Joint Parametric Reconstruction and Motion Correction Framework for Dynamic PET Data

Jieqing Jiao[1,⋆], Alexandre Bousse[2,⋆], Kris Thielemans[2], Pawel Markiewicz[1], Ninon Burgos[1], David Atkinson[3], Simon Arridge[1], Brian F. Hutton[2], and Sébastien Ourselin[1,4]

[1] Translational Imaging Group, Centre for Medical Image Computing, UCL, UK
[2] Institute of Nuclear Medicine, University College London, UK
[3] Centre for Medical Imaging, University College London, UK
[4] Dementia Research Centre, Institute of Neurology, University College London, Queen Square, WC1N 3BG, UK

Abstract. In this paper we propose a novel algorithm for jointly performing data based motion correction and direct parametric reconstruction of dynamic PET data. We derive a closed form update for the penalised likelihood maximisation which greatly enhances the algorithm's computational efficiency for practical use. Our algorithm achieves sub-voxel motion correction residual with noisy data in the simulation-based validation and reduces the bias of the direct estimation of the kinetic parameter of interest. A preliminary evaluation on clinical brain data using [^{18}F]Choline shows improved contrast for regions of high activity. The proposed method is based on a data-driven kinetic modelling method and is directly applicable to reversible and irreversible PET tracers, covering a range of clinical applications.

Keywords: Dynamic PET, direct parametric reconstruction, motion correction, optimisation transfer, kinetic analysis.

1 Introduction

Dynamic Positron Emission Tomography (PET) imaging in conjunction with appropriate tracer kinetic models allow for the estimation of biological parameters that are essential for disease understanding, clinical diagnosis and drug development. Conventionally, the raw PET data in the form of photon counts recorded by the detectors are first reconstructed to provide the temporal images of the spatial distribution of the PET tracer (activity), and a selected kinetic model is then applied to the time activity curves on a voxel/region basis to derive the biological parameters of interest. However, subject motion and photon count statistics are two of the fundamental issues in PET imaging and, if not accounted for, may lead to significant errors in the kinetic quantification for any clinical decision making. To address the uncertainties in photon emissions, direct parametric reconstruction approaches [1,2] have been developed to derive

⋆ These authors contributed equally to this work.

P. Golland et al. (Eds.): MICCAI 2014, Part I, LNCS 8673, pp. 114–121, 2014.

the kinetic parameters directly from the raw PET data, with the incorporation of kinetic modelling. This improves the reconstruction by utilising the complete spatio-temporal information rather than inappropriately ignoring the temporal dependencies within the data. To address the subject motion, a framework is proposed in [3] to estimate subject motion in PET data with steady activity from simultaneous PET-MR reconstruction. In [4], the motion correction approach has been developed for reconstructed dynamic PET images using kinetic model based registration. To the best of our knowledge, so far there has not been any work yet to address both the kinetics and motion in a unified framework.

In this work, we propose a joint motion correction and parametric reconstruction framework for dynamic PET data, in which the subject motion and tracer kinetics are estimated directly from the raw photon counts, by maximising a penalised log-likelihood with respect to the motion parameters and the kinetic parameters. The problem of head movements in brain imaging is addressed here, thus a rigid transformation model is used and the motion parameters are updated by a trust region algorithm [5]. For modelling tracer kinetics, the data-driven basis function based approach is applied for generality, and we derive here a closed form update for the kinetic parameters in the penalised log-likelihood, by applying the optimisation transfer technique to decouple the dependencies between the voxels and also to separate each kinetic parameter for every voxel. This greatly improves the algorithm's convergence speed and also leads to a parallelisable solution of this problem. The simulation-based validation shows that the proposed method achieves a sub-voxel size motion correction residual in noisy data, and that the parametric reconstruction is improved with reduced bias. We also apply the proposed algorithm to real clinical [^{18}F]Choline data, and the results show that it enhances the detection of a [^{18}F]Choline hot spot in the brain which is otherwise difficult to identify due to motion and noise.

2 Method

Discretising the continuous 3-D+t PET activity $f(x,t)$, $x \in \mathbb{R}^3$, $t \in \mathbb{R}^+$, on a regular spatial grid $\boldsymbol{x} = [x_1, \ldots, x_{n_v}] \in \mathbb{R}^{n_v \times 3}$, and discrete temporal time $\boldsymbol{t} \in \mathbb{R}^{n_t}$ obtains $\boldsymbol{f} = f(\boldsymbol{x}, \boldsymbol{t}) \in \mathbb{R}^{n_v \times n_t}$, where n_v and n_t are the numbers of voxels and time frames respectively. Denote the PET system detected photon counts by $\boldsymbol{g} \in \mathbb{R}^{n_d \times n_t}$ where n_d is the number of detector pairs. Given the activity distribution $\boldsymbol{f}$, $\boldsymbol{g}$ is a Poisson random vector with parameter $\bar{\boldsymbol{g}}(\boldsymbol{f}) = \boldsymbol{P}\boldsymbol{f} + \boldsymbol{r}$ where $\boldsymbol{P} \in \mathbb{R}^{n_d \times n_v}$ is the system matrix *i.e.* $[\boldsymbol{P}]_{i,j} = p_{i,j}$ is the probability that an event occurring in voxel j is detected by the detector pair i, and $\boldsymbol{r} \in \mathbb{R}^{n_d \times n_t}$ is the expected scattered and random events. The dynamic activity $\boldsymbol{f} = [\boldsymbol{f}_1, \ldots, \boldsymbol{f}_{n_t}]$, $\boldsymbol{f}_l \in \mathbb{R}^{n_v}$ being the activity at frame l, can be described by the tracer kinetic model $\mathcal{F}$ parameterised by $\boldsymbol{\vartheta} = \vartheta(\boldsymbol{x})$, $\boldsymbol{\vartheta} \in \mathbb{R}^{n_v \times n_k}$ where n_k is the number of kinetic parameters in the model, so that $\boldsymbol{f} = \mathcal{F}(\boldsymbol{\vartheta}) = [\mathcal{F}_1(\boldsymbol{\vartheta}), \ldots, \mathcal{F}_{n_t}(\boldsymbol{\vartheta})]$, $\mathcal{F}_l(\boldsymbol{\vartheta}) \in \mathbb{R}^{n_v}$ for all l. The kinetic model is assumed to be applied at each voxel independently *i.e.* $[\mathcal{F}_l(\boldsymbol{\vartheta})]_j = \mathcal{F}_l(\boldsymbol{\vartheta}_j)$, $\boldsymbol{\vartheta}_j \in \mathbb{R}^{n_k}$ being the kinetic parameter vector at voxel j. When the subject motion is present, the activity $\boldsymbol{f}$ is warped.

The motion can be modelled as a warping operator $\mathcal{W}$ parameterised by $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{n_t}] \in \mathbb{R}^{n_w \times n_t}$ where n_w is the number of transformation parameters. Thus the activity at frame l is redefined as $\boldsymbol{f}_l = \mathcal{W}_{\boldsymbol{\alpha}_l}\mathcal{F}_l(\boldsymbol{\vartheta})$. The warping operator $\mathcal{W}_{\boldsymbol{\alpha}_l}$ is determined by the interpolation and its motion function $\varphi_{\boldsymbol{\alpha}_l} : \mathbb{R}^3 \to \mathbb{R}^3$, so that $\boldsymbol{f}_l = \mathcal{W}_{\boldsymbol{\alpha}_l}\mathcal{F}_l(\boldsymbol{\vartheta}) = \mathcal{F}_l \circ \vartheta(\varphi_{\boldsymbol{\alpha}_l}(\boldsymbol{x}))$. For kinetic model $\mathcal{F}$ and warping operator $\mathcal{W}$, the expected photon counts can be written as

$$\bar{\boldsymbol{g}}(\boldsymbol{\vartheta}, \boldsymbol{\alpha}) = [\boldsymbol{P}\mathcal{W}_{\boldsymbol{\alpha}_1}\mathcal{F}_1(\boldsymbol{\vartheta}), \ldots, \boldsymbol{P}\mathcal{W}_{\boldsymbol{\alpha}_{n_t}}\mathcal{F}_{n_t}(\boldsymbol{\vartheta})] + \boldsymbol{r}. \tag{1}$$

Omitting the constant terms, the log-likelihood of the detected events $\boldsymbol{g}$ is

$$L(\boldsymbol{g}|\boldsymbol{\vartheta}, \boldsymbol{\alpha}) = \sum\nolimits_{l=1}^{n_t} \sum\nolimits_{i=1}^{n_d} (g_{i,l} \log \bar{g}_{i,l}(\boldsymbol{\vartheta}, \boldsymbol{\alpha}_l) - \bar{g}_{i,l}(\boldsymbol{\vartheta}, \boldsymbol{\alpha}_l)),$$

where $\bar{g}_{i,l}(\boldsymbol{\vartheta}, \boldsymbol{\alpha}_l) = \sum_{j=1}^{n_v} p_{i,j}[\mathcal{W}_{\boldsymbol{\alpha}_l}\mathcal{F}_l(\boldsymbol{\vartheta})]_j + r_{i,l}$. Since the penalties $\boldsymbol{\vartheta}$ and $\boldsymbol{\alpha}$ are independent, the penalised log-likelihood (PL) function is $\Phi(\boldsymbol{\vartheta}, \boldsymbol{\alpha}) = L(\boldsymbol{g}|\boldsymbol{\vartheta}, \boldsymbol{\alpha}) - \beta U(\boldsymbol{\vartheta}) - \gamma V(\boldsymbol{\alpha})$, where $U(\boldsymbol{\vartheta})$ and $V(\boldsymbol{\alpha})$ are penalty functions, β and γ are hyperparameters. $\boldsymbol{\vartheta}$ and $\boldsymbol{\alpha}$ can then be estimated as $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\vartheta}}) \in \arg\max_{\boldsymbol{\vartheta}, \boldsymbol{\alpha}} \Phi(\boldsymbol{\vartheta}, \boldsymbol{\alpha})$. Here we find a maximum by updating $\boldsymbol{\vartheta}$ and $\boldsymbol{\alpha}$ in alternation.

Given the measured photon counts $\boldsymbol{g}$ and current estimation of kinetic parameters $\boldsymbol{\vartheta}$, $\boldsymbol{\alpha}$ is updated by maximising $E(\boldsymbol{\alpha}) = L(\boldsymbol{g}|\boldsymbol{\vartheta}, \boldsymbol{\alpha}) - \gamma V(\boldsymbol{\alpha})$. The gradient of log-likelihood part, $\nabla_{\boldsymbol{\alpha}} L$ can be derived by applying the chain rule, as $\frac{\partial L}{\partial \alpha_{q,l}} = \sum_{i=1}^{n_d} (\frac{g_{i,l}}{\bar{g}_{i,l}} - 1)\frac{\partial \bar{g}_{i,l}}{\partial \alpha_{q,l}}$, with $\frac{\partial \bar{g}_{i,l}}{\partial \alpha_{q,l}} = \sum_{j=1}^{n_v} p_{i,j} \langle \nabla_{\varphi_{\boldsymbol{\alpha}_l}(x_j)} \mathcal{F}_l \circ \vartheta(\varphi_{\boldsymbol{\alpha}_l}(x_j)), \frac{\partial \varphi_{\boldsymbol{\alpha}_l}(x_j)}{\partial \alpha_{q,l}} \rangle$ where $\langle \cdot, \cdot \rangle$ is the inner product in $\mathbb{R}^3$. Furthermore, an approximated second order Taylor expansion gives

$$L(\boldsymbol{\alpha} + \Delta\boldsymbol{\alpha}) \approx L(\bar{\boldsymbol{g}}(\boldsymbol{\vartheta}, \boldsymbol{\alpha})) + \nabla_{\bar{\boldsymbol{g}}} L^\top \boldsymbol{J}_{\boldsymbol{\alpha}}(\bar{\boldsymbol{g}})\Delta\boldsymbol{\alpha} + \frac{1}{2}\Delta\boldsymbol{\alpha}^\top \boldsymbol{J}_{\boldsymbol{\alpha}}^\top(\bar{\boldsymbol{g}})[\nabla_{\bar{\boldsymbol{g}}}^2 L]\boldsymbol{J}_{\boldsymbol{\alpha}}(\bar{\boldsymbol{g}})\Delta\boldsymbol{\alpha} \tag{2}$$

where $\boldsymbol{J}_{\boldsymbol{\alpha}}(\bar{\boldsymbol{g}}) \in \mathbb{R}^{(n_d \times n_t) \times (n_w \times n_t)}$ is the Jacobian of $\bar{\boldsymbol{g}}$ w.r.t. $\boldsymbol{\alpha}$ and $\Delta\boldsymbol{\alpha}$ is a small perturbation on $\boldsymbol{\alpha}$ (reshaped into a vector). Therefore the Hessian matrix can be approximated by $H_{\boldsymbol{\alpha}}(L) \approx \boldsymbol{J}_{\boldsymbol{\alpha}}^\top(\bar{\boldsymbol{g}})[\nabla_{\bar{\boldsymbol{g}}}^2 L]\boldsymbol{J}_{\boldsymbol{\alpha}}(\bar{\boldsymbol{g}})$. Adding the gradient and Hessian of V allows to maximise Φ with respect to $\boldsymbol{\alpha}$ using a trust region algorithm [5]. Here the rigid brain motion is addressed, and no prior V is considered on $\boldsymbol{\alpha}$.

With the estimated $\boldsymbol{\alpha}$, the direct parametric reconstruction with motion compensation is performed to update $\boldsymbol{\vartheta}$ as the maximiser of the penalised log likelihood $G(\boldsymbol{\vartheta}) = L(\boldsymbol{g}|\boldsymbol{\vartheta}, \boldsymbol{\alpha}) - \beta U(\boldsymbol{\vartheta})$. Each warping operator $\mathcal{W}_{\boldsymbol{\alpha}_l}$ is replaced by its discrete version $\boldsymbol{W}_{\boldsymbol{\alpha}_l} \in \mathbb{R}^{n_v \times n_v}$. The composition $\boldsymbol{P}\mathcal{W}_{\boldsymbol{\alpha}_l}$ in (1) becomes a matrix product $\boldsymbol{P}\boldsymbol{W}_{\boldsymbol{\alpha}_l} = \boldsymbol{P}^l$ which represents the motion-compensated system matrix at frame l. We considered a quadratic prior $U(\boldsymbol{\vartheta}) = \frac{1}{8}\sum_{l,j}\sum_{m \in \mathcal{N}_j} \omega_{j,m}(\mathcal{F}_l(\boldsymbol{\vartheta}_j) - \mathcal{F}_l(\boldsymbol{\vartheta}_m))^2$, where $\mathcal{N}_j$ denotes the 26 voxels neighbourhood of voxel j and $\omega_{j,m}$ is the weighting factor equal to the inverse distance between voxels j and m.

To separate the voxels in $G(\boldsymbol{\vartheta})$, the optimisation transfer technique used in [1] was first applied on $L(\boldsymbol{g}|\boldsymbol{\vartheta}, \boldsymbol{\alpha})$ and $U(\boldsymbol{\vartheta})$ to derive surrogate functions $Q^L(\boldsymbol{\vartheta}|\boldsymbol{\vartheta}^{[k]})$ and $Q^U(\boldsymbol{\vartheta}|\boldsymbol{\vartheta}^{[k]})$, where $\boldsymbol{\vartheta}^{[k]}$ is the estimation at iteration k. For the log-likelihood, the surrogate function is $Q^L(\boldsymbol{\vartheta}|\boldsymbol{\vartheta}^{[k]}) = \sum_l \sum_j p_j^l (f_{j,l}^{\mathrm{em},[k]} \log(\mathcal{F}_l(\boldsymbol{\vartheta}_j)) - \mathcal{F}_l(\boldsymbol{\vartheta}_j))$, where $f_{j,l}^{\mathrm{em},[k]} = \frac{\mathcal{F}_l(\boldsymbol{\vartheta}_j^{[k]})}{p_j^l} \sum_i p_{i,j}^l \frac{g_{i,l}}{\bar{g}_{i,l}(\boldsymbol{\vartheta}^{[k]})}$, $p_{i,j}^l = [\boldsymbol{P}^l]_{i,j}$ and $p_j^l = \sum_i p_{i,j}^l$. For

the quadratic penalty, $Q^U(\boldsymbol{\vartheta}|\boldsymbol{\vartheta}^{[k]}) = \frac{1}{2}\sum_{l,j}\omega_j(f_{j,l}^{\mathrm{reg},[k]} - \mathcal{F}_l(\boldsymbol{\vartheta}_j))^2$, where $\omega_j = \sum_{m\in\mathcal{N}_j}\omega_{j,m}$, and $f_{j,l}^{\mathrm{reg},[k]} = \frac{1}{2\omega_j}\sum_{m\in\mathcal{N}_j}\omega_{j,m}(\mathcal{F}_l(\boldsymbol{\vartheta}_j^{[k]}) + \mathcal{F}_l(\boldsymbol{\vartheta}_m^{[k]}))$. The resulting surrogate separates the voxels so that each $\boldsymbol{\vartheta}_j$ can be obtained independently by maximising $q_j(\boldsymbol{\vartheta}_j|\boldsymbol{\vartheta}^{[k]}) = q_j^L(\boldsymbol{\vartheta}_j|\boldsymbol{\vartheta}^{[k]}) - \beta q_j^U(\boldsymbol{\vartheta}_j|\boldsymbol{\vartheta}^{[k]})$, where $q_j^L(\boldsymbol{\vartheta}_j|\boldsymbol{\vartheta}^{[k]}) = \sum_l p_j^l(f_{j,l}^{\mathrm{em},[k]}\log(\mathcal{F}_l(\boldsymbol{\vartheta}_j)) - \mathcal{F}_l(\boldsymbol{\vartheta}_j))$ and $q_j^U(\boldsymbol{\vartheta}_j|\boldsymbol{\vartheta}^{[k]}) = \sum_l \frac{1}{2}\omega_j(f_{j,l}^{\mathrm{reg},[k]} - \mathcal{F}_l(\boldsymbol{\vartheta}_j))^2$.

Generally, maximising $q_j(\boldsymbol{\vartheta}_j|\boldsymbol{\vartheta}^{[k]})$ can be solved by a non-linear optimisation algorithm at each voxel j. Note that the activity depends on the form of the kinetic model $\mathcal{F}$, and both linear and non-linear models can be used. In this work, a closed-form solution to update $\boldsymbol{\vartheta}_j^{[k]}$ was derived with linear kinetic models $\mathcal{F}(\boldsymbol{\vartheta}) = \boldsymbol{\vartheta}\boldsymbol{B}$, where $\boldsymbol{B}\in\mathbb{R}^{n_k\times n_t}$ is the temporal basis matrix.

With $\mathcal{F}_l(\boldsymbol{\vartheta}_j) = \sum_q \vartheta_{j,q}b_{q,l}$, $b_{q,l} = [\boldsymbol{B}]_{q,l}$, the optimisation transfer can be applied again on the log part $q_j^L(\boldsymbol{\vartheta}_j|\boldsymbol{\vartheta}^{[k]})$ and the quadratic part $q_j^U(\boldsymbol{\vartheta}_j|\boldsymbol{\vartheta}^{[k]})$ to separate each of the n_k parameters $\vartheta_{j,q}\in\mathbb{R}^+$. A separable surrogate function for $q_j(\boldsymbol{\vartheta}_j|\boldsymbol{\vartheta}^{[k]})$ at sub-iteration r is

$$\begin{aligned}\tilde{q}_j(\boldsymbol{\vartheta}_j|\boldsymbol{\vartheta}_j^{[k,r]}) &= \sum_{q=1}^{n_k}\tilde{b}_{j,q}(\vartheta_{j,q}^{\mathrm{em},[k,r]}\log(\vartheta_{j,q}) - \vartheta_{j,q}) - \beta\frac{1}{2}\omega_j a_q\left(\vartheta_{j,q} - \vartheta_{j,q}^{\mathrm{reg},[k,r]}\right)^2 \\ &= \sum_{q=1}^{n_k}\ell_{j,q}^{[k,r]}(\vartheta_{j,q}),\end{aligned} \tag{3}$$

$\vartheta_{j,q}^{\mathrm{em},[k,r]} = \tilde{b}_{j,q}^{-1}\vartheta_{j,q}^{[k,r]}\sum_l \frac{p_j^l b_{q,l} f_{j,l}^{\mathrm{em},[k]}}{\mathcal{F}_l(\boldsymbol{\vartheta}_j^{[k,r]})}$ and $\vartheta_{j,q}^{\mathrm{reg},[k,r]} = \vartheta_{j,q}^{[k,r]} - \frac{\sum_l b_{q,l}(\mathcal{F}_l(\boldsymbol{\vartheta}_j^{[k,r]}) - f_{j,l}^{\mathrm{reg},[k]})}{a_q}$ are the intermediate parametric values at voxel j, $\tilde{b}_{j,q} = \sum_l p_j^l b_{q,l}$, and $a_q = \sum_l b_{q,l}\sum_{q'} b_{q',l}$. The quadratic surrogate was obtained using De Pierro's technique [6]. Note that $\vartheta_{j,q}^{\mathrm{em},[k,r]}$ corresponds to the nested EM update from [7] whereas $\vartheta_{j,q}^{\mathrm{reg},[k,r]}$ is a quasi-Newton update of $\vartheta_{j,q}^{[k,r]}$ to minimise $\sum_l(\mathcal{F}_l(\boldsymbol{\vartheta}_j^{[k,r]}) - f_{j,l}^{\mathrm{reg},[k]})^2$. Each $\ell_{j,q}^{[k,r]}$ in (3) has a unique maximiser which corresponds to the $r+1$ inner-update of the optimisation w.r.t. $\boldsymbol{\vartheta}$ with the current $f_{j,l}^{\mathrm{em},[k]}$ and $f_{j,l}^{\mathrm{reg},[k]}$:

$$\vartheta_{j,q}^{[k,r+1]} = \frac{1}{2\beta\omega_j a_q}\left(\beta\omega_j a_q\vartheta_{j,q}^{\mathrm{reg},[k,r]} - \tilde{b}_{j,q} + \sqrt{\Delta_{j,q}(\vartheta_{j,q}^{\mathrm{em},[k,r]}, \vartheta_{j,q}^{\mathrm{reg},[k,r]})}\right), \tag{4}$$

where $\Delta_{j,q}(\vartheta_{j,q}^{\mathrm{em},[k,r]}, \vartheta_{j,q}^{\mathrm{reg},[k,r]}) = (\beta\omega_j a_q\vartheta_{j,q}^{\mathrm{reg},[k,r]} - \tilde{b}_{j,q})^2 + 4\beta\omega_j a_q\tilde{b}_{j,q}\vartheta_{j,q}^{\mathrm{em},[k,r]}$. For given $\boldsymbol{\alpha}$, the penalised likelihood function is concave and a monotonic convergence is guaranteed for $\boldsymbol{\vartheta}$ [1]. The overall scheme is summarised in Algorithm 1.

To relate the linear coefficients $\boldsymbol{\vartheta}$ directly with the transport rate constants in the compartmental model to represent the tracer kinetics, the temporal basis $\boldsymbol{B}$ is defined as in spectral analysis [8], which is able to describe differing tracer behaviour across all voxels in the image, as $b_{q,l} = \int_{t_{l,s}}^{t_{l,e}} e^{-\phi_q\tau}\star C_p(\tau)d\tau$, where $\star$ is the convolution, $t_{l,s}$, $t_{l,e}$ are the starting and ending times of frame l, $\boldsymbol{\phi} = [\phi_1,\ldots,\phi_q,\ldots,\phi_{n_k}]$ are the pre-chosen kinetic rate constants from a physiologically plausible range, and $C_p(t)$ is the plasma input function. Note that if needed, the blood component can be added to the temporal basis to account for

the blood volume in the measured PET data as $b_{0,l} = \int_{t_{l,s}}^{t_{l,e}} C_B(\tau)d\tau$ where $C_B(t)$ is the activity in blood. To ensure the sufficient representation of the kinetics, usually the temporal basis $\boldsymbol{B}$ is over-complete. The under-determined problem is constrained by the non-negativity of $\boldsymbol{\vartheta}$, which represents the transport rate constants so intrinsically $\boldsymbol{\vartheta} \geq \mathbf{0}$.

Algorithm 1. Joint 4-D parametric reconstruction and motion correction

Input: PET projection data $\boldsymbol{g}$, the basis function $\boldsymbol{B}$, regularisation parameter β
Output: Motion-corrected parametric images $\boldsymbol{\vartheta}$ and the motion estimate $\boldsymbol{\alpha}$.
Initialisation $\boldsymbol{\alpha} = \mathbf{0}$, $\boldsymbol{\vartheta} = \mathbf{0.01}$;
while *not converged* **do**
  $\boldsymbol{\vartheta}^{[0]} = \boldsymbol{\vartheta}$;
  for k=1,...,K **do**
    Compute $f_{j,l}^{\mathrm{em},[k]}$ and $f_{j,l}^{\mathrm{reg},[k]}$ from $\boldsymbol{\vartheta}^{[k-1]}$ for all j, l ;
    $\boldsymbol{\vartheta}^{[k,0]} = \boldsymbol{\vartheta}^{[k-1]}$;
    for r=1,...,R **do**
      Compute $\vartheta_{j,q}^{\mathrm{em},[k,r]}, \vartheta_{j,q}^{\mathrm{reg},[k,r]}$ from $\boldsymbol{\vartheta}_j^{[k,r-1]}$ for all j, q ;
      Update $\boldsymbol{\vartheta}^{[k,r]}$ by (4) ;
    end
    $\boldsymbol{\vartheta}^{[k]} = \boldsymbol{\vartheta}^{[k,R]}$;
  end
  $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}^{[K]}$;
  Optimise $\boldsymbol{\alpha}$ using the trust region algorithm with the gradient and Hessian derived from (2) and the current estimate of $\boldsymbol{\vartheta}$;
end

3 Results

3.1 Simulation-Based Validation

The proposed algorithm was firstly validated using simulated [^{11}C]Raclopride data. 60-min dynamic PET scans were generated based on the Zubal brain phantom, with various kinetics defined on background, grey matter, white matter, cerebellum, putamen and caudate nucleus. The time activity curves were generated based on the two-tissue compartment model using kinetic parameters derived from clinical studies in conjunction with a population input function. Rigid head movements were introduced by transforming the activity images at various time points (Fig. 1), in accordance with the expected amplitude and frequency of head motion that happens within a brain scan. In practice, detection of such motion events can be performed by applying the PCA technique proposed in [9] with a moving time window for dynamic PET data. The activity images were defined on a grid of $128 \times 128 \times 80$ voxels and then used to generate PET projections (5 mm FWHM resolution). Poisson noise was added to the projections and reconstruction was then performed with ideal pre-reconstruction correction for randoms, scatter and attenuation (e.g. derived from the simultaneous CT or MRI [10]). For comparison, the direct parametric reconstruction without motion correction (*direct*) [1] and the indirect parametric reconstruction with motion correction (*indirect+MC*) [4] were also performed in addition to the proposed algorithm (*direct+MC*). In all the methods we used spectral analysis [8] for kinetic modelling with the same temporal basis functions $\boldsymbol{B}$ for which $n_k = 16$,

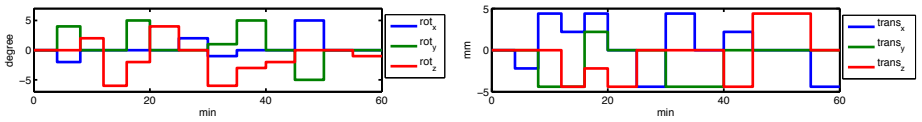

Fig. 1. Simulated random rigid head movements

and ϕ_q is spaced logarithmically in the range of $[0.0001s^{-1}, 1s^{-1}]$. Different β values were applied to each reconstruction. For the *indirect+MC* method, the quadratic penalty with β was applied to reconstructing the activity images by MLEM [6]. Motion correction accuracy was quantified by the target registration error (TRE) [11] averaged over time, shown in Fig. 2. For the parametric recon-

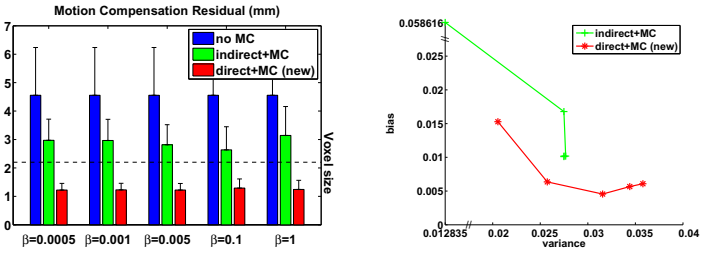

Fig. 2. Left: Motion correction performance quantified by time-averaged target registration error (TRE). The proposed *direct+MC* achieves sub-voxel size(< 2.2 mm) MC accuracy with noisy data, the β range has little impact on the motion correction performance due to the low smoothing level; Right: Bias versus variance trade-off of the volume of distribution (V_T) images of [^{11}C]Raclopride reconstructed by *indirect+MC* and *direct+MC (new)*. Data points from left to right correspond to β values of 1, 0.1, 0.005, 0.001 and 0.0005. The proposed *direct+MC* algorithm achieves lower bias.

struction, the outcome measures of interest, which is the volume of distribution V_T for [^{11}C]Raclopride was calculated using $\boldsymbol{\vartheta}$, by $V_T = \sum_q \frac{\vartheta_q}{\phi_q}$ [12]. The corresponding bias and variance plots of the parametric reconstruction of V_T are shown in Fig. 2, and Fig. 3 shows the ground truth and reconstructed V_T images by all these methods from one simulated scan.

3.2 Clinical Data

The proposed algorithm was also applied to reconstruct clinical [^{18}F]Choline data from a patient scanned for 44 mins using a Siemens Biograph mMR scanner. The reconstruction was performed by the proposed algorithm integrated with STIR [13]. The attenuation correction was conducted during the iterations using the attenuation map repositioned by the current motion estimate to account for the mismatch caused by motion. The kinetic parameter of interest

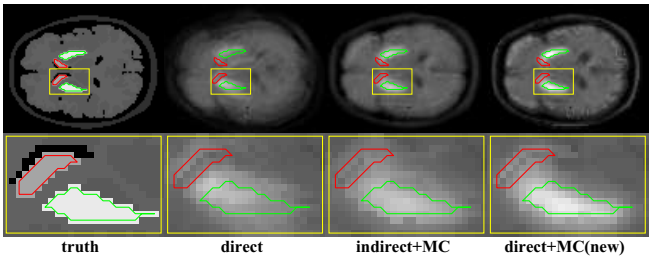

Fig. 3. Selected slice (z : 45/80) of the reconstructed V_T images of [^{11}C]Raclopride when the penalty weight $\beta = 0.005$. The direct reconstruction with no motion correction (*direct*) results in severely blurred images. The indirect reconstruction with post MC in the image space (*indirect+MC*) loses the spatial contrast due to inadequate MC. The proposed joint direct reconstruction with motion correction (*direct+MC*) better preserves the ROIs (e.g. putamen).

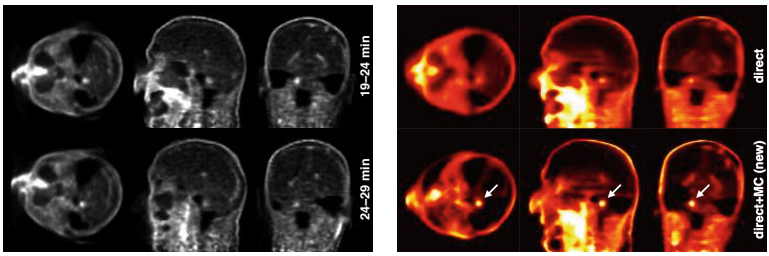

Fig. 4. Left: Two [^{18}F]Choline time frames reconstructed by MLEM to illustrate the data. Right: The K_I images reconstructed by the *direct* method without motion correction, and by the proposed *direct+MC* algorithm. The identification of the [^{18}F]Choline hot spot is greatly enhanced by the proposed method.

K_I was calculated for [^{18}F]Choline, which is the irreversible uptake rate constant from plasma defined by $K_I = \vartheta_{n_k}$ (corresponds to $\phi_{n_k} = 0$) [12]. For the temporal basis functions $\boldsymbol{B}$, $n_k = 16$ and ϕ_q is logarithmically in the range of $[0.000001s^{-1}, 0.01s^{-1}]$. Fig. 4 shows the reconstruction of K_I by the *direct* method without motion correction and by the proposed *direct+MC* algorithm. The result demonstrates that, by applying the proposed method, the reconstructed K_I image shows the [^{18}F]Choline hot spot which is otherwise difficult to identify due to motion or noise.

4 Discussion and Conclusion

This work proposed a novel algorithm for joint motion correction and parametric reconstruction of dynamic PET data, which addresses two of the fundamental issues in PET imaging in a unified framework. In particular, a parallelisable closed form solution of a sub-problem derived here has greatly enhanced the computational efficiency for practical use. Initial evaluation on clinical data has

shown improved identification of the kinetic activity in the regions of interest. The proposed method can adapt to different PET tracers without the need of reimplementation, and future evaluation over a range of clinical applications will determine the full utility of the method.

Acknowledgement. EPSRC (EP/H046410/1, EP/J020990/1, EP/K005278), the MRC (MR/J01107X/1), the NIHR Biomedical Research Unit (Dementia) at UCL and the National Institute for Health Research University College London Hospitals Biomedical Research Centre (NIHR BRC UCLH/UCL High Impact Initiative), Roger Gunn and Graham Searle from Imanova Ltd (kinetic analysis).

References

1. Wang, G., Qi, J.: An optimization transfer algorithm for nonlinear parametric image reconstruction from dynamic PET data. IEEE. Trans. on Med. Imag. 31(10), 1977–1988 (2012)
2. Cheng, X., Navab, N., Ziegler, S.I., Shi, K.: Direct Parametric Image Reconstruction of Rapid Multi-tracer PET. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part III. LNCS, vol. 8151, pp. 155–162. Springer, Heidelberg (2013)
3. Pedemonte, S., Bousse, A., Hutton, B.F., Arridge, S., Ourselin, S.: 4-D generative model for PET/MRI reconstruction. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part I. LNCS, vol. 6891, pp. 581–588. Springer, Heidelberg (2011)
4. Jiao, J., Searle, G.E., Tziortzi, A.C., Salinas, C.A., Gunn, R.N., Schnabel, J.A.: Spatio-temporal pharmacokinetic model based registration of 4D PET neuroimaging data. Neuroimage 84, 225–235 (2014)
5. Madsen, K., Nielsen, H.B., Tingleff, O.: Methods for non-linear least-squares problems. Technical report, Technical University of Denmark (2004)
6. De Pierro, A.R.: A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. IEEE Trans. on Med. Imag. 14(1), 132–137 (1995)
7. Wang, G., Qi, J.: Acceleration of the direct reconstruction of linear parametric images using nested algorithms. Phys. Med. Biol. 55, 1505–1517 (2010)
8. Cunningham, V.J., Jones, T.: Spectral analysis of dynamic PET studies. J. Cereb. Blood Flow Metab. 13(1), 15–23 (1993)
9. Thielemans, K., Schleyer, P., Dunn, J., Marsden, P.K., Manjeshwar, R.M.: Using PCA to detect head motion from PET list mode data. In: IEEE Nucl. Sci. Syp. & Med. Im. Conf. (2013)
10. Burgos, N., Cardoso, M.J., Modat, M., Pedemonte, S., Dickson, J., Barnes, A., Duncan, J.S., Atkinson, D., Arridge, S.R., Hutton, B.F., Ourselin, S.: Attenuation correction synthesis for hybrid PET-MR scanners. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part I. LNCS, vol. 8149, pp. 147–154. Springer, Heidelberg (2013)
11. Fitzpatrick, J., West, J., Maurer Jr., C.R.: Predicting error in rigid-body point-based registration. IEEE Trans. on Med. Imag. 17(5), 694–702 (1998)
12. Gunn, R.N., Gunn, S.R., Cunningham, V.J.: Positron emission tomography compartmental models. J. Cereb. Blood Flow Metab. 21(6), 635–652 (2001)
13. Thielemans, K., Tsoumpas, C., Mustafovic, S., Beisel, T., Aguiar, P., Dikaios, N., Jacobson, M.W.: STIR: software for tomographic image reconstruction release 2. Phys. Med. Biol. 57(4), 867–883 (2012)