An Optimization Transfer Algorithm for Nonlinear Parametric Image Reconstruction from Dynamic PET Data

Guobao Wang and **Jinyi Qi**
Department of Biomedical Engineering, University of California, Davis, CA 95616

Guobao Wang: gbwang@ucdavis.edu; Jinyi Qi: qi@ucdavis.edu

Abstract

Direct reconstruction of kinetic parameters from raw projection data is a challenging task in molecular imaging using dynamic positron emission tomography (PET). This paper presents a new optimization transfer algorithm for penalized likelihood direct reconstruction of nonlinear parametric images that is easy to use and has a fast convergence rate. Each iteration of the proposed algorithm can be implemented in three simple steps: a frame-by-frame maximum likelihood EM-like image update, a frame-by-frame image smoothing, and a pixel-by-pixel time activity curve fitting. Computer simulation shows that the direct algorithm can achieve a better bias-variance performance than the indirect reconstruction algorithm. The convergence rate of the new algorithm is substantially faster than our previous algorithm that is based on a separable paraboloidal surrogate function. The proposed algorithm has been applied to real 4D PET data.

Index Terms

Image reconstruction; kinetic modeling; penalized maximum likelihood; parametric imaging

I. Introduction

Parametric imaging using dynamic positron emission tomography (PET) provides important information for biological research and clinical diagnosis [1]–[3]. Direct reconstruction methods combine kinetic modeling and emission image reconstruction into a single formula and estimate parametric images directly from raw projection data (see [4], [5] for surveys on this topic). Compared with conventional indirect methods which reconstruct a sequence of emission images first and then estimate kinetic parameters from time activity curves (TACs) in a separate step, direct reconstruction methods can accurately model the statistics of noise in the kinetic parameter estimation process.

One drawback of direct reconstruction with compartment models is that the optimization algorithms are usually more complex than indirect methods, because the kinetic parameters are related to the expectation of PET data nonlinearly. Most existing algorithms [6]–[12] were derived for a specific kinetic model. Recently, a generalized algorithm [13] was proposed for direct reconstruction. It uses a paraboloidal surrogate function [15] in the sinogram domain to transfer the penalized maximum-likelihood (ML) problem at each iteration into a pixel-wise nonlinear least squares (NLS) formulation. However, the paraboloidal surrogate function requires the expectation of the background events (randoms

and scatters) to be nonzero in all sinogram bins, which may limit its applications. Furthermore, the convergence rate of this algorithm can be very slow if the level of background events is low.

In this paper, we present a new algorithm that uses the expectation-maximization (EM) [16] based surrogate function in the image domain for direct reconstruction of nonlinear parametric images in dynamic PET. Compared with the paraboloidal surrogate, the EM surrogate does not require a positive background. We have previously used the same EM surrogate for the ML reconstruction of linear parametric images and achieved fast convergence rate [17]. For linear parametric models, the EM surrogate function can be easily maximized by using the well-known ML EM update [18], [19], which results in the so-called nested EM algorithm [17]. The penalized likelihood reconstruction of nonlinear parametric images is more challenging because of the nonlinear relationship between the kinetic parameters and noisefree data. Here we developed a modified Levenberg-Marquardt algorithm to maximize the surrogate function of the penalized likelihood. Part of this work was previously presented at the 2010 IEEE Nuclear Science Symposium and Medical Imaging conference [20]. A similar work was also presented for maximum likelihood reconstruction at the same conference by Matthews *et al* [21]. Rather than solving the ML problem directly, the method by Matthews *et al* converts the ML problem to a weighted least squares problem and determines the weight using a one-step-late approach. Without a proper line search, the one-step-late approach does not guarantee a monotonic convergence to the ML solution, although convergence has often been observed empirically. In comparison, the proposed method guarantes monotonic convergence and is able to regularize image reconstruction through the penalized likelihood formulation.

The rest of the paper is organized as follows. Section II describes the general formulation of direct reconstruction for dynamic PET data. Section III presents the details of the proposed algorithm under the optimization transfer framework. Computer simulations are given in Section IV to validate the performance of the proposed algorithm and to compare with our previous algorithm and other algorithms. An application of the proposed algorithm to real data is given in Section V. Finally conclusions are drawn in Section VI.

II. Direct Reconstruction of Parametric Images

Dynamic PET can monitor the spatiotemporal distribution of radioactive tracers *in vivo*. The concentration of the radiotracer at a given pixel at time t can be described by the following tracer kinetic model [22]

$$C_T(t;\boldsymbol{\kappa}, f_v) = (1-f_v)h(t;\boldsymbol{\kappa}) \otimes C_p(t) + f_v C_{\mathrm{wb}}(t) \quad (1)$$

where $\boldsymbol{\kappa}$ is a vector that contains all the kinetic parameters that determine the tracer uptake in the tissue, f_v is the fractional volume of blood in the tissue, $h(t; \boldsymbol{\kappa})$ is the impulse response function, $C_p(t)$ is the tracer concentration in plasma, $C_{\mathrm{wb}}(t)$ is the tracer concentration in whole blood, and "$\otimes$" denotes the convolution operator. The model in (1) is general and includes compartmental models and simplified linear models as special cases. For example, the impulse response $h(t; \boldsymbol{\kappa})$ of the commonly used three-compartment model is given by

$$\boldsymbol{h}(t;\boldsymbol{\kappa})=\frac{K_1}{\Delta\alpha}(k_3+k_4-\alpha_1 \quad \alpha_2-k_3-k_4)e^{-\boldsymbol{\alpha}(t)}, \quad (2)$$

where $\boldsymbol{\kappa}=\{K_1, k_2, k_3, k_4\}$ are the transport rate constants between compartments, $\Delta a = a_2 - a_1$ with

$$\alpha_{1,2}=\tfrac{1}{2}(k_2+k_3+k_4)\mp\tfrac{1}{2}\left[(k_2+k_3+k_4)^2-4k_2k_4\right]^{1/2}$$

, and $e^{-\boldsymbol{\alpha}(t)} = (e^{-a_1t}, e^{-a_2t})^T$. For simplicity, we will use $\theta_j \in IR^{n_p}$ to denote the collection of transport rates and fractional volume $\{\boldsymbol{\kappa}, f_v\}$ for pixel j and $\boldsymbol{\theta} \equiv \{\theta_j\}$ to denote the parametric images. n_p is the total number of kinetic parameters in a kinetic model.

A dynamic PET scan is often divided into multiple consecutive time frames, with each frame containing coincidence events recorded from the start of the frame till the end of the frame. The image intensity at pixel j in time frame m, $x_m(\theta_j)$, is then given by

$$x_m(\boldsymbol{\theta}_j)=\int_{t_{m,s}}^{t_{m,e}} C_T(\tau;\boldsymbol{\theta}_j)e^{-\lambda\tau}d\tau, \quad (3)$$

where $t_{m,s}$ and $t_{m,e}$ denote the start and end times of frame m, respectively, and λ is the decay constant of the radiotracer. PET measured data can be modeled as a collection of independent Poisson random variables with the expected projection $\bar{\boldsymbol{y}}_m(\boldsymbol{\theta})$ in time frame m related to the dynamic image $\boldsymbol{x}_m(\boldsymbol{\theta})$ through an affine transform [13],

$$\overline{\boldsymbol{y}}_m(\boldsymbol{\theta})=\boldsymbol{P}\boldsymbol{x}_m(\boldsymbol{\theta})+\boldsymbol{r}_m, \quad (4)$$

where the ith element of $\bar{\boldsymbol{y}}_m(\boldsymbol{\theta})$, $\bar{y}_{im}(\boldsymbol{\theta})$, denotes the expected measurement of detector pair i in frame m, the (i, j)th element of $\boldsymbol{P} \in IR^{n_i \times n_j}$, p_{ij}, is the probability of detecting an event originated in pixel j by detector pair i, and $\boldsymbol{r}_m \in IR^{n_i}$ is the expectation of scattered and random events in the mth frame. n_i and n_j are the total number of the detector pairs and voxels, respectively.

The goal of dynamic PET is to estimate $\boldsymbol{\theta}$ from dynamic PET measurements. Let y_{im} denote the coincidence data detected by detector pair i in time frame m and $\boldsymbol{y} \equiv \{y_{im}\}_{i=1,m=1}^{n_i,n_m}$ be the collection of all the measurements. The log-likelihood function of the dynamic PET data set, omitting constants that are independent of $\boldsymbol{\theta}$, is

$$L(\boldsymbol{y}|\boldsymbol{\theta})=\sum_{m=1}^{n_m}\sum_{i=1}^{n_i}y_{im}\log\overline{y}_{im}(\boldsymbol{\theta})-\overline{y}_{im}(\boldsymbol{\theta}), \quad (5)$$

where n_m is the total number of time frames. Direct reconstruction finds the solution by maximizing the penalized likelihood function

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \Phi(\boldsymbol{\theta}), \quad \Phi(\boldsymbol{\theta}) = L(\boldsymbol{y}|\boldsymbol{\theta}) - \beta U(\boldsymbol{\theta}). \quad (6)$$

where $U(\boldsymbol{\theta})$ is a smoothness penalty and β is the regularization parameter that controls the tradeoff between the resolution and noise. The smoothness penalty can be applied either on the kinetic parameters $\boldsymbol{\theta}$ or on the dynamic image $\{\boldsymbol{x}_m(\boldsymbol{\theta})\}$, depending on the application. Here we use the following quadratic penalty applied on the dynamic image

$$U(\boldsymbol{\theta}) - \frac{1}{4}\sum_{m=1}^{n_m}\sum_{j=1}^{n_j}\sum_{l\in\mathcal{N}_j}\frac{1}{2}\gamma_{jl}(x_m(\boldsymbol{\theta}_j) - x_m(\boldsymbol{\theta}_l))^2 \quad (7)$$

where $\mathcal{N}_j$ denotes the neighborhood of pixel j and γ_{jl} is the weighting factor equal to the inverse distance between pixels j and l. We use the eight nearest neighboring pixels in 2D and 26 nearest neighbors in 3D.

III. Proposed Algorithm

A. Maximum likelihood reconstruction

For a concave function f, we have the following inequality [14]

$$f\left(\sum_{j=0}^{n_j}w_j a_j\right) \geq \sum_{j=0}^{n_j}\frac{w_j b_j}{\sum_{l=0}^{n_j}w_l b_l}f\left(\frac{\sum_{l=0}^{n_j}w_l b_l}{b_j}a_j\right) \quad (8)$$

where $\{w_j\}$, $\{a_j\}$ and $\{b_j\}$ are positive vectors.

Using the above inequality and setting $w_j = p_{ij}$, $a_j = x_m(\boldsymbol{\theta}_j)$, $b_j = x_m(\boldsymbol{\theta}_j^n)$ for $j > 0$ and $w_0 = r_{im}$, $a_0 = b_0 = 1$, we get

$$\begin{aligned}\log\bar{y}_{im}(\boldsymbol{\theta}) &= \log\left(\sum_{j=1}^{n_j}p_{ij}x_m(\boldsymbol{\theta}_j) + r_{im}\right)\\ &\geq \sum_{j=1}^{n_j}\frac{p_{ij}x_m(\boldsymbol{\theta}_j^n)}{\bar{y}_{im}(\boldsymbol{\theta}^n)}\log x_m(\boldsymbol{\theta}_j) + c_1^n + c_2^n\end{aligned} \quad (9)$$

where the superscript n denotes the estimate at iteration n. The constants

$c_1^n = \sum_{j=1}^{n_j}\frac{p_{ij}x_m(\boldsymbol{\theta}_j^n)}{\bar{y}_{im}(\boldsymbol{\theta}^n)}\log\frac{\bar{y}_{im}(\boldsymbol{\theta}^n)}{x_m(\boldsymbol{\theta}_j^n)}$ and $c_2^n = \frac{r_{im}}{\bar{y}_{im}(\boldsymbol{\theta}^n)}\log\bar{y}_{im}(\boldsymbol{\theta}^n)$ depend on the known estimate $\boldsymbol{\theta}^n$, but are independent of the unknown parameter $\boldsymbol{\theta}$ and hence will be omitted in the following optimization.

Then the following surrogate function can be obtained for the original likelihood function $L(\boldsymbol{y}|\boldsymbol{\theta})$:

$$Q_1(\boldsymbol{\theta};\boldsymbol{\theta}^n) = \sum_{j=1}^{n_j} p_j \left(\sum_{m=1}^{n_m} \hat{x}_{jm}^{\mathrm{em},n} \log x_m(\boldsymbol{\theta}_j) - x_m(\boldsymbol{\theta}_j) \right) \quad (10)$$

where

$$p_j = \sum_{i=1}^{n_i} p_{ij} \quad (11)$$

and $\{ \hat{x}_{jm}^{\mathrm{cm},n} \}$ is an intermediate dynamic image given by

$$\hat{x}_{jm}^{\mathrm{em},n} = \frac{x_m(\boldsymbol{\theta}_j^n)}{p_j} \sum_{i=1}^{n_i} p_{ij} \frac{y_{im}}{\overline{y}_{im}(\boldsymbol{\theta}^n)}. \quad (12)$$

Equation (12) resembles the ML EM update for a dynamic PET image.

It is easy to prove that the surrogate function in (10) satisfies

$$Q_1(\boldsymbol{\theta};\boldsymbol{\theta}^n) - Q_1(\boldsymbol{\theta}^n;\boldsymbol{\theta}^n) \leq L(\boldsymbol{y}|\boldsymbol{\theta}) - L(\boldsymbol{y}|\boldsymbol{\theta}^n), \quad (13)$$

$$\nabla Q_1(\boldsymbol{\theta}^n;\boldsymbol{\theta}^n) = \nabla L(\boldsymbol{y}|\boldsymbol{\theta}^n), \quad (14)$$

where ∇ denotes the gradient with respect to $\boldsymbol{\theta}$. The optimization of the original likelihood function can now be transferred to the maximization of the surrogate function

$$\boldsymbol{\theta}^{n+1} = \arg \max_{\boldsymbol{\theta}} Q_1(\boldsymbol{\theta};\boldsymbol{\theta}^n) \quad (15)$$

which guarantees

$$L(\boldsymbol{y}|\boldsymbol{\theta}^{n+1}) \geq L(\boldsymbol{y}|\boldsymbol{\theta}^n) \quad (16)$$

and the equality holds only when $\boldsymbol{\theta}^n$ achieves the maximum of $L(\boldsymbol{y}|\boldsymbol{\theta})$ because of the matched gradient condition in (14).

It is worth noting that the optimization in (15) is separable for pixels and can be solved by the following pixel-wise optimization:

$$\boldsymbol{\theta}_j^{n+1} = \arg \max_{\boldsymbol{\theta}_j} q_1(\boldsymbol{\theta}_j;\hat{x}_j^{\mathrm{em},n}) \quad (17)$$

where $\hat{x}_j^{\mathrm{cm},n} \equiv \{\hat{x}_{jm}^{\mathrm{cm},n}\}_{m=1}^{n_m}$ and

$$q_1(\boldsymbol{\varphi}; \boldsymbol{f}^{\mathrm{em}}) \triangleq \sum_{m=1}^{n_m} f_m^{\mathrm{em}} \log x_m(\boldsymbol{\varphi}) - x_m(\boldsymbol{\varphi}) \quad (18)$$

with $\boldsymbol{\phi} \in IR^{n_p}$ and $\boldsymbol{f}^{\mathrm{em}} \in IR^{n_m}$.

Note that $q_1(\boldsymbol{\theta}_j; \hat{\boldsymbol{x}}_j^{\mathrm{cm},n})$ shares the same form as the Poisson log-likelihood function. However, it does not suggest that $\hat{x}_{jm}^{\mathrm{cm},n}$ follows an independent Poisson distribution because $q_1(\boldsymbol{\theta}_j; \hat{\boldsymbol{x}}_j^{\mathrm{cm},n})$ is just a surrogate function for the optimization.

Compared with the formula given by Carson and Lange in [23], the introduction of $\hat{x}_{jm}^{\mathrm{cm},n}$ simplifies the implementation of the maximum likelihood estimation. When $x_m(\boldsymbol{\theta}_j)$ is a linear model of $\boldsymbol{\theta}_j$, an EM algorithm can be used to maximize $q_1(\boldsymbol{\theta}_j; \hat{\boldsymbol{x}}_j^{\mathrm{cm},n})$, which results in the nested EM algorithm in [17]. For nonlinear parametric image reconstruction, the fitting problem in (17) can be solved by the modified Levenberg-Marquardt algorithm to be presented in Section III-C.

B. Penalized likelihood reconstruction

The penalized likelihood function $\Phi(\boldsymbol{\theta})$ contains a penalty term $U(\boldsymbol{\theta})$ defined by (7). Using De Pierro's decoupling rule [14], [24]

$$(x_m(\boldsymbol{\theta}_j) - x_m(\boldsymbol{\theta}_l))^2 \leq \frac{1}{2}\left(2x_m(\boldsymbol{\theta}_j) - x_m(\boldsymbol{\theta}_j^n) - x_m(\boldsymbol{\theta}_l^n)\right)^2 + \frac{1}{2}\left(2x_m(\boldsymbol{\theta}_l) - x_m(\boldsymbol{\theta}_j^n) - x_m(\boldsymbol{\theta}_l^n)\right)^2 \quad (19)$$

we can construct a separable surrogate function $Q_2(\boldsymbol{\theta}; \boldsymbol{\theta}^t)$ for the regularization term $U(\boldsymbol{\theta})$:

$$Q_2(\boldsymbol{\theta}; \boldsymbol{\theta}^n) = \frac{1}{2} \sum_{m=1}^{n_m} \sum_{j=1}^{n_j} w_{jm}^{\mathrm{reg}}\left(\hat{x}_{jm}^{\mathrm{reg},n} - x_m(\boldsymbol{\theta}_j)\right)^2 \quad (20)$$

where

$$w_{jm}^{\mathrm{reg}} \triangleq \sum_{l \in \mathcal{N}_j} \gamma_{jl} \quad (21)$$

and $\{\hat{x}_{jm}^{\mathrm{reg},n}\}$ is an intermediate smoothed dynamic image obtained at iteration n by

$$\hat{x}_{jm}^{\mathrm{reg},n} = \frac{1}{2w_{jm}^{\mathrm{reg}}} \sum_{l \in \mathcal{N}_j} \gamma_{jl}\left(x_m(\boldsymbol{\theta}_j^n) + x_m(\boldsymbol{\theta}_l^n)\right). \quad (22)$$

It is easy to verify that the surrogate function satisfies

$$Q_2(\boldsymbol{\theta}; \boldsymbol{\theta}^n) - Q_2(\boldsymbol{\theta}^n; \boldsymbol{\theta}^n) \geq U(\boldsymbol{\theta}) - U(\boldsymbol{\theta}^n). \quad (23)$$

$$\nabla Q_2(\boldsymbol{\theta}^n;\boldsymbol{\theta}^n)=\nabla U(\boldsymbol{\theta}^n). \quad (24)$$

Combining the surrogate functions for the log-likelihood and regularization terms gives us the overall surrogate function

$$Q(\boldsymbol{\theta};\boldsymbol{\theta}^n)=Q_1(\boldsymbol{\theta};\boldsymbol{\theta}^n)-\beta Q_2(\boldsymbol{\theta};\boldsymbol{\theta}^n) \quad (25)$$

for the penalized likelihood function $\Phi(\boldsymbol{\theta})$. Again, the surrogate function $Q(\boldsymbol{\theta};\boldsymbol{\theta}^n)$ is separable for pixels, so the large-scale optimization of the penalized likelihood function $\Phi(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ can be transferred into the following small-scale pixel-wise optimization:

$$\boldsymbol{\theta}_j^{n+1}=\arg\max_{\boldsymbol{\theta}_j} q_1\left(\boldsymbol{\theta}_j;\hat{\boldsymbol{x}}_j^{\mathrm{em},n}\right)-q_2\left(\boldsymbol{\theta}_j;\hat{\boldsymbol{x}}_j^{\mathrm{reg},n},\frac{\beta}{p_j}\boldsymbol{w}_j^{\mathrm{reg}}\right) \quad (26)$$

where $\hat{\boldsymbol{x}}_j^{\mathrm{reg},n}\equiv\left\{\hat{x}_{jm}^{\mathrm{reg},n}\right\}_{m=1}^{n_m}, \boldsymbol{w}_j^{\mathrm{reg}}\equiv\{w_{jm}^{\mathrm{reg}}\}_{m=1}^{n_m}$ and

$$q_2(\boldsymbol{\varphi};\boldsymbol{f}^{\mathrm{reg}},\boldsymbol{\omega}^{\mathrm{reg}})\triangleq\frac{1}{2}\sum_{m=1}^{n_m}\omega_m^{\mathrm{reg}}(f_m^{\mathrm{reg}}-x_m(\boldsymbol{\varphi}))^2 \quad (27)$$

with $\boldsymbol{f}^{\mathrm{reg}}, \boldsymbol{\omega}^{\mathrm{reg}}\in IR^{n_m}$. Note that $q_2\left(\boldsymbol{\theta}_j;\hat{\boldsymbol{x}}_j^{\mathrm{reg},n},\frac{\beta}{p_j}\boldsymbol{w}_j^{\mathrm{reg}}\right)=\frac{\beta}{p_j}q_2\left(\boldsymbol{\theta}_j;\hat{\boldsymbol{x}}_j^{\mathrm{reg},n},\boldsymbol{w}_j^{\mathrm{reg}}\right)$.

The objective function in (26) resembles a penalized Poisson likelihood for one-dimensional curve fitting. It can be solved by many existing nonlinear optimization algorithms. Here we use the modified Levenberg-Marquardt algorithm.

C. Levenberg-Marquardt algorithm

The Levenberg-Marquardt (LM) algorithm [25] is usually used to solve nonlinear least squares problems. Here we modify it slightly to solve the penalized Poisson likelihood curve fitting problem in (26) which is equivalent to solving the following minimization problem

$$\hat{\boldsymbol{\varphi}}=\arg\min_{\boldsymbol{\varphi}} q(\boldsymbol{\varphi}) \quad (28)$$

$$q(\boldsymbol{\varphi})=-q_1(\boldsymbol{\varphi};\boldsymbol{f}^{\mathrm{em}})+q_2(\boldsymbol{\varphi};\boldsymbol{f}^{\mathrm{reg}},\boldsymbol{\omega}^{\mathrm{reg}}) \quad (29)$$

Given an estimate at iteration k, the second-order Taylor approximation of $q(\boldsymbol{\varphi})$ is

$$\tilde{q}(\boldsymbol{\varphi})=q(\boldsymbol{\varphi}^k)+\boldsymbol{g}^T(\boldsymbol{\varphi}-\boldsymbol{\varphi}^k)+\frac{1}{2}(\boldsymbol{\varphi}-\boldsymbol{\varphi}^k)^T\boldsymbol{H}(\boldsymbol{\varphi}-\boldsymbol{\varphi}^k) \quad (30)$$

where $\boldsymbol{g}\in IR^{n_p}$ is the gradient vector and $\boldsymbol{H}\in IR^{n_p\times n_p}$ is the Hessian matrix:

$$g = J^T \dot{q}^k, \quad H = J^T \text{diag}[\ddot{q}^k] J \quad (31)$$

where $J \in IR^{nm \times np}$ is the Jacobian matrix of time activity curve with the (m, p)th element

being $\dfrac{\partial x_m(\varphi)}{\partial \varphi_p}$. The vectors $q^{k\cdot} \in IR^{nm}$ and $q^{k\cdot\cdot} \in IR^{nm}$ are the first and second derivatives of the function $q(\phi)$ with respect to $x_m(\phi)$ at iteration k, respectively:

$$\dot{q}_m^k = \left(1 - \frac{f_m^{\text{cm}}}{x_m(\varphi^k)} \right) + \omega_m^{\text{reg}} \left(x_m(\varphi^k) - f_m^{\text{reg}} \right) \quad (32)$$

$$\ddot{q}_m^k = \frac{f_m^{\text{cm}}}{x_m^2(\varphi^k)} + \omega_m^{\text{reg}} \quad (33)$$

A new estimate of ϕ can be obtained by minimizing the following objective function

$$\hat{\varphi}(\mu) = \arg \min_{\varphi} \tilde{q}(\varphi) + \frac{\mu}{2} (\varphi - \varphi^k)^T D[H_{pp}](\varphi - \varphi^k) \quad (34)$$

subject to the bound constraints $\underline{\phi} \leq \phi \leq \bar{\phi}$, where $\underline{\phi}$ and $\bar{\phi}$ are the lower and upper bounds of the parameters, respectively. μ is the damping parameter. $D[H_{pp}]$ is a diagonal matrix with the pth diagonal element being H_{pp}. We use a coordinate descent algorithm for the quadratic optimization in (34). The method described in [26] is used to find the optimal damping parameter μ^* that results in the maximum decrease in the original objective function, i.e.,

$$\mu^* = \arg \min_{\mu} q(\hat{\varphi}(\mu)). \quad (35)$$

Then the estimate at iteration $(k + 1)$ is set to

$$\varphi^{k+1} = \hat{\varphi}(\mu^*). \quad (36)$$

The nonlinear fitting optimization therefore guarantees monotonic convergence to a local optimum.

The generalized algorithm and the modified LM algorithm are summarized in the algorithmic tables 1 and 2, respectively. A comparison with our previous algorithms [13] is given in Table I.

D. Convergence

The modified LM algorithm guarantees an increase in the surrogate function given in (26). Using the property of the surrogate function, we then have

$$\Phi(\boldsymbol{\theta}^{n+1})-\Phi(\boldsymbol{\theta}^n) \geq Q(\boldsymbol{\theta}^{n+1};\boldsymbol{\theta}^n)-Q(\boldsymbol{\theta}^n;\boldsymbol{\theta}^n) \geq 0 \quad (37)$$

If $\boldsymbol{\theta}^n$ is the maximizer of $Q(\boldsymbol{\theta}; \boldsymbol{\theta}^n)$, i.e. $\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n$, we have

$$\nabla Q(\boldsymbol{\theta}^n;\boldsymbol{\theta}^n)=\nabla Q(\boldsymbol{\theta}^{n+1};\boldsymbol{\theta}^n)=0 \quad (38)$$

according to the Karush-Kuhn-Tucker condition. Combining the above equation with equations (14) and (24), we get

$$\nabla \Phi(\boldsymbol{\theta}^n) \equiv \nabla Q(\boldsymbol{\theta}^n;\boldsymbol{\theta}^n)=\mathbf{0}. \quad (39)$$

indicating $\boldsymbol{\theta}^n$ also maximizes the original objective function $\Phi(\boldsymbol{\theta})$.

Thus, the whole algorithm guarantees a monotonic convergence to a local optimum for the nonlinear parametric image reconstruction. When a linear kinetic model is used, the convergence is also global because the penalized log-likelihood function is concave.

IV. Computer Simulations

We performed computer simulations to validate the proposed algorithm. A 60-minute ^{18}F-FDG brain scan was simulated using one slice of the Zubal phantom [27] that consisted of gray matter, white matter and a small tumor inside the white matter, as shown in Fig. 1(a). The scanning schedule of dynamic PET data consisted of 24 time frames: 4×20 s, 4×40 s, 4×60 s, 4×180 s and 8×300 s. The time activity curve of each region was generated using a three-compartment model and an analytical blood input function in Feng's model [28], and shown in Fig. 1(b). The activity is shown in an arbitrary unit (a.u.) that is proportional to Bq. Since the scaling factor has no effect on image reconstruction algorithms, we did not estimate it here. The kinetic parameters f_v, K_1, k_2, k_3, k_4 used for gray matter and white matter are taken from [29] and listed in Table II. The TACs were integrated for each time frame and forward projected to generate dynamic sinograms. The forward projection model included object self-attenuation. The expectation of scatter sinogram was obtained by using the SimSET simulation. The scatter fraction was set to 20%. A 20% uniform randoms was also included. Poisson noise was then generated, resulting in an expected total number of events over the 60 minutes equal to 20 million.

Fifty independent and identically-distributed noisy datasets were generated and processed independently by the direct reconstruction method and indirect method to estimate the images of kinetic parameters. The regularization parameters in both reconstruction methods were the same and covered a wide range of values: 0, 1×10^{-4}, 3×10^{-4}, 1×10^{-3}, 3×10^{-3}, 1×10^{-2}. Both methods used 200 iterations in the reconstruction step to guarantee convergence. At each iteration of the direct algorithm, two iterations of the modified Levenberg-Marquardt algorithm were used. In the indirect method, dynamic images were first reconstructed frame-by-frame using the regular MAP-EM algorithm [24] for static image reconstruction, then followed by the usual Levenberg-Marquardt algorithm with 100 iterations to estimate the kinetic parameters pixel-by-pixel. The weighting factor in the

kinetic fitting in the indirect method was chosen to be the squared frame duration divided by the count of the frame. The initial values of the kinetic parameters in both reconstructions were set to 0.01 for all pixels. The lower bounds were set to 10^{-5} and the upper bounds were 1.0 for all the kinetic parameters. The influx rate

$$K_i = \frac{K_1 k_3}{k_2 + k_3},$$

which is the major parameter of interest in FDG PET studies [30], was calculated from the estimated micro-parameters for evaluating the two methods.

A. Comparison between direct and indirect reconstructions

Fig. 2 shows the true and reconstructed K_i images by the direct method and indirect method. The regularization parameter used in this case was $\beta = 3 \times 10^{-4}$. In comparison, the result of the direct reconstruction is less noisy. The corresponding bias and variance images of the reconstructions are shown in Fig. 3. While the two methods have similar biases, the variance of the direct reconstruction is much lower than that of the indirect method. Fig. 4 plots the total squared bias against the total variance in the whole brain, by varying the regularization parameter β with $\beta = 1 \times 10^{-2}$ corresponding to the leftmost point and $\beta = 0$ to the rightmost point. It shows that the direct reconstruction results in less variance at any given bias level than the indirect method.

We also calculated the bias versus standard deviation tradeoff curve for region of interest (ROI) quantification, which takes into account of spatial correlation between image pixels. The ROI was chosen to be the tumor region. The trade-off curves are plotted in Fig. 5. Again, the standard deviation of the direct reconstruction is less than that of the indirect method at all bias levels.

All of these results indicate that the direct reconstruction has a better quantification performance for K_i than the indirect reconstruction. While direct reconstruction using the Patlak model [31]–[33] can also generate K_i image for irreversible tracers, the Patlak model itself may introduce bias when applied to reversible tracers with $k_4 > 0$ (see [39] for example). In addition to the estimation of K_i, the proposed method can offer better parametric images of other kinetic parameters as well. As an example, Fig. 6 shows the reconstructed k_3 images by the direct approach and the indirect approach, both with the same regularization parameter $\beta = 0.0003$. Cleary the direct reconstruction result has much less noise than the indirect reconstruction result while maintaining the same spatial resolution.

B. Comparison with previous optimization transfer algorithm

The convergence rate of the proposed algorithm (the optimization transfer with EM surrogate, OT-EM) was compared with one of our previous algorithms (optimization transfer with separable paraboloidal surrogate, OT-SP) [13]. We used two different measures to evaluate the convergence speed: normalized difference of the penalized log-

likelihood and mean squared error (MSE) of reconstructed K_i images. The normalized likelihood difference at iteration n is calculated by

$$L_n = \frac{\Phi^{\mathrm{ref}} - \Phi_n}{\Phi^{\mathrm{ref}} - \Phi_1} \quad (40)$$

where Φ^{ref} is the reference value obtained at convergence. The advantage of using the normalized likelihood difference is that we can easily compare the convergence rate of different algorithms using a log-scale plot. The mean squared error of a reconstructed image at iteration n is calculated by

$$E_n = \frac{1}{n_j} \sum_{j=1}^{n_j} (x_j^n - x_j^{\mathrm{ref}})^2 \quad (41)$$

where x^n is the image reconstructed at iteration n, x^{ref} is the reference image obtained at convergence.

To investigate the effect of background (randoms and scatters) fraction *BF* on the convergence speed, we simulated three levels of background in the sinograms: *BF* = 10%, *BF* = 20% and *BF* = 40%. The OT-EM and OT-SP algorithms used the same regularization parameter $\beta = 3 \times 10^{-4}$ and started from the same uniform initials. The results of the OT-SP at iteration 5000 were used as the reference for calculating the normalized likelihood difference and mean squared error.

Fig. 7 shows the comparisons of the normalized likelihood difference and mean squared errors of the two direct algorithms as a function of iteration number. Both results show that the OT-EM algorithm is faster than the OT-SP algorithm. The speed-up of the OT-EM over the OT-SP becomes greater as the background fraction reduces, because the OT-SP algorithm uses a non-tight paraboloidal surrogate function and hence becomes slower when the fraction of background is low. When the background fraction is at 10%, the OT-SP requires more than 300 iterations to reach the same likelihood value and mean squared error as that achieved by the OT-EM with 50 iterations. The computational costs per iteration of the OT-SP and OT-EM are comparable. The former requires one more backprojection to compute the curvature of the paraboloidal surrogate, while the latter takes more computation time in the fitting step to calculate the logarithm. The real CPU times are about 9 seconds per iteration for the OT-EM and 10 seconds per iteration for the OT-SP, when implemented in MATLAB running on a 2.4 GHz CPU.

C. Comparison with other algorithms

We compared the OT-EM algorithm with two direct reconstruction algorithms proposed by others [11], [21]. Both algorithms solve the image space maximum likelihood problem using a weighted least squares formulation. Uniform weight was used in [11] and nonuniform weight derived from an one-step-late approach was used in [21]. To our knowledge no proof of convergence has been established for either of the algorithm.

Fig. 8 compares the normalized likelihood difference and mean squared errors of the OT-EM algorithm and the algorithms by [11] and [21] for the Zubal data. Three different count levels (0.2M, 2M and 20M) were studied. The regularization parameter β in the OT-EM was set to 0 because the other two algorithms are only applicable to maximum likelihood reconstruction. The result of the algorithm in [21] at iteration 5000 was used as the reference. The plots show that the OT-EM is faster than the algorithm in [11] at all the count levels and can achieve faster convergence and lower MSE than the algorithm in [21] for lower-count data.

To demonstrate the potential non-monotonicity of the empirical algorithms, we also simulated a two-pixel toy example using a system matrix $P=\left[\begin{array}{ccc} 0.5 & 0.8 & 0.2 \\ 0.5 & 0.2 & 0.8 \end{array}\right]^{T}$ and two-tissue compartment model with the kinetic parameters $[f_v, K_1, k_2, k_3, k_4]^T = [0.01, 0.1, 0.01, 0.01, 0.01]^T$ for pixel 1 and $[f_v, K_1, k_2, k_3, k_4]^T = [0.01, 1.0, 1.0, 0.01, 0.01]^T$ for pixel 2. The input function and scanning schedule are the same as those used in the Zubal phantom simulation. One thousand events were simulated with a 20% background. Initial guess was 0.03 for all of the kinetic parameters. The lower bound for parameter estimate was set to 10^{-5} for all the parameters and the upper bound was 1.0 for f_v and 2.0 for K_1 to k_4.

Fig. 9(a) shows the normalized likelihood difference of the OT-EM and the algorithms in [11] and [21], using the result of the algorithm in [21] at iteration 5000 as the reference point. The trajectories of the estimated k_2 of the two pixels are plotted in figure 9(b) for the three algorithms, where the isocontours of the likelihood were drawn by varying k_2 parameter values while fixing other parameters at their reference values. The estimates by the algorithms in [11] and [21] at iteration 5000 are also marked in the figure. Both the OT-EM and the algorithm in [21] converged to the same optimum solution, but the OT-EM was much faster. The algorithm in [11] was non-monotonic and converged to a solution away from the optimum.

V. Application to Real 4D Data

We have applied the OT-EM algorithm to a dynamic brain scan of a nonhuman primate. The data were acquired on a microPET P4 scanner (Siemens, Knoxville, TN). The radiotracer was ^{11}C-SCH 23390 which binds to dopamine D1 receptors in the brain [34]. The scan consists of 20 time frames over 45 minutes: 10×60 s, 5×120 s, 5×300 s. A transmission scan was used for estimating the attenuation map. Randoms were pre-corrected using the delayed window technique. The expected scatter sinograms were estimated by the single-scatter simulation method [35]. Images were reconstructed using an array of $128 \times 128 \times 63$ voxels with a voxel size of $1.0 \times 1.0 \times 1.2$ mm^3.

The simplified reference tissue model (SRTM) was used for kinetic fitting of the ^{11}C-SCH 23390 data [36]. The cerebellum was chosen as the reference region. The tissue time activity curve is modeled by [37], [38]

$$C_T(t) = \theta_1 C_r(t) + \theta_2 C_r(t) \otimes \exp(-\theta_3 t). \quad (42)$$

where θ_1 is the relative delivery rate from plasma to tissue, θ_2 and θ_3 can be interpreted as the transport rates of the target tissue relative to the reference region. $C_r(t)$ is the reference TAC. The distribution volume ratio (DVR) of interest is calculated by

$$\mathrm{DVR}=\theta_1+\theta_2/\theta_3, \quad (43)$$

which is related to the binding potential (BP) via DVR = BP + 1.

We used both the indirect and direct methods to reconstruct the DVR parametric image. In the indirect method, the regular MAP-EM algorithm with 200 iterations was used to reconstruct dynamic emission images, then followed by the LM fitting with 100 iterations to obtain the weighted least squares estimates of the kinetic parameters θ_1, θ_2 and θ_3. The weighting factor in the fitting was set to the squared time duration divided by the total count of each frame [37]. The initial estimates in both methods were $\theta_1 = 1.0$, $\theta_2 = 1.0$, $\theta_3 = 0.1$. The lower and upper bounds of $(\theta_1, \theta_2, \theta_3)$ were set to $[10^{-5}, 5.0]$, $[10^{-5}, 5.0]$ and $[0.03, 0.6]$, respectively. In the direct method, the OT-EM algorithm was run for 200 iterations with two sub-iterations of the modified LM algorithm in the penalized likelihood fitting step.

Region-of-interest (ROI) quantification was performed on the DVR images estimated using the direct and indirect reconstruction methods. To define ROIs, the dynamic sinograms were summed up and reconstructed to form a single static image. One slice of the static reconstruction is shown in figure 10. Two ROIs were drawn in the striatal regions and four spheres of 8 mm in diameter were drawn in the brain background. Mean uptakes in the striatal regions and the average standard deviation in the background regions were computed for comparison.

Figure 11 shows the DVR images reconstructed using the indirect and direct methods. Both methods used the same initial estimates of the kinetic parameters. The regularization parameter β for the indirect reconstruction was 100. The regularization parameter for the direct reconstruction was $\beta = 200$ in order to match the mean value of the striatal regions to that of the indirect reconstruction. Fig. 12 plots the mean uptake of the striatal ROIs versus background noise curve by varying the regularization parameter with $\beta = 0$ corresponding to the rightmost point and $\beta = 1000$ corresponding to the leftmost point. The second rightmost point on the indirect curve and the third rightmost point on the direct curve correspond to the images shown in figure 11, respectively. In comparison, the direct reconstruction reduces noise in the brain at any matched mean value of the striatum ROIs.

The brain data were also reconstructed using the OT-SP algorithm with $\beta = 200$. The penalized likelihood of the proposed OT-EM is compared with that of the OT-SP as shown in Fig. 13. Both algorithms started from the same initialization. The OT-EM is clearly faster than the OT-SP algorithm.

VI. Conclusions

This paper presents a new optimization transfer approach for penalized likelihood direct reconstruction of nonlinear parametric images using the EM surrogate functions. Each iteration of the proposed algorithm consists of an EM-like dynamic image update, a spatial image smoothing, and a pixel-wise penalized likelihood kinetic fitting. The algorithm does not require a positive background in sinogram data and hence complements our previous generalized algorithm for direct reconstruction. Computer simulations validate that the direct reconstruction algorithm can have a better quantification performance than the indirect algorithm for estimation of parametric images. Application of the algorithm to a real brain scan also showed promising results.

Acknowledgments

References

1. Kordower JH, Emborg ME, Bloch J, et al. Neurodegeneration prevented by lentiviral vector delivery of GDNF in primate models of Parkinson's disease. Science. 2000; 290(5492):767–773. [PubMed: 11052933]

2. Gill SS, Patel NK, Hotton GR, et al. Direct brain infusion of glial cell line-derived neurotrophic factor in Parkinson disease. Nature Medicine. 2003; 9(5):589–595.

3. Visser E, Philippens M, Kienhorst L, Kaanders J, Corstens F, de Geus-Oei L, Oyen W. Comparison of tumor volumes derived from glucose metabolic rate maps and SUV maps in dynamic 18F-FDG PET. Journal of Nuclear Medicine. 2008; 49(6):892–898. [PubMed: 18483085]

4. Tsoumpas C, Turkheimer FE, Thielemans K. A survey of approaches for direct parametric image reconstruction in emission tomography. Med Phys. 2008; 35(9):3963–3971. [PubMed: 18841847]

5. Rahmim A, Tang J, Zaidi H. Four-dimensional image reconstruction strategies in dynamic PET: Beyond conventional independent frame reconstruction. Med Phy. 2009; 36(8):3654–3670.

6. Matthews J, Bailey D, Price P, Cunningham V. The direct calculation of parametric images from dynamic PET data using maximum-likelihood iterative reconstruction. Physics in Medicine and Biology. 1997; 42(6):1155–1173. [PubMed: 9194135]

7. Reutter BW, Gullberg GT, Huesman RH. Kinetic parameter estimation from attenuated SPECT projection measurements. IEEE Transactions on Nuclear Science. 1998; 45(6):3007–3013.

8. Meikle SR, Matthews JC, Cunningham VJ, Bailey DL, Livieratos L, Jones T, Price P. Parametric image reconstruction using spectral analysis of PET projection data. Physics in Medicine and Biology. 1998; 43(3):651–666. [PubMed: 9533143]

9. Vanzi E, Formiconi AR, Bindi D, Cava GL, Pupi A. Kinetic parameter estimation from renal measurements with a three-headed SPECT system: A simulation study. IEEE Transactions on Medical Imaging. 2004; 23(3):363–373. [PubMed: 15027529]

10. Kamasak ME, Bouman CA, Morris ED, et al. Direct reconstruction of kinetic parameter images from dynamic PET data. IEEE Trans Med Im. 2005; 24(5):636–650.

11. Reader A, Matthews J, Sureau F, et al. Iterative kinetic parameter estimation within fully 4D PET image reconstruction. IEEE Nucl Sci Syp & Med Im Conf (3). 2006:1752–56.

12. Yan J, Wilson BP, Carson RE. Direct 4D list mode parametric reconstruction for PET with a novel EM algorithm. IEEE Nucl Sci Syp & Med Im Conf. 2008

13. Wang G, Qi J. Generalized algorithms for direct reconstruction of parametric images from dynamic PET data. IEEE Trans Med Im. 2010; 28(11):1717–1726.

14. Lange K, Hunter DR, Yang I. Optimization transfer using surrogate objective functions. Journal of Computational and Graphical Statistics. 2000; 9(1):1–20.

15. Fessler JA, Erdogan H. A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction. 1998 IEEE Nuclear Science Syposium and Medical Imaging Conference. 1998; 2:1132–5.

16. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B. 1977; 39(1):1–38.

17. Wang G, Qi J. Acceleration of direct reconstruction of linear parametric images using nested algorithms. Phys Med Bio. 2010; 55(5):1505–1517. [PubMed: 20157226]

18. Shepp LA, Vardi Y. Maximum likelihood reconstruction for emission tomography. IEEE Transactions on Medical Imaging. 1982; 1(2):113–122. [PubMed: 18238264]

19. Lange K, Carson RE. EM reconstruction algorithms for emission and transmission tomography. Journal of Computer Assisted Tomography. 1984; 8(2):306–316. [PubMed: 6608535]

20. Wang G, Qi J. Direct reconstruction of nonlinear parametric images for dynamic PET using nested optimization transfer. IEEE Nuclear Science Symposium and Medical Imaging Conference. Nov. 2010 (abstract only).

21. Matthews JC, Angelis GI, Kotasidis FA, Markiewicz PJ, Reader AJ. Direct reconstruction of parametric images using any spatiotemporal 4D image based model and maximum likelihood expectation maximisation. IEEE Nuclear Science Symposium and Medical Imaging Conference. Nov.2010 :2435–2441.

22. Schmidt KC, Turkheimer FE. Kinetic modeling in positron emission tomography. Quarterly Journal of Nuclear Medicine. 2002; 46(1):70–85. [PubMed: 12072847]

23. Carson RE, Lange K. The EM parametric image reconstruction algorithm. Journal of America Statistics Association. 1985; 80(389):20–22.

24. De Pierro AR. A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. IEEE Transactions on Medical Imaging. 1995; 14(1):132–137. [PubMed: 18215817]

25. Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics. 1963; 11(2):431–441.

26. Madsen K, Nielsen HB, Tingleff O. Methods for non-linear least squares problems. Tech rep, Technical University of Denmark. 2004

27. Zubal IG, Harrell CR, Smith EQ, Rattner Z, Gindi G, Hoffer PB. Computerized three-dimensional segmented human anatomy. Medical Physics. 1994; 21(2):299–302. [PubMed: 8177164]

28. Feng DG, Wong KP, Wu CM, Siu WC. A technique for extracting physiological parameters and the required input function simultaneously from PET image measurements: Theory and simulation study. IEEE Transactions on Information Technology in Biomedicine. 1997; 1(4):243–254. [PubMed: 11020827]

29. Spence AM, Muzi M, Graham MM, et al. Glucose metabolism in human malignant gliomas measured quantitatively with PET, 1-[C-11]glucose and FDG: analysis of the FDG lumped constant. Journal of Nuclear Medicine. 1998; 39(3):440–448. [PubMed: 9529289]

30. Huang SC, Phelps ME, Hoffman EJ, Kuhl DE. Error sensitivity of fluorodeoxyglucose method for measurement of cerebral metabolic-rate of glucose. Journal of Cerebral Blood Flow and Metabolism. 1981; 1(4):391–401. [PubMed: 6976976]

31. Wang G, Qi J. Maximum a posteriori reconstruction of the Patlak parametric image from sinograms in dynamic PET. Physics in Medicine and Biology. 2008; 53(3):593–604. [PubMed: 18199904]

32. Tsoumpas C, Turkheimer FE, Thielemans K. Study of direct and indirect parametric estimation methods of linear models in dynamic positron emission tomography. Medical Physics. 2008; 35(4):1299–1309. [PubMed: 18491524]

33. Tang J, Kuwabara H, Wong DF, Rahmim A. Direct 4D reconstruction of parametric images incorporating anato-functional joint entropy. Physics in Medicine and Biology. 2010; 55(15): 4261–4272. [PubMed: 20647600]

34. Farde L, Halldin C, Sone-Elander S, Sedvall G. PET analysis of human dopamine receptor subtypes using ^{11}C-SCH 23390 and ^{11}C-raclopride. Psychopharmacology. 1987; 92(3):278–284. [PubMed: 2957716]

35. Watson CC. New, faster, image-based scatter correction for 3D PET. IEEE Transactions on Nuclear Science. 2000; 47(4):1587–1594.

36. Lammertsma AA, Hume SP. Simplified reference tissue model for PET receptor studies. Neuroimage. 1996; 4(3):153–158. [PubMed: 9345505]

37. Gunn RN, Lammertsma AA, Hume SP, Cunningham VJ. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. Neuroimage. 1997; 6(4):279–287. [PubMed: 9417971]

38. Ichise M, Liow JS, Lu JQ, Takano A, Carson RE. Linearized reference tissue parametric imaging methods: application to [11C]DASB Positron Emission Tomography studies of the serotonin transporter in human brain. Journal of Cerebral Blood Flow and Metabolism. 2003; 22(10):1271–1281. [PubMed: 12368666]

39. Yu AS, Lin HD, Huang SC, Phelps ME, Wu HM. Quantication of cerebral glucose metabolic rate in mice using 18F-FDG and small-animal PET. Journal of Nuclear Medicine. 2009; 50(6):966–973. [PubMed: 19443595]
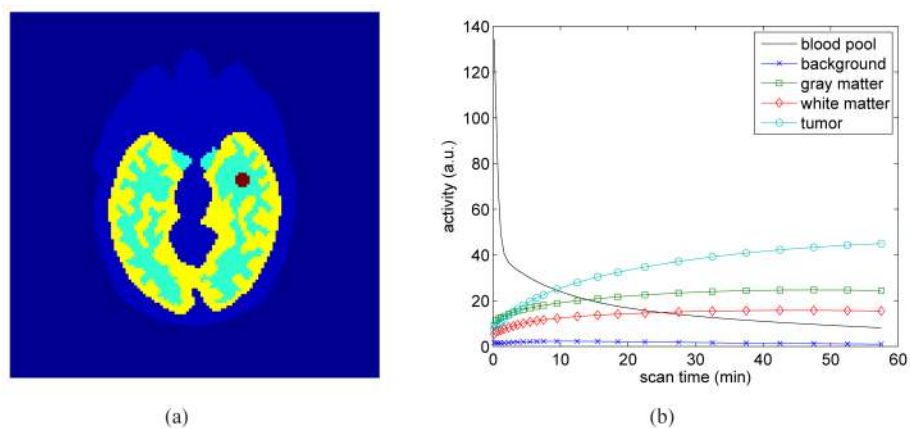
Fig. 1.
The simulation settings. (a) a Zubal brain phantom composed of gray matter, white matter and a small tumor; (b) regional time activity curves
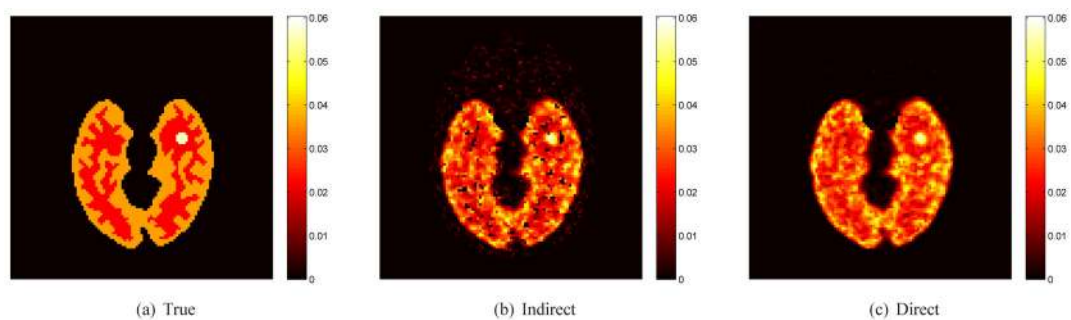
(a) True   (b) Indirect   (c) Direct

Fig. 2.
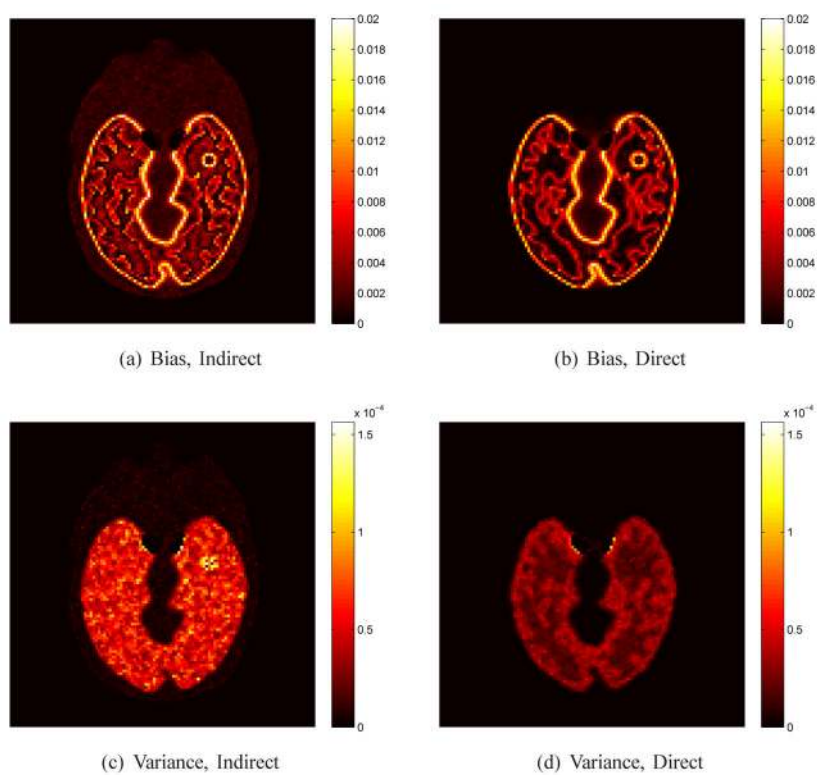True and reconstructed K_i parametric images using the indirect and direct algorithms ($\beta = 0.0003$).

Fig. 3.
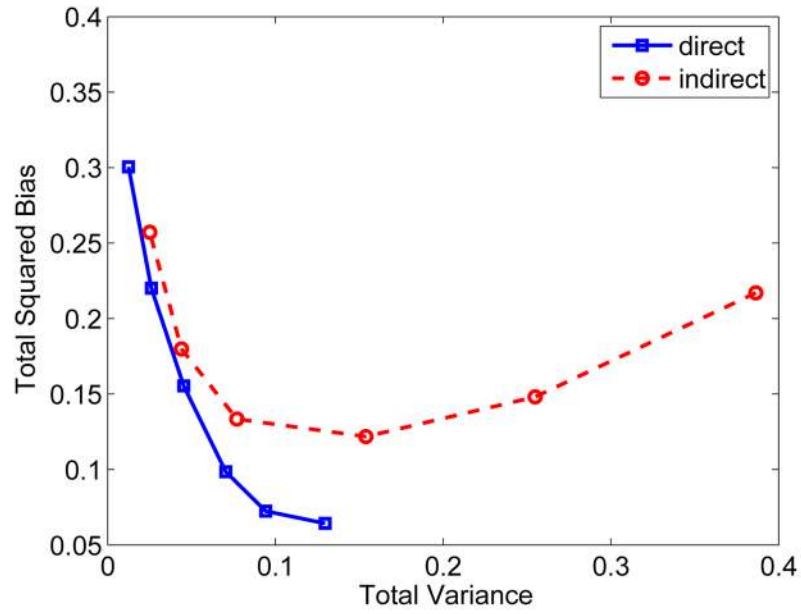The bias and variance images of K_i image reconstructed using the direct and indirect algorithms ($\beta = 0.0003$).

Fig. 4.
The total squared bias versus variance tradeoff of the influx rate K_i image reconstructed using the direct and indirect algorithms in the brain. Different points on the curves were obtained by varying the parameter β.
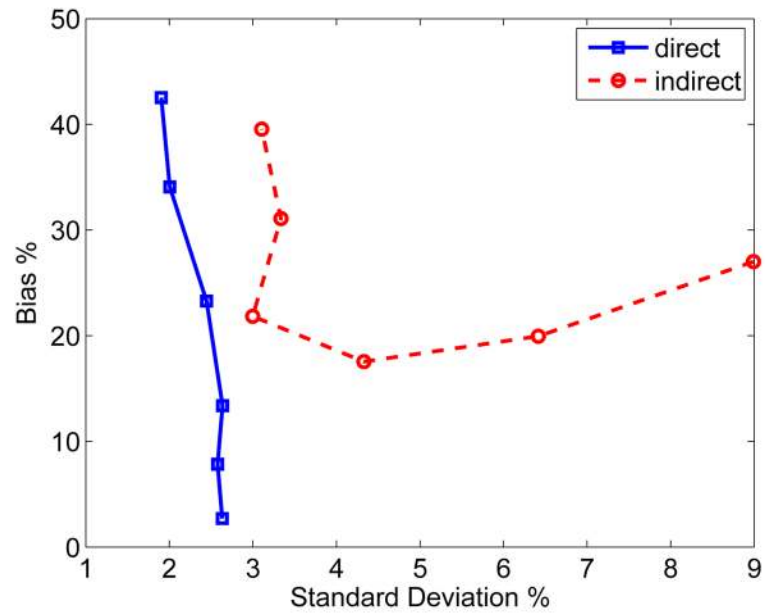
Fig. 5.
The bias versus standard deviation tradeoff of tumor ROI quantification on K_i images reconstructed using the direct and indirect algorithms. Different points on the curves were obtained by varying the parameter β.
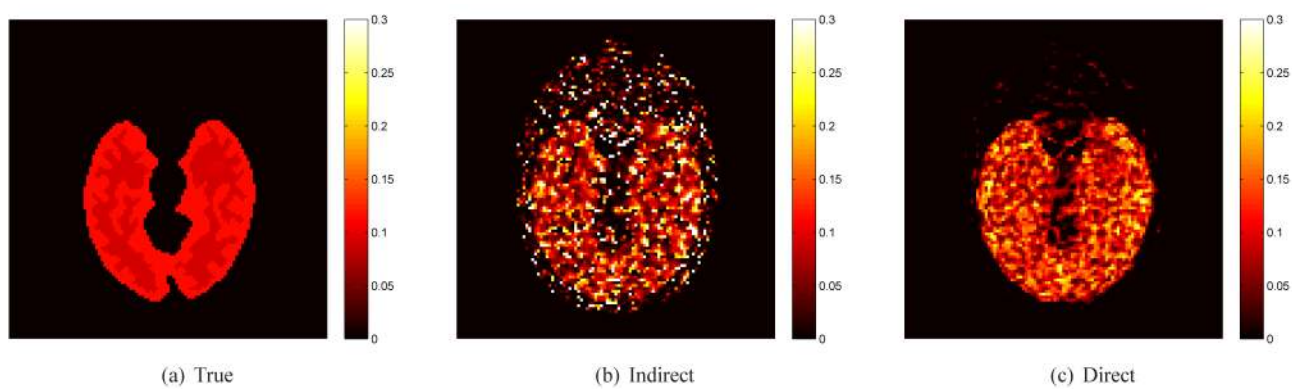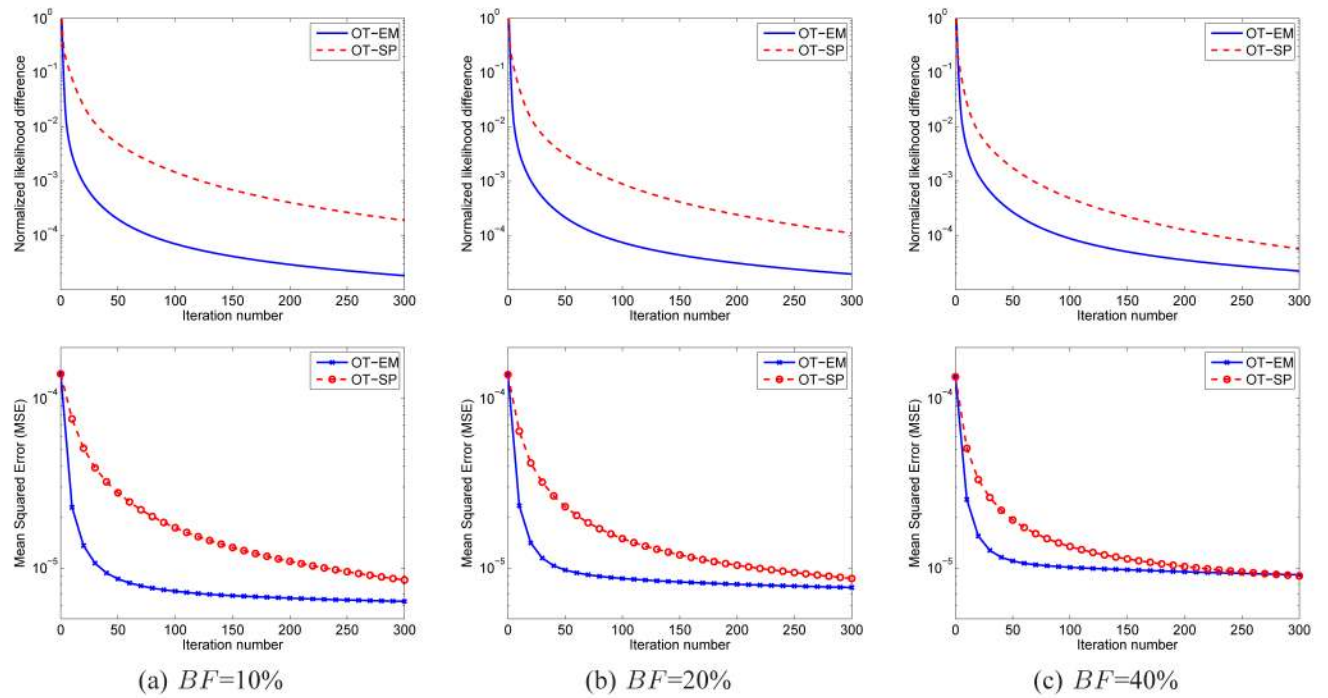
(a) True   (b) Indirect   (c) Direct

Fig. 6.
The true and reconstructed k_3 images by the indirect and direct algorithms ($\beta = 0.0003$).

Fig. 7.
Comparisons of normalized likelihood (first row) and mean squared error (second row) of the penalized likelihood direct reconstructions using the optimization transfer approaches with the SP surrogate and with the EM surrogate for three different background fractions: (a) $BF = 10\%$, (b) $BF = 20\%$ and (c) $BF = 40\%$.
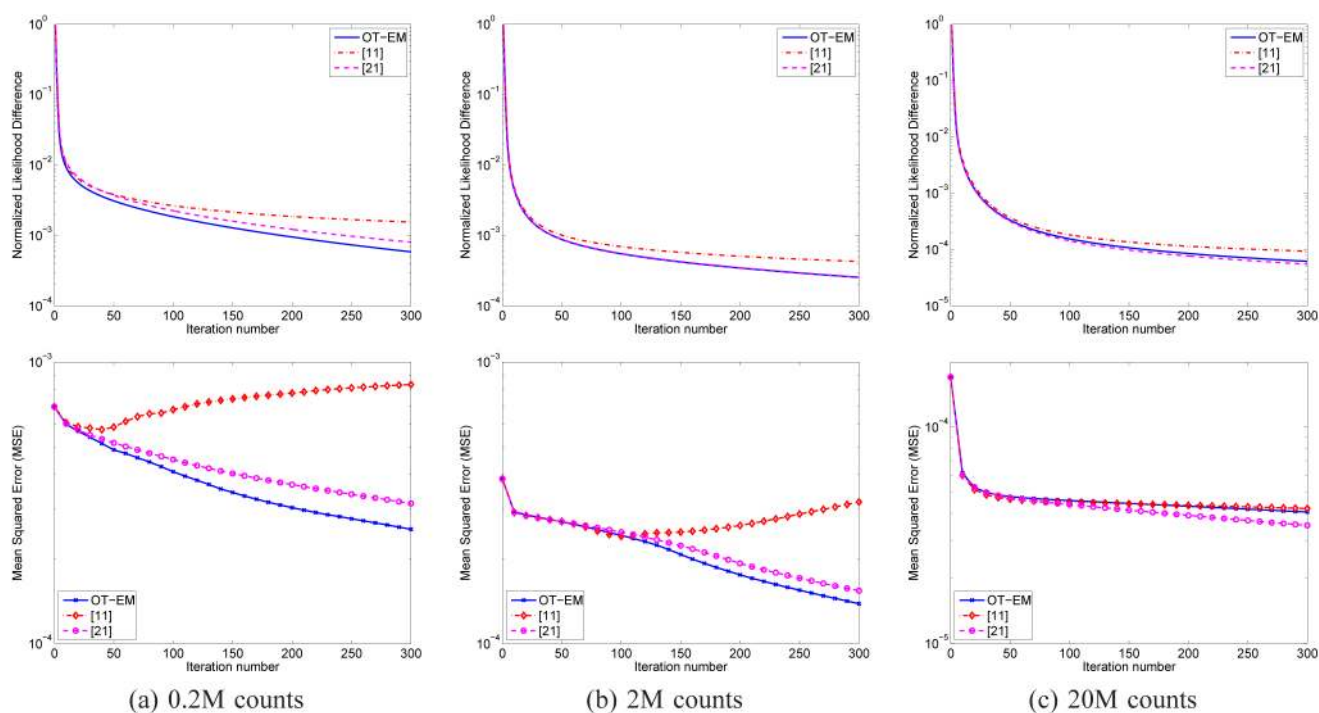
Fig. 8.
Comparison of the normalized likelihood difference (first row) and mean squared error (second row) of the maximum likelihood reconstruction using the OT-EM algorithm and the algorithms in [11] and [21] at three different count levels.
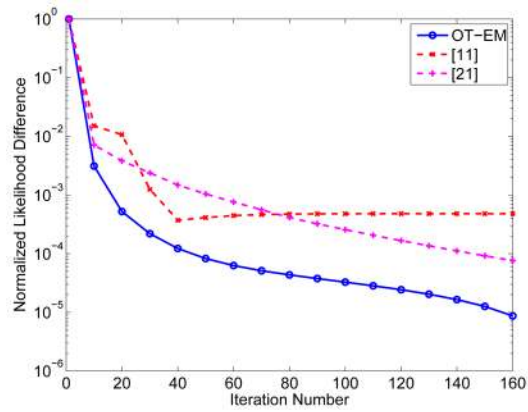
(a) normalized likelihood
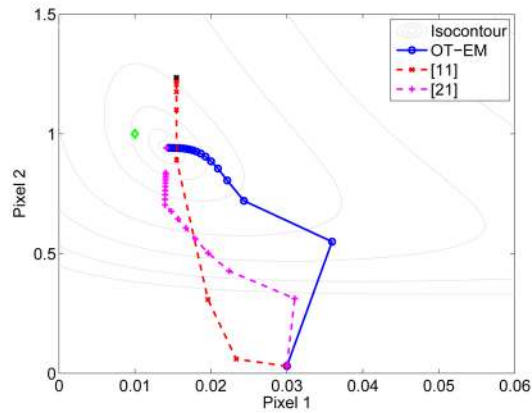
(b) iterative trajectory

Fig. 9.
Comparison of (a) the likelihood and (b) the trajectory of estimated k_2 by different direct reconstruction algorithms in the two-pixel simulation. The noise-free ground truth is marked by $\diamond$ in (b). All algorithms started from the same initialization.
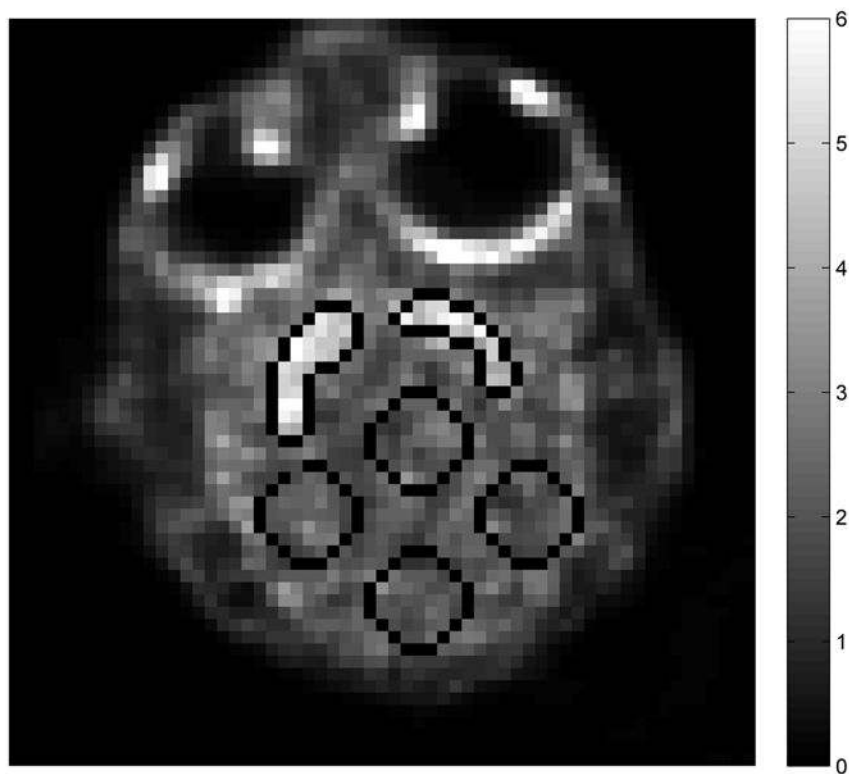
Fig. 10.
Striatal targets and background ROIs drawn on the reconstructed image of the sinogram summed over the whole dynamic scans.
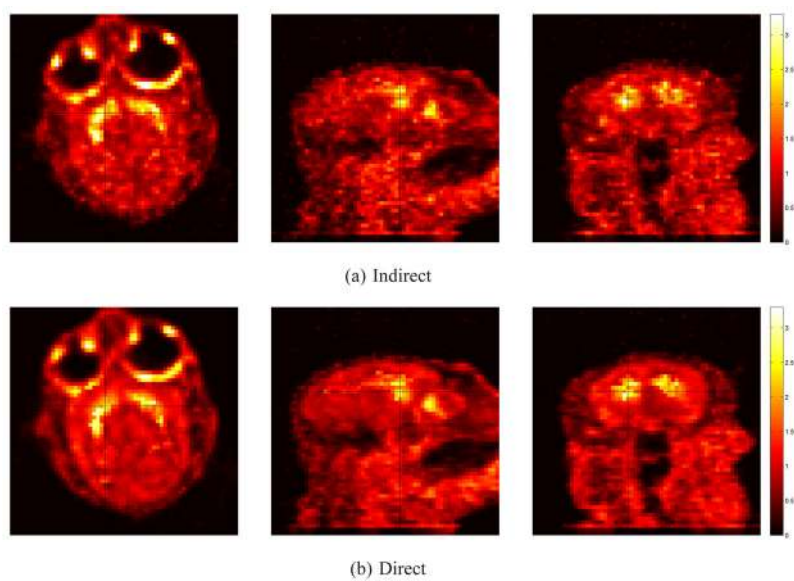
(a) Indirect

(b) Direct

Fig. 11.
The estimated DVR images of the real primate brain data using (a) the indirect and (b) the direct methods. The three columns represent the transverse, sagittal and coronal slices of each 3D volume, respectively.

Fig. 12.
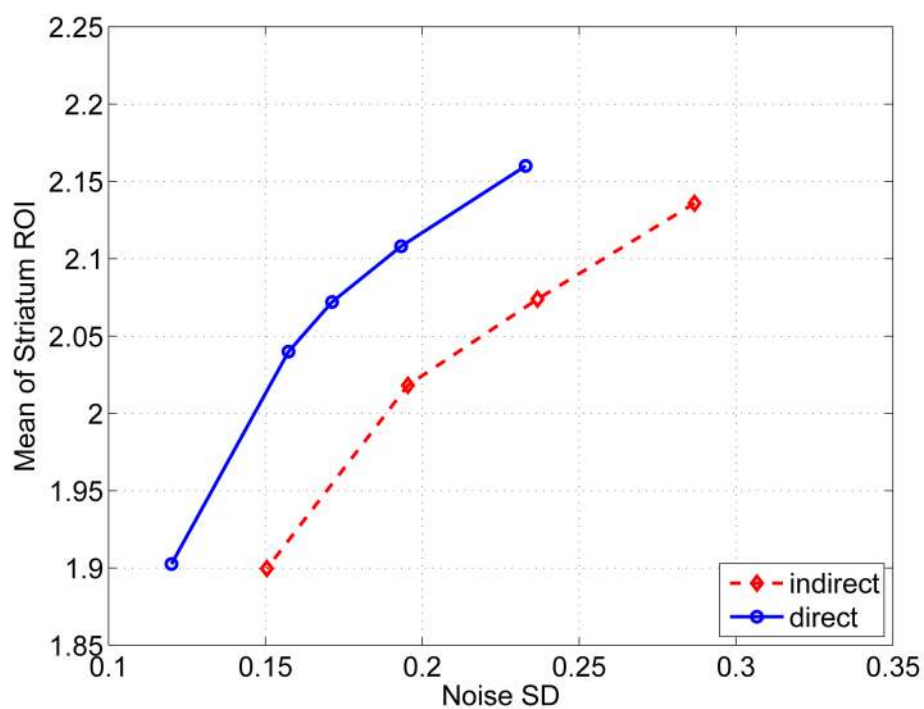ROI mean of the striatum regions versus background standard deviation of the indirect and direct reconstructions. The curves are plotted by varying the regularization parameter β.
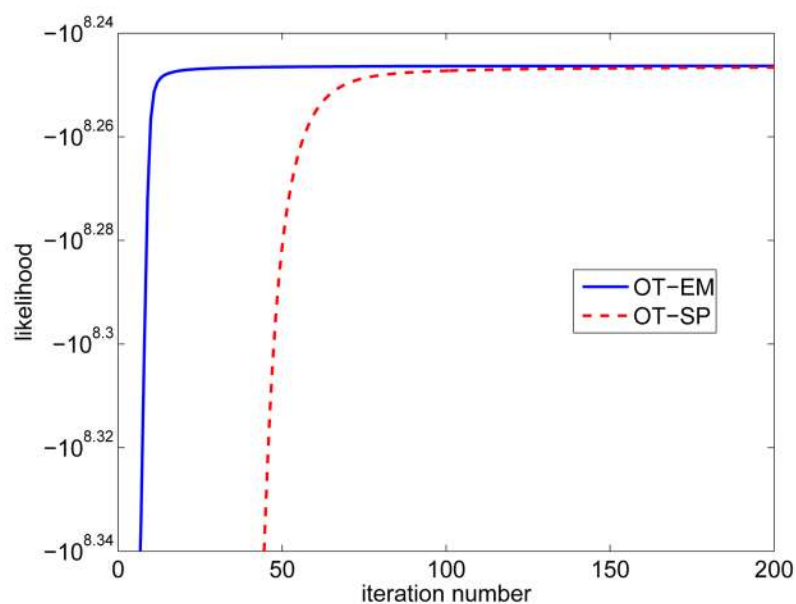
Fig. 13.
Comparison of convergence rate of the penalized likelihood direct reconstruction using the OT-SP and OT-EM algorithms for the real 4D primate data. Both approaches start from the same initialization.

TABLE I

Comparison between the proposed algorithm and the previous algorithm for direct reconstruction.

	Prevous Algorithms [13]	Proposed Algorithm
Surrogate Function	paraboloidal surrgoate	EM surrogate
Monotonic Convergence	yes	yes
Reconstruction Step	paraboloidal update	EM update
Smoothing Step	spatial filtering	spatial filtering
Fitting Step	least squares fitting by the Levenberg-Marquardt (LM) algorithm	penalized likelihood fitting by the modified LM algorithm
Compartment type	any types	any types
Background Events	must be positive	no requirement

TABLE II

The kinetic parameters used in the FDG PET simulation. The unit of K_1-k_4 is min^{-1}.

	f_v	K_1	k_2	k_3	k_4
Background	0.01	0.010	0.100	0.001	0.001
Gray matter	0.05	0.116	0.254	0.116	0.011
White matter	0.03	0.059	0.149	0.090	0.013
Tumor	0.04	0.088	0.055	0.096	0.001

Algorithm 1

The generalized algorithm for direct reconstruction of parametric images

---

1:    Input parameters: Maximum iteration number for reconstruction MaxIter; Regularization parameter β

2:    Initialize the parametric images $\boldsymbol{\theta}$,

3:    Compute the dynamic image $\boldsymbol{x} := \{x_m(\boldsymbol{\theta}_j)\}$ by Eq.(3);

4:    **for** $n = 1$ to MaxIter **do**

5:        Frame-wise EM image reconstruction:

$$\hat{x}_{jm}^{\mathrm{em}} := \frac{x_m(\boldsymbol{\theta}_j)}{p_j} \sum_{i=1}^{n_i} p_{ij} \frac{y_{im}}{\bar{y}_{im}(\boldsymbol{\theta})}$$

where $p_j = \sum_{i=1}^{n_i} p_{ij}$ and $\bar{y}_{im}(\boldsymbol{\theta})$ is calculated by Eq. (4);

6:        Frame-wise image smoothing if $\beta > 0$:

$$\hat{x}_{jm}^{\mathrm{reg}} := \frac{1}{2w_{jm}^{\mathrm{reg}}} \sum_{l \in \mathcal{N}_j} \gamma_{jl} \left(x_m(\boldsymbol{\theta}_j) + x_m(\boldsymbol{\theta}_l) \right),$$

where the weight w_{jm}^{reg} is calculated by Eq.(21);

7:        Pixel-wise penalized likelihood fitting

$$(\boldsymbol{\theta}_j, \boldsymbol{x}_j) := \mathrm{PLFit} \left(\boldsymbol{\theta}_j, \hat{\boldsymbol{x}}_j^{\mathrm{em}}, \hat{\boldsymbol{x}}_j^{\mathrm{reg}}, \frac{\beta}{p_j} \boldsymbol{w}_j^{\mathrm{reg}} \right)$$

8:    **end for**

9:    **return** the parametric images $\boldsymbol{\theta}$

Algorithm 2

PLFit $\phi_0, f^{\text{em}}, f^{\text{reg}}, \omega^{\text{reg}}$: the Levenberg-Marquardt algorithm for penalized likelihood fitting

---

1:      Set the parameters τ and ε to small values, e.g. $\tau = 10^{-3}$, $\varepsilon = 10^{-9}$

2:      Set initals: $v = 2$, $\phi = \phi_0$

3:      Calculate the gradient g and Hessian H by Eq. (31)

4:      $\mu := \tau \max_p \{[H]_{pp}\}$

5:      **for** $k = 1$ to $k_{\max}$ **do**

6:          Get ϕ_{new} by solving the quadratic optimization (34)

7:          **if** $\|\phi_{\text{new}} - \phi\| \leq \varepsilon (\|\phi\| + \varepsilon)$ **then**

8:              **return** ϕ and $\{x_m(\varphi)\}_{m=1}^{n_m}$

9:          **else**

10:             $\rho := \frac{q(\varphi) - q(\varphi_{\text{new}})}{\tilde{q}(\varphi) - \tilde{q}(\varphi_{\text{new}})}$

11:         **if** $\rho > 0$ **then**

12:             $\phi := \phi_{\text{new}}$

13:             Update g and H using Eq. (31)

14:             $\mu := \mu \max\{\frac{1}{3}, 1 - (2\rho - 1)^3\}; \; v := 2$

15:         **else**

16:             $\mu := \mu v; \; v := 2v$

17:         **end if**

18:     **end if**

19: **end for**

20:     return ϕ and $\{x_m(\varphi)\}_{m=1}^{n_m}$