Direct 4D PET reconstruction with discrete tissue types

Michele Scipioni

***Abstract*— Dynamic positron emission tomography (dPET) is known for its ability to extract spatiotemporal information of a radio tracer in living tissue. In this paper, a novel direct reconstruction framework is presented, which include concurrent clustering as a potential aid in addressing high levels of noise typical of voxel-wise kinetic modeling. Core assumption is that the imaged volume is formed by a finite number of different functional regions, and that voxel-wise time courses are determined by the functional cluster they belong to. Probabilistic Graphical Modeling (PGM) theory is used to describe the problem, and to derive the inference strategy. The proposed iterative estimation scheme provides concurrent estimate of kinetic parameter maps, activity images, and segmented clusters. Simulation studies and exploratory application to real data are performed to validate the proposal.**

I. INTRODUCTION

Positron emission tomography (PET) is a molecular imaging modality enabling measurements of radio tracer distributions *in vivo*. In addition to conventional static acquisitions, dynamic scans allow to measure the kinetics of the radiotracer in living tissues, and to follow biological processes to better understand metabolism and drug action. Dynamic PET scans are complex spatiotemporal processes [1], which represent a major challenge for analysis, as they require to simultaneously handle spatial and temporal information.

The conventional (*indirect*) approach to estimate parametric maps from 4D PET data begins with the independent reconstruction of a sequence of 3D emission images from dynamic projection data, followed by the voxel-wise application of a kinetic model (KM) to the reconstructed time-activity curves (TACs) [1]. Failing to incorporate in the reconstruction information from multiple time frames may lead to enhanced noise in dynamic activity images, which then propagates to parametric maps estimates [2], because of the low signal-to-noise ratio (SNR) of each TAC.

Several Bayesian methods have been proposed to tackle the ill-posedness of the problem of PET reconstruction through the introduction of prior models acting as regularization factors. Traditional choices for priors aim to enforce spatial constraints derived by local-neighborhood kernels or additional high resolution anatomic images. A different, post-reconstruction strategy [3] seeks to reduce noise in kinetic analysis by grouping together voxels with similar kinetics, and identifying the boundaries of each functional region. Clustering analysis using principle component analysis and supplemented by a mixed Gaussian model to help classification of noisy dataset [4] has been thoroughly investigated, as well as combined hierarchical and K-means cluster analysis [3], or constrained mixture representations [5]. All these solutions, however, are still ignoring the temporal correlation of the activity, while reconstructing different time frames.

With the *direct* 4D reconstruction [1], by contrast, the tracer kinetics can be estimated directly from the raw projection data, allowing better noise modeling and hence more accurate results in terms of images and kinetic maps. One major drawback is the increased complexity of the optimization algorithms, especially when dealing with nonlinear compartment models. Moreover, current models for direct estimation are based on a deterministic description of voxel-wise TACs captured by the chosen KM, considering the photon counting process as the only source of uncertainty.

In this work, a new modeling strategy is introduced, based on the key assumptions that (i) activity time course is subject to uncertainty even if the parameters of the underlying dynamic process are known, and that (ii) there exist finite states of tissue, each one characterized by different kinetics and time courses. This leads to a hierarchical Bayesian model, which is formulated using the formalism of Probabilistic Graphical Modeling (PGM) [6]. Describing all variables involved as random variables, the inference of the joint probability distribution arising from the model can be addressed using a gradient based algorithm for image reconstruction, parametric maps estimation, and clustering.

II. METHODS

A. Theory

In 4D PET imaging, both activity image $\boldsymbol{x} = \{x_{jm}\}$ and measured counts $\boldsymbol{y} = \{y_{im}\}$ are functions of time, and organized into multiple consecutive time frames, $m \in \{1, \dots, M\}$, each containing all the events detected in a fixed time interval. Let $j \in \{1, \dots, J\}$ be the voxel index, and $i \in \{1, \dots, I\}$ the index of a sinogram line of response (LOR). The geometry of the acquisition system and the attenuation determine the probability $\boldsymbol{P} = \{p_{ij}\}$ of a photon emitted by voxel j being detected by LOR i. Approximating p_{ij} as time-invariant, the probability to observe counts y_{im} in detector bin i, given activity $\boldsymbol{x}_m$ at time frame m is:

$$p(y_{im}|\boldsymbol{x}) = Poisson\left(\sum_{j=1}^{J} p_{ij}x_{jm}; y_{im}\right) \qquad (1)$$

Let us now introduce the hypothesis that there exist finite states of tissue, each characterized by different dynamics. Therefore, each voxel-wise TAC is modeled as a sample from a Gaussian Mixture of G different classes, with voxels within the same functional region sharing similar kinetic behavior, and variations only due to noise:

Michele Scipioni is with the Department of Information Engineering at the University of Pisa (DII-UNIPI, Pisa, Italy) and with the National Research Council Institute of Clinical Physiology (CNR-IFC, Pisa, Italy).

$$p(x_{j:}|\boldsymbol{z},\boldsymbol{\theta}) = \prod_{m=1}^{M} \mathcal{N}\Big(x_{jm}; \mu_{gm}(\theta_{g:}), \sigma_{gm}(\theta_{g:})\Big) \quad (2)$$

Let $\boldsymbol{z} = \{z_{jg}\}$ be a set of auxiliary hidden random variables, such that the hidden state (i.e. cluster membership) of voxel j is expressed with a one-of-G representation: $\sum_{g=1}^{G} z_{jg} = 1$ and $z_{jg} \in \{0,1\}$. The prior probability $p(z_{jg} == 1)$ of voxel j belonging to tissue g when no observation of its time course is available is determined by the hidden states of its neighbors C_j, i.e. $p(z_{jg} == 1|z_{C_j g})$.

Let $f(\boldsymbol{\theta}; t)$ represent a generic kinetic model function, with $\boldsymbol{\theta} = \{\theta_{gp}\}$ vector of model parameters. The mean $\boldsymbol{\mu}_g$ and standard deviation $\boldsymbol{\sigma}_g$ of cluster g are expressed as functions of $f(\theta_{g:}; t)$. The cluster variance $\boldsymbol{\sigma}_g$ represents the uncertainty about the ability of the kinetic model to completely capture the variability of activity time courses, even when parameters $\boldsymbol{\theta}$ and membership labels $\boldsymbol{z}$ are known. In this work, no prior assumption is imposed on the values of the kinetic parameters, i.e. $p(\theta) \sim \mathcal{U}(0, \infty)$.

The probabilistic graph in Fig. 1 depicts the causal relationship between the random variables involved in this model: final goal is to infer the value of three latent variables, namely the kinetic parameter vector $\boldsymbol{\theta}$, the discrete tissue state $\boldsymbol{z}$ for each voxel, and the dynamic activity image $\boldsymbol{x}$, given the observation of projection data $\boldsymbol{y}$. This can be achieved by maximizing the following joint distribution:

$$p(y, x, \theta, z) = p(\theta)p(z)p(x|z,\theta)p(y|x) \ . \quad (3)$$

B. Algorithm

The joint optimization in (3) can be performed by alternating the optimization of the marginal posterior of each variable, conditioned to the provisional estimates of *all* the others. When applied to each node in the graph in turn, this procedure defines a single cycle of an iterative algorithm to update all the variables by alternating the following steps:

$$z^{(n+1)} = \arg\max_{z} p(z|x,\theta,y)\big|_{x^{(n)},\ \theta^{(n)}} \quad (4)$$

$$\theta^{(n+1)} = \arg\max_{\theta} p(\theta|x,z,y)\big|_{x^{(n)},\ z^{(n+1)}} \quad (5)$$

$$x^{(n+1)} = \arg\max_{x} p(x|z,\theta,y)\big|_{z^{(n+1)},\ \theta^{(n+1)}} \quad (6)$$

Two Expectation-Maximization (EM) algorithms are used to maximize (i) the conditional posterior of the parameters of the mixture and (ii) the conditional probability of activity time series $\boldsymbol{x}$ (one-step-late (OSL) variant of MLEM [7]).

1) Updating the mixture components, given $\boldsymbol{x}$ and $\boldsymbol{\theta}$: From Fig. 1, $\boldsymbol{z}$ and $\boldsymbol{y}$ are conditionally independent, given $\boldsymbol{x}$ and $\boldsymbol{\theta}$: the hidden state z_{jg} of each voxel can then be updated maximizing the conditional posterior of the Gaussian mixture model, $p(z|x,\theta) = p(z)p(x|z,\theta)/p(x)$, instead of (4):

$$\alpha_{jg}^{(n+1)} = \frac{\pi_{jg}^{(n)} \prod_m \mathcal{N}\left(x_{jm}^{(n)}, \mu_{gm}^{(n)}, \sigma_{gm}^{(n)}\right)}{\sum_{g'} \pi_{jg'}^{(n)} \prod_m \mathcal{N}\left(x_{jm}^{(n)}, \mu_{g'm}^{(n)}, \sigma_{g'm}^{(n)}\right)} \ , \quad (7)$$

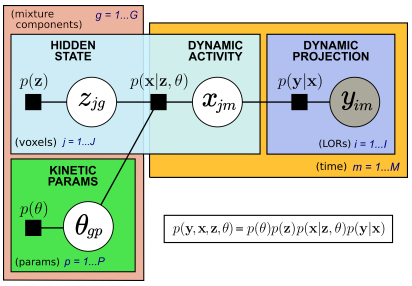

Fig. 1. Probabilistic Graphical Model representation of 4D PET imaging, with voxel-wise TACs related to a hidden kinetic state, and mean activity of each cluster function of a set of kinetic parameters. White background circles are latent random variables; gray background circle is the observed variable (raw counts); black squares represent probability distributions used to model each variable; colored plates convey dimensionality information.

where, for easier readability, α_{jg} symbolizes the posterior probability $p(z_{jg} == 1|x,\theta)$, and π_{jg} the prior probability $p(z_{jg} == 1)$. Values for all variables other than the conditional posterior α_{jg} are fixed at their provisional estimate.

2) Updating the kinetic parameters $\boldsymbol{\theta_g}$ for each cluster $\boldsymbol{g}$: To do this, the most likely (i.e. mean) time course needs to be computed first for each class g:

$$\hat{\mu}_{gm}^{(n+1)} = \frac{\sum_j \alpha_{jg}^{(n+1)} x_{jm}^{(n)}}{\sum_j \alpha_{jg}^{(n+1)}} \ . \quad (8)$$

Then, the new estimate of the hidden state $\boldsymbol{z}$ is frozen, and the estimate of kinetic parameters $\boldsymbol{\theta}$ is updated by fitting the chosen kinetic model to each curve $\hat{\boldsymbol{\mu}}_g$, from (8):

$$\boldsymbol{\theta}_g^{(n+1)} = \arg\min_{\boldsymbol{\theta}_g} \sum_m \|\hat{\mu}_{gm}^{(n+1)} - f(t_m, \boldsymbol{\theta}_g)\|_{\boldsymbol{W}}^2 \ . \quad (9)$$

This minimization can be done using any weighted non-linear least-squares method. Here, a gradient descent with Levenberg-Marquardt (LM) pre-conditioning was used [8]. Cluster means $\boldsymbol{\mu}$ and variances $\boldsymbol{\sigma}$ can, therefore, be expressed as function of $\boldsymbol{\theta}^{(n+1)}$:

$$\begin{aligned} \mu_{gm}^{(n+1)} &= f(\boldsymbol{\theta}_g^{(n+1)}; t_m) \ , \\ \sigma_{gm}^{2(n+1)} &= \frac{\sum_j \left[\alpha_{jg}^{(n+1)} \left(x_{jm}^{(n)} - \mu_{gm}^{(n+1)}\right)^2\right]}{\sum_j \alpha_{jg}^{(n+1)}} \ . \end{aligned} \quad (10)$$

3) Updating the activity time series $\boldsymbol{x}$, given kinetic and mixture parameters: This can be done maximizing (3) marginalized with respect to everything except for $\boldsymbol{x}$: $p(x|z,\theta,y) \propto p(x|z,\theta)p(y|x)$, where $p(y|x)$ is the Poisson data-fidelity term (1) and $p(x|z,\theta)$ is a prior term which incorporates information about the hidden functional states and kinetic modeling. The OSL [7] method can be used:

$$x_{jm}^{(n+1)} = \frac{x_{jm}^{(n)}}{\sum_i p_{ij} + \dfrac{\partial \log p(x|z,\theta)}{\partial x_{jm}^{(n)}}} \sum_i p_{ij} \frac{y_{im}}{\sum_{j'} p_{ij'} x_{j'm}^{(n)}} \ , \quad (11)$$

with:

$$\log p(x|z,\theta) = \sum_{j,m,g} \alpha_{jg}^{(n+1)} \frac{\left(x_{jm}^{(n)} - \mu_{gm}^{(n+1)}\right)^2}{\sigma_{gm}^{2\,(n+1)}} \,. \qquad (12)$$

This quadratic potential extends the idea of voxel neighborhoods to functionally homogeneous clusters, and σ_{gm}^2 provides an auto-tuning of the prior weight.

4) Updating the global multinomial probability of each cluster $\boldsymbol{g}$: Lastly, the prior probability π_{jg} of each voxel can be updated based on the hidden states of its neighbors:

$$\pi_{jg}^{(n+1)} = \frac{1}{\text{size}(C_j)} \sum_{k \in C_j} \alpha_{kg}^{(n+1)} \,. \qquad (13)$$

5) Determining the number of mixture components $\boldsymbol{G}$: The optimal number of clusters is usually an unknown of the problem. Here, we adopted an *on-line* update strategy, using a varying number of clusters until an optimum is reached. After each iteration, we check if it is possible to reduce the number of classes following criteria of *redundancy* (overlapping clusters) and *significance* (too sparse posterior map).

III. Experiments

A. Simulation

Dynamic [^{18}F]FDG PET scans were simulated for a Siemens Biograph mMR scanner using the geometric phantom in Fig. 2. The scanning schedule consisted of 24 time frames over 60 minutes: 4x20s, 4x40s, 4x60s, 4x180s, 8x300s. Regional TACs shown in Fig. 2 were generated according to an irreversible bi-compartmental model and assigned to different phantom regions. Resulting noise-free dynamic activity images were forward projected to simulate dynamic sinograms, and Poisson noise was generated producing an expected total number of events of ∼20 million.

B. Real Data

A brain dynamic [^{18}F]FDG PET scan was performed on the Siemens Biograph mMR scanner. The Institutional Review Board approved all experimental procedures involving human subjects. First 40 minutes of listmode raw data were binned into a total of 24 dynamic frames: 12x10s, 2x30s, 3x60s, 2x120s, 4x300s, 1x600s. The vendor software was used to extract the data correction matrices for each frame, including normalization factors, scattered and random counts estimates, and MR-based attenuation maps.

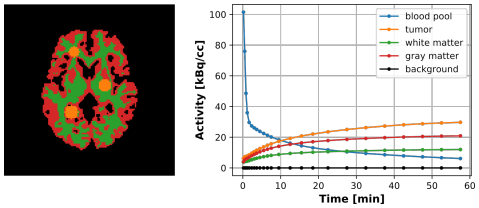

Fig. 2. Synthetic data used in simulation. (a) Digital phantom used in the simulation study, composed by four different regions mimicking the response of gray matter, white matter, and tumor tissue. (b) Input function and time-activity curves assigned to each region of the phantom.

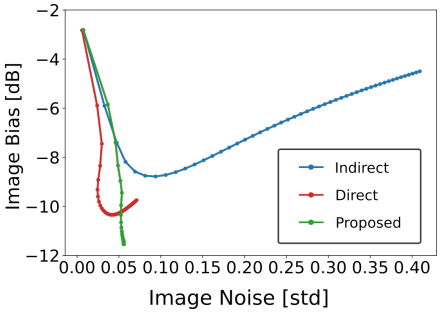

Fig. 3. Plot of the trade-off between bias [dB] and image noise for different reconstruction methods, as resulting from the simulation study.

IV. Results and Discussion

Simulated and real data were reconstructed independently by the proposed DirectCluster algorithm, plus a Direct [2] and an Indirect (OSEM + post-reconstruction fitting) alternatives, using a single subset of projections and 100 iterations.

Fig. 3 shows a comparison of the three algorithms, in terms of bias-vs-noise trade-off computed on the reconstructed simulated activity images over 100 iterations. The convergence rate of Direct (red) and DirectCluster (green) is slower than conventional OSEM (blue), because of the additional temporal correlations to account for, but this is compensated by a greater bias reduction. While both kinetic-informed methods can limit the increase of noise with reconstruction iterations, the achievement of an overall convergence in the estimate of the optimal number of kinetic clusters and of their mean time courses (after ∼30 iterations) allows the proposed DirectCluster method to keep the noise almost constant, while continuing reducing the bias.

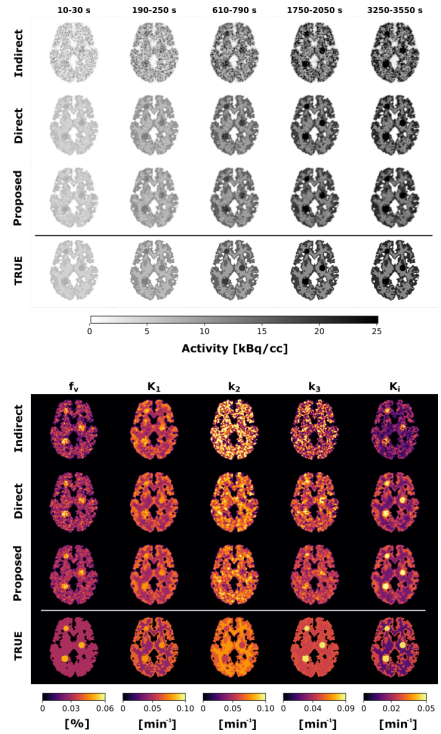

Fig. 4. Simulation study after 100 iterations. (top) Five time frames extracted from the whole time series. (bottom) Parametric maps of the 4 model's micro-parameters, plus the macro-parameter K_i.

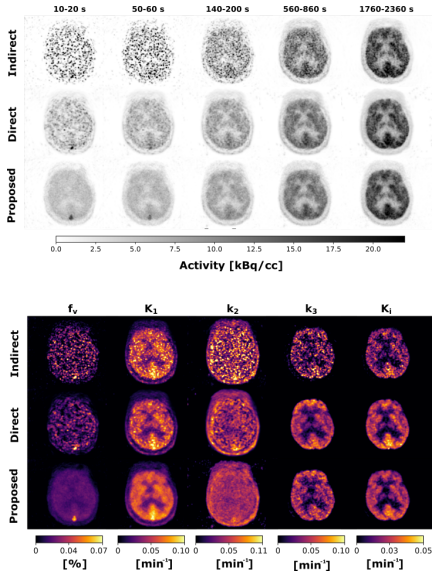

Fig. 5. Real [^{18}F]FDG human data, after 100 iterations. (top) Five example time frames with different length of acquisition. (bottom) Parametric maps of the four model micro-parameters, plus the macro-parameter K_i.

Fig. 4 compares the result after 100 iterations in terms of (top) activity time frames, $\boldsymbol{x}$, and (bottom) the kinetic maps, $\boldsymbol{\theta}$. Informing the reconstruction with knowledge coming from kinetic modeling has a clear impact on the quality of both images and parametric maps. The proposed DirectCluster method is able to compete, where not outperform, the quality of the maps produced by the direct method. It is especially interesting to notice the greater capability of recovering edges between different tissues, which is a direct effect of the clustering information encoded in the log-prior distribution (12) used to regularize the Poisson likelihood in (11).

Fig. 5 shows the same kind of images, as estimated from the reconstruction of the real dataset. It is noteworthy to see how a constraint that correlates the activity of successive time points allows to recover meaningful details in early frames: those are especially challenging for a frame-independent reconstruction approach, because of the low signal-to-noise ratio, due to short acquisition windows ($\Delta t \leq 10s$) required to follow the fast kinetics of the arterial phase. The limited number of cluster means used by the DirectCluster method to compute the kinetic-based prior allows for a great SNR enhancement, outperforming also the Direct approach.

Furthermore, the proposed method provides also an estimate of the posterior distribution $p(z|x,\theta)$ of the hidden states for each voxel j, resulting from (4) and (7). Each map in Fig. 6 shows the posterior probability of each voxel having a time course similar to one of the kinetic clusters identified in the volume, and they can be read also as cluster membership probability maps. A part from the active role these posterior maps play in the inference, we anticipate them being of help in understanding tissue kinetic behavior.

V. CONCLUSIONS

In this work a cluster-based direct reconstruction algorithm was developed and evaluated, which decouples the overall reconstruction problem allowing concurrent estimation

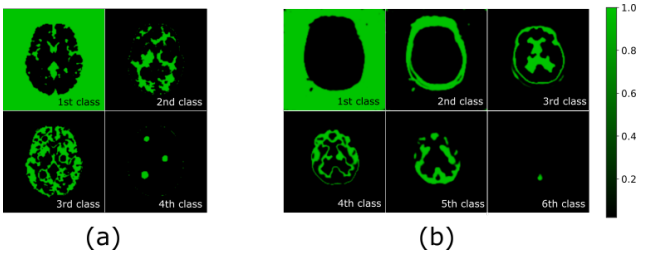

Fig. 6. Conditional posterior maps. (a) Simulation: background and zero activity; white matter; gray matter; tumors. (b) Patient data: background; non-brain tissue; CSF; white matter; gray matter; mostly blood tissue.

of tomographic images, kinetic parameters, and functional clusters. This approach is based on two critical assumptions. The first is that the imaged volume contains a finite number of tissue types, each with a unique kinetic behavior identified by a certain TAC. Then, each voxel-wise TAC is modeled as a sample from a Gaussian Mixture, with different physiological responses for each component. If one or more of these assumptions is not satisfied, then the performance of the proposed method may suffer. While the idea of using mixture representations is not new in the field of PET kinetic modeling, the aim of this work is rather to show how incorporating the clustering step within the reconstruction may assist the estimate of dynamic images and parametric maps. This idea presents several advantages compared to existing methods: it is simple to implement; it is flexible to an arbitrary choice of the kinetic model; it is faster than conventional direct methods, as the kinetic modeling is applied only to a few cluster means, at each iteration; and it grants meaningful noise reduction, via a boundary-aware averaging. Finally, work is in progress to extend the proposed method with a model selection strategy applied at each tomographic iteration, before model fitting: this would allow to select an optimal model for each cluster mean, simproving accuracy in maps estimation.

REFERENCES

[1] A. J. Reader and J. Verhaeghe, "4D image reconstruction for emission tomography," *Phys. Med. Biol.*, vol. 59, no. 22, pp. R371–R418, 2014.
[2] M. Scipioni, A. Giorgetti, D. D. Latta, S. Fucci, V. Positano, L. Landini, and M. F. Santarelli, "Direct parametric maps estimation from dynamic PET data: An iterated conditional modes approach," *J Heal. Eng*, vol. 21, p. 14, 2018.
[3] W. Zhu, J. Ouyang, Y. Rakvongthai, N. J. Guehl, D. W. Wooten, G. El Fakhri, M. D. Normandin, and Y. Fan, "A Bayesian spatial temporal mixtures approach to kinetic parametric images in dynamic positron emission tomography: Bayesian inference of kinetic model for PET image," *Med. Phys.*, vol. 43, no. 3, pp. 1222–1234, Feb. 2016.
[4] G. I. Angelis and S. R. Meikle, "Cluster-based Direct Estimation of Parametric Maps of Dopamine Response in Dynamic PET Data," in *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. Atlanta, GA: IEEE, Oct. 2017, pp. 1–3.
[5] F. O'Sullivan and A. Saha, "Use of ridge regression for improved estimation of kinetic constants from PET data," *IEEE Trans. Med. Imaging*, vol. 18, no. 2, pp. 115–125, 1999.
[6] C. M. Bishop, *Pattern Recognition and Machine Learning*, ser. Information Science and Statistics. New York: Springer, 2006.
[7] P. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imaging*, vol. 9, no. 1, pp. 84–93, Mar. 1990.
[8] M. Scipioni, A. Giorgetti, D. D. Latta, S. Fucci, V. Positano, L. Landini, and M. F. Santarelli, "Accelerated PET kinetic maps estimation by analytic fitting method," *Comput. Biol. Med.*, vol. 99, pp. 221–35, 2018.