Data acquired in PET, also called events, are pairs of anti-parallel annihilation photons emitted following the decay of radioactive nuclei through positron emission. Let y_j^g , with $j \in [1, J]$ and $g \in [1, G]$, be the gated sinograms bin representing the number of events observed by the j^{th} detector pair at gate g . Its corresponding random variable, denoted Y_j^g , is a Poisson process [4]. The likelihood that the gated sinograms $\{y_j^g\}^{J,G}$ were produced by a 3D radiotracer distribution λ is defined as $\mathcal{L}(\lambda) := \Pr_\lambda(\{Y_j^g\}^{J,G} = \{y_j^g\}^{J,G})$. Its maximum, noted as λ^* , should be the true radiotracer distribution.

Let λ_i^{g*} denote the true number of events originating from voxel $i \in [1, I]$ at gate g and P be the system matrix with P_{ji} defining the probability of an event originating from the voxel i to be detected by the detector pair j . Hence, the expected value of Y_j^g is $[P\lambda^{g*}]_j$. These definitions flow directly from the motionless model applied to each gated sinogram [4].

For simplicity, let λ be the reference gate image. We assume that for any choice of reference gate and $\forall g \in [1, G]$, there exists a deformation matrix W^g such that $\lambda^{g*} = W^g \lambda^*$. If the $\{W^g\}^G$ are known, then $Y_j^g \sim \text{Poisson}([PW^g \lambda^*]_j)$. The key step to build the 3.5D models is to assume that the random variables $\{Y_j^g\}^{J,G}$ are all independent.

2.1 The classical model: 3.5D

The independence hypothesis is used to define the joint probability as a product resulting in the following equivalence: $\Pr_\lambda(\{Y_j^g\}^{J,G} = \{y_j^g\}^{J,G}) = \prod_{j,g} \Pr_\lambda(Y_j^g = y_j^g)$. From there, we substitute the Poisson density functions, apply the natural logarithm on the result and remove the additive terms that do not depend on λ . Thus, the objective function to be maximized for the classical model is:

$$\mathcal{L}(\lambda; W) = -\sum_{g,j} [P^g \lambda]_j + \sum_{g,j} y_j^g \ln([P^g \lambda]_j) \quad (1)$$

with $P^g = PW^g$ [2]. Applying the MLEM technique to equation (1) yields the following iterative scheme:

$$\lambda^{\text{new}} = \left(\lambda \oslash \sum_g (P^g)' \mathbf{1} \right) \otimes \left(\sum_g (P^g)' (y^g \oslash P^g \lambda) \right) \quad (2)$$

where $\mathbf{1}$ is a column vector of 1 with J elements and $\oslash$ and $\otimes$ are respectively the element wise division and multiplication.

2.2 The new model: 3.5D_Merge

In this case, the independence hypothesis and the sum property of independent Poisson processes are used to define $\{\bar{Y}_j\}^J$ where $\bar{Y}_j \sim \text{Poisson}([\sum_g PW^g \lambda^*]_j)$. This implies that $\Pr_\lambda(\{Y_j^g\}^{J,G} = \{y_j^g\}^{J,G}) = \prod_j \Pr_\lambda(\bar{Y}_j = \bar{y}_j)$. By applying the same steps as for the classical model, we obtain the objective function of the new model:

$$\mathcal{L}(\lambda; \bar{W}) = -\sum_j [\bar{P} \lambda]_j + \sum_j \bar{y}_j \ln([\bar{P} \lambda]_j) \quad (3)$$

where $\bar{P} = P\bar{W}$, $\bar{W} = \sum_g W^g$ and $\bar{y} = \sum_g y^g$. Applying the MLEM technique to equation (3) yields the following iterative scheme:

$$\lambda^{\text{new}} = (\lambda \oslash (\bar{P})' \mathbf{1}) \otimes ((\bar{P})' (\bar{y} \oslash \bar{P} \lambda)) \quad (4)$$

3 Comparison of the models

We show that the 3.5D_Merge model, in contrast to the 3.5D model, uses roughly G times less computational resources per iteration and can only warrant a concave objective function.

3.1 Per-iteration computational cost

The 3.5D_Merge model uses only one deformation matrix, $\bar{W}$, and one dataset, $\bar{y}$, while the 3.5D model explicitly preserves each $\{W^g\}^G$ and $\{y^g\}^G$, thus resulting in a G -fold saving in memory space. The actual memory saved from merging the $\{W^g\}^G$ is dependent on their sparsity structure, typically very close to a G -fold memory reduction.

Matrix-vector multiplications represent the most expensive operation in a PET iterative scheme. Consequently, its computational cost can be summarized as the number of time that the system matrix is manipulated. Thus, the 3.5D MLEM computational cost (equation (2)) increases linearly with G , while it is constant for the 3.5D_Merge MLEM (equation (4)). This observation holds for most PET iterative schemes since gradient evaluations require system matrix projections.

3.2 Concavity of their likelihood function

The Hessian of the 3.5D and 3.5D_Merge models are

$$\nabla^2 \mathcal{L}(\lambda; W) = -\sum_g (P^g)' \text{diag} \{y^g \oslash (P^g \lambda)^2\} P^g \quad (5)$$

$$\nabla^2 \mathcal{L}(\lambda; \bar{W}) = -(\bar{P})' \text{diag} \{\bar{y} \oslash (\bar{P} \lambda)^2\} \bar{P} \quad (6)$$

where $(v)^2 = v \otimes v$ and $\text{diag} \{v\}$ is a diagonal matrix constructed from a vector v . Without loss of generality, we define $g = 1$ as the reference gate. Thus, W^1 is the identity matrix.

For $g \in [1, G]$, we assume that $y^g > \vec{0}$ and $W^g \lambda > \vec{0}$. If P is full column rank, we have $y^g \oslash (PW^g \lambda)^2 > \vec{0}, \forall \lambda > \vec{0}$.

3.5D: Let $f^g = P^g x$ with $x \in \mathbb{R}^I$.

$$\begin{aligned} x' \nabla^2 \mathcal{L}(\lambda; W) x &= -\sum_g \|f^g\|_{\text{diag}\{y^g \odot (P^g \lambda)^2\}}^2 \\ &\leq -\|f^1\|_{\text{diag}\{y^1 \odot (P^1 \lambda)^2\}}^2 < 0 \end{aligned}$$

Hence, (5) is negative definite and as such the 3.5D model possesses a unique global maximum.

3.5D_Merge: We define $\bar{f} = \bar{P}x$ with $x \in \mathbb{R}^I$.

$$\begin{aligned} x' \nabla^2 \mathcal{L}(\lambda; \bar{W}) x &= -x' (\bar{P})' \text{diag}\{\bar{y} \odot (\bar{P} \lambda)^2\} \bar{P} x \\ &= -\|\bar{f}\|_{\text{diag}\{\bar{y} \odot (\bar{P} \lambda)^2\}}^2 \leq 0 \end{aligned}$$

We cannot, without more information, ensure that $\bar{W}x = x + \sum_{g \neq 1} W^g x \neq \vec{0}$. Thus, (6) is at least negative semi-definite and this model could have multiple global maxima.

4 Experimental setup

Our aim is to provide qualitative and quantitative evidence that the 3.5D_Merge model provides images at least as good as those obtained by the 3.5D model. We will use the best image estimate of the reference gate obtained from each model as the comparison basis.

4.1 Data acquisition

Experimental and simulated datasets were used in this study. The *real* dataset was acquired with a LabPET8 small animal PET camera [5] whereas the *simulated* dataset was generated from the 3D system matrix used for the *real* reconstructions. P is a full column rank matrix since it was defined with $J = 2\,526\,720$ detector pairs and $I = 541\,596$ image voxels.

The *real* acquisition consisted of a 12.5-minute scan of the left ventricle of a healthy rat, 30 minutes after injection of ^{18}F -FDG. The start of each heart cycle was detected by using an electrocardiograph. The events were distributed into eight sinograms corresponding to eight uniformly divided time frames of each heart cycle. A normalization scan was performed to obtain the detector efficiency correction factors.

The phantoms used for the creation of the *simulated* sinograms were defined as a simple focaloid in a uniform background. The focaloid position, size and intensity through the gated sequence were modeled on the left ventricle observed in the 3D reconstructions of the *real* sequence. The background intensity was chosen so that the data counts in the *simulated* and *real* datasets would be similar.

The *simulated* sinograms were obtained in two steps. First, the theoretical sinogram of each gate was generated by multiplying its phantom with the system matrix. Then, Poisson noise, generated by the method *gsl_ran_poisson* from the *gsl* library [6], was added to the sinograms. A different seed was chosen for each phantom of the sequence.

4.2 Comparison context and methods

To obtain the best estimate with each model, we used the algorithms defined in equations (2) and (4). However, likelihood models are known to over-fit noise-contaminated dataset [7]. Hence, a stopping criteria must be used to find the “true” best image for a given dataset and model. The estimate with maximum cross-correlation (CC) to the ground truth of the reference gate image was used to fulfill this task. The CC was restricted to a cubic region of interest encompassing the left ventricle. The 3D estimates are used as a baseline for the comparison.

By construction, the ground truth of the *simulated* sequence is known. Thus, we can numerically compare the quality of the estimates produced by each algorithm. However, the *real* sequence ground truth is unknown. It was approximated by averaging the 3D reconstructions of gates 1, 7 and 8, which are visually in the same

motion state. The number of iterations applied for the 3D *real* reconstructions was the number of iterations needed to obtain the 3D reference gate estimate with the smallest sum of voxel intensities. This criteria was chosen by observing that, for the *simulated* reconstructions, it was almost equivalent to taking the estimate that has the maximum correspondence to the ground truth [8, A.12]. Since *real* images are defined with complex structures, we rely on visual assessment for their comparisons.

4.3 Deformation matrices

The deformation matrices of the two models were estimated from the 3D reconstructions of the gated sinograms. These estimates were evaluated by a script built from ITK [9]. The motion was defined as a non-rigid deformation model discretized by 25 uniformly distributed B-splines of order 3 in each spatial dimension. The mean squared error was used as the objective function and B-splines of order 3 were used for the interpolation functions. The resulting optimization problem was solved using the L-BFGS algorithm of ITK.

Due to partial volume effect, non-rigid registration of PET images sometimes results in suboptimal deformation [10]. To minimize its impact on this experiment, the gates 3 and 4, which are the farthest to the reference gate in term of motion, were removed from the *real* reconstructions.

5 Results

For the *simulated* acquisition, the 3.5D models produce estimates of equivalent CC, which are notably closer to the ground truth than those produced by the 3D model (Figure 1). The number of iterations needed to attain the best solutions for the 3.5D models is much larger than with the 3D.

From Figure 2, we can observe that the 3.5D_Merge algorithm reaches its best solution almost seven times faster than the 3.5D algorithm, while still requiring more resources than the 3D MLEM.

As for the *real* reconstructions, visual inspection of the best estimate of each model indicates that the classical 3.5D model creates a slightly sharper image than the 3.5D_Merge model, even though it used only half the iterations (Figures 3c and 3d). They both reproduce the features observed in the ground truth approximation (Figure 3a), while having higher contrast and sharper edges than the best 3D MLEM estimate (Figure 3b). However, some differences in the radiotracer distribution along the left ventricle wall can be noticed.

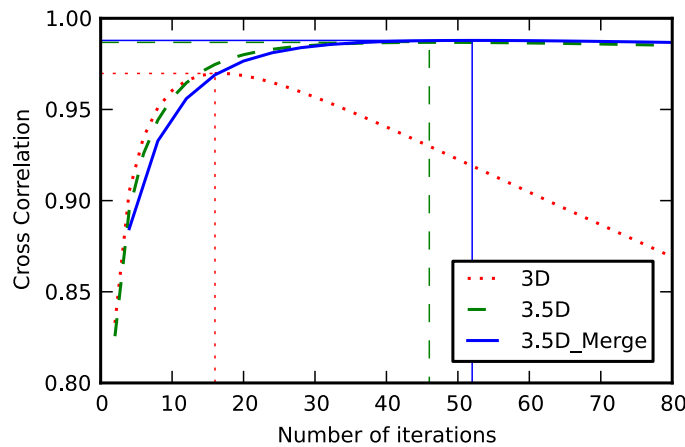


Figure 1: Comparison of the 3D, 3.5D and 3.5D_Merge models. Each curve represents the CC between the estimates produced by each MLEM algorithm applied to the *simulated* dataset and its ground truth as a function of the number of iterations used. The straight lines indicate their respective best estimate.

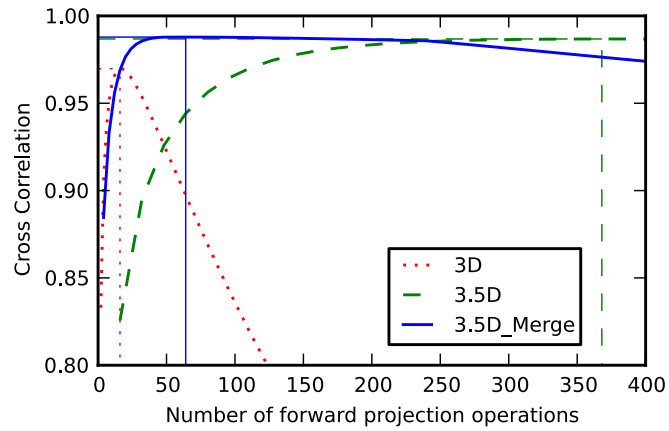
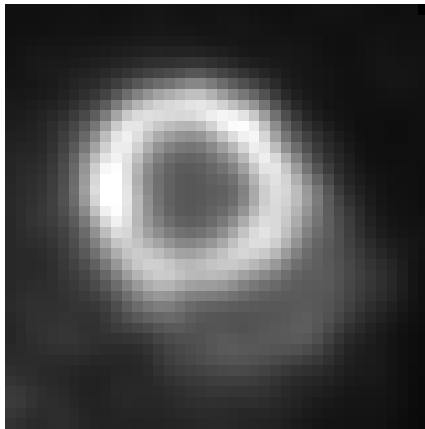
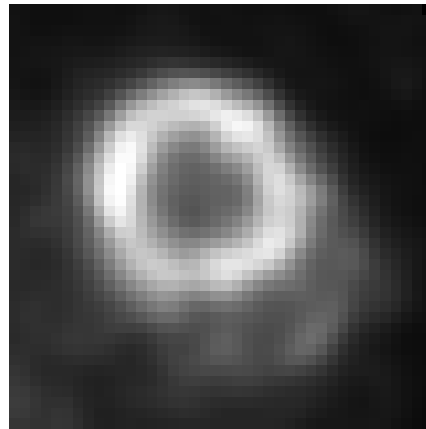


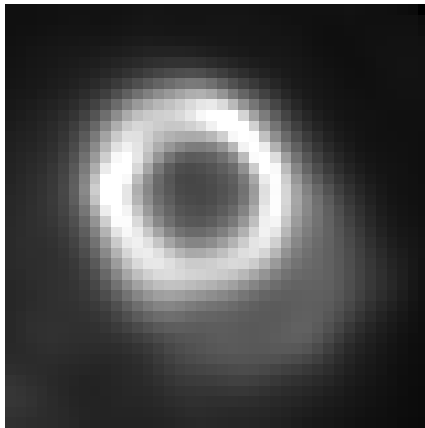
Figure 2: Comparison of the 3D, 3.5D and 3.5D_Merge models. Each curve represents the CC between the estimates produced by each MLEM algorithm applied to the *simulated* dataset and its ground truth as a function of the number of system matrix projections. The straight lines indicate their respective best estimate.



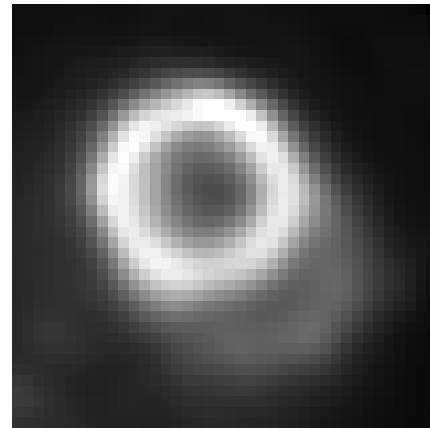
(a) Mean of the 3D obtained from gates 1, 7 and 8 each with 14 iterations



(b) 3D, 14 iterations



(c) 3.5D, 18 iterations



(d) 3.5D_Merge, 30 iterations

Figure 3: Comparison of the best reconstructions of four methods applied on the *real* dataset. In all cases a 2D slice of the left ventricle is shown using the same grey scale.

6 Discussion

The 3.5D and 3.5D_Merge MLEM are very similar, only differing by how the correction factors are evaluated in the projection space (the part with the second element-wise division in equations (2) and (4)). In the 3.5D model, these factors are evaluated for each gated sinogram and back-projected independently, with $(P^g)'$, whereas they are considered globally in the proposed 3.5D_Merge model. Both models achieve better performance than the *de facto* 3D model, which can be expected since the 3.5D approaches use all of the gated sinograms in their model. The results obtained from the *simulated* acquisition indicate that the two models produce very similar estimates. With the *real* dataset it is not straightforward to assess the quality of the best estimate for each model since the ground truth is unknown. However, our results indicate that the new model is a proper candidate for the 3.5D approach.

As stated previously, the 3.5D_Merge model uses significantly less system matrix projection operations per-iteration than the 3.5D model. With the current experimental setup, it represents an eight-fold, or six-fold in the *real* reconstruction case, reduction of computing resources. If we consider the total computing cost of each reconstruction made in our experiment, which is the product of the number of iterations and the cost per-iteration of the algorithm, we can conclude that the 3.5D_Merge is a considerably less expensive alternative to the classical approach. This would be especially advantageous when using finer discretization of the motion or incorporating another cyclic motion e.g. breathing motion. However, the gain in memory requirement might become negligible if the $\{W^g\}^G$ must be computed on the fly.

From a theoretical stand point, the possible lack of a unique optimal solution for the 3.5D_Merge model, as shown in Section 3.2, is a source of concern. On the other hand, the concavity of equation (3) ensures that the set of its optimal solutions is convex. Thus, it might not have much impact on the quality of PET reconstructions since the iterative process is often stopped way before asymptotic convergence. We believe that the properties inherent to matrices defining invertible motion are the key to warrant strict concavity. Also, it can be conjectured that this model is more robust in presence of low counts, due to the way correction factors are computed. A comparison of the spectral radius of their Hessian matrices would provide more insight into this property.

7 Conclusion

The 3.5D_Merge model was obtained by re-interpreting the independence hypothesis of the projection data along the gating axis. By exploiting the interdependence of the parameters defining the stochastic processes behind the gated sinograms, the resulting iterative scheme yields a single merged sinogram. Numerical and qualitative comparisons of this model to the classical 3.5D model indicate that the 3.5D_Merge model yields equivalent image accuracy and quality at a reduced cost in computational resources.

References

- [1] A. Rahmim, O. Rousset, and H. Zaidi, Strategies for motion tracking and correction in PET, *PET Clinics*, 2(2) 251–266, 2007.
- [2] T. Li et al., Model-based image reconstruction for four-dimensional PET, *Medical Physics*, 33(5) 1288–1298, 2006.
- [3] M Reyes et al., Model-based respiratory motion compensation for emission tomography image reconstruction, *Physics in Medicine and Biology*, 52(12) 3579, 2007.
- [4] K. Lange and R. Carson, EM reconstruction algorithms for emission and transmission tomography, *Journal of Computer Assisted Tomography*, 8(2) 306–16, 1984.
- [5] M. Bergeron et al., Performance evaluation of the LabPET APD-based digital PET scanner, *IEEE Transactions Medical Imaging*, 56(1) 10–16, 2009.
- [6] M. Galassi et al., GNU scientific library reference manual-(v1. 12), Network Theory Ltd, 2009.
- [7] T. J. Herbert, Statistical stopping criteria for iterative maximum likelihood reconstruction of emission images, *Physics in Medicine and Biology*, 35(9) 1221, 1990.

-
- [8] M. Toussaint, Reconstruction multidimensionnelle de type mouvement cyclique en TEP: revue et nouvelle approche, M. Sc. Thesis, Université de Sherbrooke, Sherbrooke, 2015.
 - [9] L. Ibáñez and Insight Software Consortium, The ITK software guide: updated for ITK v2.4, Kitware, 2005.
 - [10] D. Tenbrinck et al., Motion correction in positron emission tomography considering partial volume effects in optical flow estimation, in Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on, 1233–1236, 2010.