Joint reconstruction of dynamic PET activity and kinetic parametric images using total variation constrained dictionary sparse coding

Haiqing Yu¹, Shuhang Chen¹, Yunmei Chen² and Huafeng Liu¹,³

¹ College of Optical Science and Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, People’s Republic of China
² Department of Mathematics, University of Florida, 458 Little Hall Gainesville, FL 32611-8105, United States of America

E-mail: liuhf@zju.edu.cn

Received 11 October 2016, revised 14 January 2017
Accepted for publication 20 February 2017
Published 24 March 2017

Abstract

Dynamic positron emission tomography (PET) is capable of providing both spatial and temporal information of radio tracers in vivo. In this paper, we present a novel joint estimation framework to reconstruct temporal sequences of dynamic PET images and the coefficients characterizing the system impulse response function, from which the associated parametric images of the system macro parameters for tracer kinetics can be estimated. The proposed algorithm, which combines statistical data measurement and tracer kinetic models, integrates a dictionary sparse coding (DSC) into a total variational minimization based algorithm for simultaneous reconstruction of the activity distribution and parametric map from measured emission sinograms. DSC, based on the compartmental theory, provides biologically meaningful regularization, and total variation regularization is incorporated to provide edge-preserving guidance. We rely on techniques from minimization algorithms (the alternating direction method of multipliers) to first generate the estimated activity distributions with sub-optimal kinetic parameter estimates, and then recover the parametric maps given these activity estimates. These coupled iterative steps are repeated as necessary until convergence. Experiments with synthetic, Monte Carlo generated data, and real patient data have been conducted, and the results are very promising.

¹ Author to whom any correspondence should be addressed.

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Keywords: dynamic PET reconstruction, joint reconstruction, direct estimation, total variation, dictionary sparse coding

(Some figures may appear in colour only in the online journal)

1. Introduction

It has been the fundamental goal of many efforts in the dynamic positron emission tomography (PET) imaging community to elucidate underlying biological and physiological processes in living tissues through reconstructing the spatiotemporal distribution of an injected radio tracer. With a suitable kinetic model, such as a compartmental model, the additional temporal information can be extracted from measured data—as in direct methods, or reconstructed distribution—as in indirect schemes, and can be generalized into physiological parametric images which are of greater clinical use in tracking treatments of diseases and tumor diagnosis compared to conventional static PET imaging techniques [1, 2]. In practice, however, the data collecting process for each time frame is limited, due to the temporal resolution demand of acquiring a sequence of data at several time points during a continuous period of time. Therefore, in dynamic PET reconstruction problems, the inner connection between different time frames can no longer be ignored, and the use of conventional image reconstruction methods—which deal with the measured data frame by frame—will suffer from the low counts. In this work, tracer kinetics are integrated into the traditional frame-by-frame reconstruction algorithm, and thereby a direct estimation of the kinetic parameters of radio tracer is simultaneously performed during the optimization of the activity distribution.

The conventional image reconstruction problem has been considered for many years by many researchers. The majority of active reconstruction falls into statistical iterative configurations [3], being usually a specific model for the measurement statistics, and the physical systems with designed priors [4]. Among these iterative algorithms, two methods are often mentioned: these are the maximum a posteriori (MAP) [5–9] and the maximum likelihood (ML) based reconstruction algorithms, which have been widely examined by researchers ever since Shepp and Vardi presented their work in 1982 [10–16]. Based on the research in [17, 18], with extra constraints, it is possible to reduce image corruption caused by low counts (via such methods as TV regularization [19, 20] and hierarchical regularization [21], which is also effective in edge-preserving) or further achieve a joint reconstruction task by integrating other information into the image reconstruction problems, as described in [22, 23].

Basically, there are two groups of approaches to the reconstruction of tracer kinetics from dynamic PET data: indirect and direct reconstruction. Indirect reconstruction efforts treat the image reconstruction and the dynamic parameter estimation as two different problems. This is based on the premise that the activity distributions are known as prior information, and the challenge is to estimate the kinetic parameters by fitting a mathematical model to the time activity curve (TAC) extracted from activity distributions [24]. On the other hand, the direct method combines those two steps into one problem, and the parametric images are directly estimated from the raw emission data [25–29]. In 1985, Carson and Lange first presented their research on direct parametric reconstruction based on expectation maximization (EM) algorithm [30]. Inspired by their work, later in 2005, Kamasak et al presented their research on the direct reconstruction of kinetic parameter images [31]. According to Kamasak, direct reconstruction methods often outperform indirect methods, whereas the solution becomes more complicated. In both direct and indirect methods, a suitable model is needed to fit the tracer kinetics and represent TACs by mathematical functions. Among these models, the nonlinear
compartmental model is widely used to represent the connection between kinetic parameters and the underlying physiological processes. Representative works include spectral analysis [32], DEPICT [33] and sparse Bayesian learning [34]. Based on the theory of the compartment model, TACs are the sum of the activity concentration within each compartment. By pre-defining a suitable range of the basis function, the model fitting procedure can be treated as a regression problem with a biologically reasonable sparsity prior, which makes the model more robust to noise.

Distinct from either indirect or direct estimation method, in which only one variable of activity distribution and parametric image can be directly connected with the raw sinogram data, we propose a joint estimation framework, tackling PET image reconstruction along with the direct tracer kinetic estimation process. The algorithm for our whole framework iterates between the alternating direction minimization process (providing the current estimation of the activity image), updating coefficients for the compartmental model through iterative thresholding algorithm (solution of which gives the current estimation of the coefficients for the system impulse response function) and finally minimizing the objective function.

In this paper, the joint reconstruction problem couples the traditional frame-by-frame image reconstruction method and the direct data-driven estimation of parametric images. Unlike the traditional objective function, for the reconstruction of either activity maps or kinetic parameters, our data fitting term is formed from both measured data and a biologically based compartmental model, and regularization such as total variational (TV) is added so as to improve the robustness of the results [35, 36]. As for the dictionary sparse coding (DSC), it contains an over-completed dictionary composed of kinetic basis functions and an additional sparse constraint for the determination of the most related subset of the basis function. Moreover, since both statistical and biological constraints are added in our work, the results have the potential to give a more accurate and better fit for the tracer’s fate in vivo. The hypothesis that we can use emission sinograms to track the activity distribution and tracer kinetics simultaneously is confirmed by the results of our experiments using synthetic, Monte Carlo simulated, and real data. The visually and quantitative results for the activity distribution and the suitable macro parameter are given in the section for experiments.

2. Methodology

2.1. Problem formulation

2.1.1. TV based dynamic PET image reconstruction model. Unlike static PET, dynamic PET measurements are acquired as a sequence of temporal frames. During the processing of one time frame, a coincidence event comes from the annihilation process, which converts all the mass into energy and thus emits two photons, collected by annihilation coincidence detection. The counts of near-simultaneously acquired photons collected by the opposite detector pair are organized in the format of a sinogram. Hence for the mth time frame, in which \( m = 1, 2, \ldots, M \) (with \( M \) being the total number of the frames), the sinogram can be denoted by a vector \( y_m \in \mathbb{R}^I \) containing all the data collected by the \( I \) detector pairs in all, and the corresponding activity distribution for the mth time frame can also be denoted by a vector \( x_m \in \mathbb{R}^J \), where \( J \) is the total number of discrete pixels.

In general, the aim of dynamic PET reconstruction can be reduced to the estimation of a radioactivity map \( X \in \mathbb{R}^{J \times M} \) from the raw projection data \( Y \in \mathbb{R}^{I \times M} \) measured from the detector probes of PET through a transform
\[ Y = GX + \text{noise} \]  

(1)

where \( X = [x_1, x_2, \ldots, x_m, \ldots, x_M] \), \( Y = [y_1, y_2, \ldots, y_m, \ldots, y_M] \), \( G \in \mathbb{R}^{L \times J} \) is a geometry-based probability matrix and each entry \( g_{ij} \) of \( G \) represents the possibility of detecting an emission from pixel site \( j \) at detector probe \( i \).

Total variation regularization is then applied into the reconstruction and the sub-objective function for \( X \) can be formulated as a sum of TV regularizer and a data fitting term through the transform [37, 38]

\[ L(X) = \text{TV}(X) + \frac{\alpha}{2} \|GX - Y\|^2 \]  

(2)

in which \( \| \cdot \|^2 \) refers to the square of Euclidean norm and \( \alpha \) denotes a relative weighting parameter of data fidelity term \( \|GX - Y\|^2 \). TV(\cdot) is the total variation and TV(\( X \)) represents the sum of the discrete gradient of activity map \( X \) of every pixel \( j \) at each frame \( m \). It can be calculated through

\[ \text{TV}(X) = \sum_{m=1}^{M} \sum_{j=1}^{J} \| (Dx_m)_j \| \]  

(3)

where \( (Dx_m)_j \in \mathbb{R}^2 \) is defined as the discrete partial derivatives of \( x_m \) at \( j \)th pixel. In this work, a Gaussian based noise model, instead of the Poisson noise model, is used to unify the data fitting term and the TV norm, which will further simplify the solving process of activity distribution \( X \).

2.1.2. Modeling of tracer kinetics. One widely used model of tracer kinetic fitting is the compartmental model, which is tightly related to the underlying biochemical process [39]. Based on the theory of compartmental modelling, the time activity curve (TAC) is the sum of radioactivity within each tissue compartment obtained by convolving exponential functions with the blood input function or reference tissue input function.

Normally, a PET tracer compartmental model can be better defined as a plasma input model or a reference tissue input model, and it can be generally represented by a set of first-order equations, each of which describes the tracer concentration of each tissue compartment. The solution is that the impulse response function consists solely of exponentials and a delta function [39]:

\[ C_T(t) = \left[ \phi_0 \delta(t) + \sum_{l=1}^{q} \phi_l e^{-\theta_l t} \right] \otimes C_l(t) \]  

(4)

where \( C_T(t) \) represents the tissue concentration time course at time point \( t \), while \( C_l(t) \) refers to the input (plasma or reference tissue) concentration time course. \( q \) is the total number of tissue compartments and \( l \) represents the \( l \)th tissue compartment, \( \otimes \) is the convolution operator, \( \phi_0 \) and \( \theta_l \) represent, respectively, the coefficient and the exponent of the general impulse response function, for the \( l \)th compartment. \( \delta(t) \) is the Delta function and the convolution of \( \delta(t) \), and the input function represents the concentration time course when there is no tissue compartment involved.

The measurement data for dynamic PET are obtained as a sequence of temporal frames and the tissue concentration value \( x_{jm} \) for the \( j \)th pixel at the \( m \)th time frame can be generated through an integral over the individual frames from the continuous functions, i.e. the concentration time course of radioactivity in the target tissue. Therefore, the tissue radioactivity observations, in this case each entry \( x_{jm} \) of the matrix \( X \), can be further defined as
\[ x_{jm} = \frac{1}{t_m^e - t_m^s} \int_{t_m^s}^{t_m^e} C^j_t(t) \, dt \] (5)

in which \( t_m^s \) and \( t_m^e \) denote the start and end times of the \( m \)th frame and \( C^j_t(t) \) is defined as the tissue concentration time course \( C_T(t) \) for the \( j \)th pixel. The continuous function \( C^j_t(t) \) must be normalized to the frame length to correspond to the data sampling process.

Based on the equation (4), \( C_T(t) \) is the sum of the concentration in the target tissue and hence can be further expressed as an expansion on a basis

\[ C_T(t) = \sum_{\psi=0}^{N} \psi(t) \psi(t_\psi) \] (6)

where

\[ \psi(t) = C_j(t) \]
\[ \psi(t_\psi) = \int_{t_\psi}^{t_\psi + t} e^{-\theta(t - \tau)} C_I(\tau) \, d\tau \] (7)

in which \( N \) denotes the number of pre-defined basis functions, and corresponds to the scale set of \( \theta = \{ \theta_1, \theta_2, \ldots, \theta_N \} \) since it may be pre-chosen from a physiologically plausible range \( \theta_{\text{min}} \leq \theta \leq \theta_{\text{max}} \) and each \( \theta_i \) is spaced in a suitable manner within the proposal range so as to elicit a suitable coverage of the kinetic spectrum. In addition, the range of \( \theta \) is closely connected to the tracer, i.e. the same set of \( \theta \) may not fit for all situations. \( \psi_\phi \) is defined as the corresponding coefficient in \( C_T(t) \). The pre-defined basis functions, referring to the convolution of a set of exponential functions with the input function, are discrete, correspond to the time frames and are organized in the form of a matrix \( \Psi \in \mathbb{R}^{N \times M} \) as below, which is also known as the dictionary of the proposed framework [33].

\[
\Psi = \begin{bmatrix}
\psi_{10} & \psi_{11} & \cdots & \psi_{1N} \\
\vdots & \vdots & \ddots & \vdots \\
\psi_{m0} & \psi_{m1} & \cdots & \psi_{mN} \\
\vdots & \vdots & \ddots & \vdots \\
\psi_{M0} & \psi_{M1} & \cdots & \psi_{MN}
\end{bmatrix}
\]

\[
\psi_{m0} = \frac{1}{t_m^e - t_m^s} \int_{t_m^s}^{t_m^e} C_j(t) \, dt \\
\psi_{mc} = \frac{1}{t_m^e - t_m^s} \int_{t_m^s}^{t_m^e} \int_{t_\psi}^{t_\psi + t} e^{-\theta(t - \tau)} C_I(\tau) \, d\tau \, d\tau .
\]

(8)

The data fitting process can be achieved through the equation

\[ X \cong \Psi \Phi \] (9)

in which \( \Phi \in \mathbb{R}^{(N+1) \times J} \) is defined as the matrix for the coefficient parameters of the general impulse response function, which is related to the tracer kinetics and the matrix for \( \Phi \) corresponding to \( X \) is hence reformulated as

\[
\Phi = \begin{bmatrix}
\phi_{01} & \phi_{02} & \cdots & \phi_{0J} \\
\vdots & \vdots & \ddots & \vdots \\
\phi_{N1} & \phi_{N2} & \cdots & \phi_{NJ}
\end{bmatrix}.
\]

(10)
Usually, a relatively large $N$ is chosen for the matrix to ensure a good coverage of the kinetic spectrum, which would lead to an over-complete dictionary of $\Psi$. With the prior knowledge about the solution that the activity distribution is only composed by a few compartments, a sparse constraint for the coefficients of $\Phi$ can be added to select the most related subset of the basis functions; the process of fitting the data can therefore transformed into the regression problem.

Based on the compartmental model, given a TAC, only a few elements of the dictionary should be enough to restore the parametric images, which leads to the following sub-objective function

$$S(X, \Phi) = \frac{1}{2} \|X - (\Phi \Psi)^T\|^2 + \mu \|\Phi\|_1$$

(11)

in which $\|\cdot\|_1$ refer to the $\ell_1$ norm and $\mu$ is a weighting parameter for the sparseness. The $\ell_1$ norm can be replaced by the $\ell_0$ norm to constrain the sparsity of $\Phi$. Here, the $\ell_1$ norm is chosen to make the problem easier to solve than the $\ell_0$ norm and achieve basis pursuit denoising. Then, the appropriate macro parameters, for example volume of distribution of the target tissue $V_0 \in \mathbb{R}^J$ can be calculated from $\Phi$ through the affine

$$(V_d)_j = \sum_{c=1}^{N} \frac{\theta_d}{\theta_c}$$

(12)

where $(V_d)_j$ represents the $j$th element in $V_d$, which is related to the $j$th pixel in the activity distribution field.

2.2. Solution

Our aim here is to reconstruct dynamic PET activity map $X$ and kinetic coefficient parameter $\Phi$ simultaneously. Therefore, we propose the objective function in our model by combining equations (1) and (11) as follows.

$$\mathcal{F}(X, \Phi) = TV(X) + \alpha \gamma \mu \Phi^T$$

(13)

The weighting parameter $\gamma$ is incorporated to control the relative ratio for the terms generated from kinetic modelling. $TV(X)$ is the total variation of $X$ and used to provide edge-preserving smoothing. $\|GX - Y\|^2$ and $\|X - (\Phi \Psi)^T\|^2$ are the data fitting terms for $X$ and $\Phi$. The last $\ell_1$ norm for $\Phi$ is the sparsity regularization term.

The model proposed above is then optimized by alternatively minimizing $\mathcal{F}(X, \Phi)$ with respect to $X$ and $\Phi$ as follows.

2.2.1. Activity reconstruction sub-problem. Extracting the components of $\mathcal{F}(X, \Phi)$ that concerns variable $X$ gives

$$\mathcal{F}_X(X, \Phi) = TV(X) + \alpha \gamma \mu \|X - (\Phi \Psi)^T\|^2.$$ 

(14)

Thus, the sub-problem here can be formulated as

$$\min_X \{TV(X) + \alpha \gamma \mu \|X - (\Phi \Psi)^T\|^2\}.$$ 

(15)
By introducing an auxiliary variable $w_{jm}$, problem (15) is equivalent to the following constrained problem

$$\min_{w_{jm}} \{ \sum_{m=1}^{M} \sum_{j=1}^{J} \|w_{jm}\| + \frac{\alpha}{2} (\|GX - Y\|^2 + \gamma\|X - (\Psi\Phi)^T\|^2) \}$$

s.t. $w_{jm} = (DX_m)_j$. \hspace{1cm} (16)

Therefore, the corresponding augmented Lagrangian function for problem (16) is

$$L_{\lambda}(w_{jm}, X) = \sum_{m=1}^{M} \sum_{j=1}^{J} (\|w_{jm}\| - \nu^T_{jm}((DX_m)_j - w_{jm})$$

$$+ \frac{\beta_{jm}}{2} ((DX_m)_j - w_{jm})^2$$

$$+ \frac{\alpha}{2} (\|GX - Y\|^2 + \gamma\|X - (\Psi\Phi)^T\|^2)$$ \hspace{1cm} (17)

in which $\nu_{jm}$ is the vector of the Lagrange multipliers and $\beta_{jm}$ denotes the penalty parameters (for all $j$ and $m$).

An iterative alternating minimization scheme is then applied to solve the problem above. Minimizers $X$ and $w_{jm}$ are achieved alternatively in the proposed algorithm. First, for a fixed $X$, the minimizers $w_{jm}$ can be achieved through

$$\min_{w_{jm}} \{ \sum_{m=1}^{M} \sum_{j=1}^{J} (\|w_{jm}\| - \nu_{jm}^T((DX_m)_j - w_{jm})$$

$$+ \frac{\beta_{jm}}{2} ((DX_m)_j - w_{jm})^2) \}$$ \hspace{1cm} (18)

which is formulated by extracting the relevant components from $L_{\lambda}(w_{jm}, X)$. In $\beta_{jm} > 0$ case, the $k$th updated minimizer is then given by the 2D-shrinkage formula as follows [40]

$$w_{jm}^{k+1} = \max\{ \|(DX_m)_j - w_{jm}^k\| - \frac{1}{\beta_{jm}} \|\nu_{jm}^k\|, 0\} \frac{(DX_m)_j^k - \nu_{jm}^k}{\|DX_m)_j^k\|} . \hspace{1cm} (19)$$

Next, for the fixed $w_{jm}^{k+1}$, noting that the terms $\|GX - Y\|^2$ and $\|X - (\Psi\Phi)^T\|^2$ can be linearized, the update of $X^{k+1}$ is estimated through approximately minimizing the following linearized alternating direction method of multipliers (ADMM) scheme for $X$

$$\min_{X} \{ \sum_{jm} (-\nu_{jm}^k)^T(DX_m)_j$$

$$+ \frac{\beta_{jm}}{2} ((DX_m)_j - w_{jm}^{k+1})^2$$

$$+ \alpha(G^T(GX^k - Y) + \gamma(X^k - (\Psi\Phi)^T), X)$$

$$+ \frac{\eta}{2} \|X - X^k\|^2 \} . \hspace{1cm} (20)$$

In addition, the Lagrange multipliers also need to be updated with $w_{jm}$ and $X$ through

$$\nu_{jm}^{k+1} = \nu_{jm}^k - \frac{\beta_{jm}}{\|DX_m)_j\|} (DX_m)_j - w_{jm}^{k+1}. \hspace{1cm} (21)$$

After each steepest descent step, $w_{jm}$ is updated and the process is repeated until the problem (14) is approximately solved within a prescribed tolerance.
2.2.2. Tracer kinetics sub-problem. Only $S(X, \Phi)$ contains $\Phi$ in objective function (13), so the sub-problem of $\Phi$ is to minimize $S(X, \Phi)$ with respect to $\Phi$ to get an equivalent expression of $S(X, \Phi)$ as

$$S(X, \Phi) = \mu \| \Phi \|_1 + Q(X, \Phi)$$

subject to

$$Q(X, \Phi) = \frac{1}{2} \| X - (\Psi \Phi)^T \|^2. \quad (22)$$

To treat the sub-problem, $Q(X, \Phi)$ can be linearized and the iterative thresholding algorithm [41] is then applied to minimize the sub-problem

$$\Phi^{k+1} = S_{\mu, \sigma}(\Phi^k - \frac{1}{2\sigma} \nabla Q(X^{k+1}, \Phi)) \quad (23)$$

where $\nabla Q(X^{k+1}, \Phi)$ is the derivative of $Q(X^{k+1}, \Phi)$ at $\Phi$, $S_{\mu, \sigma}$ is the soft shrinkage operator, and $\sigma$ is the step size [41].

Algorithm 1. Optimization scheme of the proposed model.

Require: Sinogram $Y$, system matrix $G$, dictionary $\Psi$.

1: Initialize: Input parameters: $\alpha, \beta, \gamma, \mu, \sigma$. Iteration number $k = 1$.

2: repeat

3: $w^{k+1}_{jm} \leftarrow \arg \min L_A(w_{jm}, X^k, v^k_{jm})$

4: $X^{k+1} \leftarrow \arg \min L_A(w^k_{jm}, X, v^k_{jm})$

5: $v^{k+1}_{jm} \leftarrow v^k_{jm} - \beta(DX^{k+1}_m - u^{k+1}_{jm})$

6: $\Phi^{k+1} = S_{\mu, \sigma}(\Phi^k - \frac{1}{\sigma} \nabla Q(X^{k+1}, \Phi))$

7: $k \leftarrow k + 1$

8: until Outer stopping criteria are satisfied.

9: Return $X, \Phi$.

The overall algorithm of the proposed model is summarized in algorithm 1. The algorithm terminates when both the variable achieve global convergence.

3. Experiments

In this section, the proposed algorithm is evaluated through experiments based on the Zubal phantom [42] synthetic data, Monte Carlo simulated data, and real patient cardiac data. Simulation results demonstrate both robustness and accuracy of reconstructed activity and parametric images under different conditions (noise level and counting rate). In addition, the selection of the tuning parameters mentioned in the previous section will also be presented as a form of variable controlling experiment in the following section.

For a part of the activity images, the reconstruction result of the proposal method is compared to the traditional ML-EM method, and quantitative evaluation is presented as a relative bias and relative variance for all the $M$ frames, which are defined as

$$\text{bias} = \frac{1}{M} \sum_{m=1}^{M} \left( \frac{1}{J} \sum_{j=1}^{J} \frac{|\hat{x}_{jm} - x_{jm}|}{x_{jm}} \right)$$

$$\text{variance} = \frac{1}{M} \sum_{m=1}^{M} \left( \frac{1}{J-1} \sum_{j=1}^{J} \frac{(|\hat{x}_{jm} - x_{jm}|^2)}{x_{jm}} \right) \quad (24)$$
where $x_{jm}$ denotes the true value of the $j$th pixel in the $m$th time frame, $\hat{x}_{jm}$ is the reconstructed value for the pixel and $\bar{x}_{jm}$ is defined as the mean value of all the pixels that belong in the same cluster with the pixel. Quantitative results for bias, variance, and the relative MSE (mean...
square error) mentioned below can only be presented for synthetic and Monte Carlo simulated data with ground truth known.

To evaluate the estimation accuracy of parametric images, the result achieved by our joint framework is compared with data-driven estimation of parametric images based on the compartmental theory (DEPICT) method [33]. The input image for DEPICT is generated through the ML-EM method so as to unify the form of input data (sinogram) for our algorithm and the DEPICT method. The parametric image of $V_d$ (volume of distribution) is presented and quantitatively evaluated by relative MSE.

| Noise Level | Method | Bias   | Variance   |
|-------------|--------|--------|------------|
|             |        | Whole  | ROI1 | ROI2 | ROI3 | Whole | ROI1 | ROI2 | ROI3 |
| 0.00        | joint + TV | 0.1169 | 0.1364 | 0.1382 | 0.0778 | 0.0243 | 0.0306 | 0.0316 | 0.0116 |
|             | joint   | 0.1644 | 0.1564 | 0.1659 | 0.1775 | 0.0395 | 0.0422 | 0.0374 | 0.0357 |
|             | ML-EM   | 0.1387 | 0.1530 | 0.2090 | 0.0937 | 0.0330 | 0.0398 | 0.0547 | 0.0151 |
| 0.02        | joint + TV | 0.1313 | 0.1483 | 0.1590 | 0.0943 | 0.0280 | 0.0342 | 0.0397 | 0.0139 |
|             | joint   | 0.1797 | 0.1740 | 0.1761 | 0.1903 | 0.0462 | 0.0502 | 0.0421 | 0.0407 |
|             | ML-EM   | 0.1461 | 0.1616 | 0.2109 | 0.1004 | 0.0354 | 0.0432 | 0.0556 | 0.0162 |
| 0.04        | joint + TV | 0.1374 | 0.1518 | 0.1659 | 0.1047 | 0.0294 | 0.0344 | 0.0435 | 0.0167 |
|             | joint   | 0.1971 | 0.1922 | 0.1885 | 0.2078 | 0.0541 | 0.0587 | 0.0475 | 0.0483 |
|             | ML-EM   | 0.1521 | 0.1672 | 0.2305 | 0.1030 | 0.0376 | 0.0453 | 0.0631 | 0.0170 |
| 0.08        | joint + TV | 0.1467 | 0.1636 | 0.1732 | 0.1102 | 0.0332 | 0.0380 | 0.0463 | 0.0211 |
|             | joint   | 0.2179 | 0.2118 | 0.2076 | 0.2312 | 0.0656 | 0.0708 | 0.0565 | 0.0596 |
|             | ML-EM   | 0.1980 | 0.1690 | 0.2630 | 0.2275 | 0.0592 | 0.0469 | 0.0802 | 0.0735 |

Figure 3. Reconstruction of $V_d$ for synthetic Zubal phantom data. (a) Groundtruth. (b) DEPICT. (c) Joint. (d) Joint + TV.
3.1. Synthetic Zubal phantom data

We first conducted the experiment on synthetic emission Zubal thorax phantom data with known radioactivity concentration as shown in figure 1. A dynamic PET image sequence (18 frames) was generated using the two-compartment model and the resolution for each frame was 128 by 128 pixels. The simulated tracer was $^{18}$F-FDG and kinetic parameters were set as in [43]. Then, the sinogram was generated based on the system matrix, which was calculated based on the tomography geometry, and different levels of random noise were modeled so as to evaluate the robustness of our algorithm.

The reconstructed activity images are partially demonstrated in figure 2, in which ‘joint + TV/joint’ refer to the proposal framework with (‘joint + TV’) and without (‘joint’) TV regularization.

It is clear that images reconstructed by the ML-EM method are blurry, with low contrast; joint has much less noise compared to ML-EM and the updated joint + TV framework achieved the most faithful representation of the ground truth compared to the other methods. Moreover, the improvement in the accuracy of the reconstructed image is clear from the quantitative data shown in tables 1 and 2. From table 2, we can see that the results for joint + TV maintain a promising value when the noise level goes higher.

Figure 3 shows the true value and the reconstructed value of $V_d$. Compared to the ground truth, our results present a better structure and the contrast is higher so the different ROIs can be identified clearly. Table 1 shows the relative MSE comparison between the joint + TV algorithm and DEPICT.

3.2. Monte Carlo simulated data

In this section, we verify our algorithm on Monte Carlo simulated brain and cardiac phantom. The Monte Carlo method can simulate the physical imaging process of PET, thus providing realistic sinogram data, and more importantly, with the known concentration distribution, reconstruction results of Monte Carlo simulated data can be evaluated against a highly reliable criterion.
3.2.1. Brain phantom. Experiments on Monte Carlo simulated brain phantom are first presented. The template of the brain phantom is shown in figure 4, in which a tumor is inserted as ROI3. In this experiment, an image sequence \( (64 \times 64 \times 18) \) based on the original brain phantom is first generated through the two-compartment model. Kinetic parameters \( k_1, k_2, k_3, k_4 \ (\text{min}^{-1}) \) and the temporal distribution of time frames are defined according to [44]. The PET scanner is simulated as real Hamamatsu SHR-22000 and the tracer is set as \(^{18}\text{F}-\text{FDG}.\) The generated measurement data has \( 64 \times 64 \) projections. In addition, the data set is simulated with four different counts \( (5 \times 10^4, 1 \times 10^5, 5 \times 10^5, 1 \times 10^6) \) in order to evaluate the robustness of the proposed algorithm.

Figure 5. Reconstructed activity maps from joint + TV method for Monte Carlo simulated brain phantom data under different counts. The first two rows correspond to the reconstructed images and subtraction images for the 9th frame. The last two rows correspond to the reconstructed images and subtraction images for the 12th frame. (a) \( 5 \times 10^4 \) counts. (b) \( 1 \times 10^5 \) counts. (c) \( 5 \times 10^5 \) counts. (d) \( 1 \times 10^6 \) counts.
### Table 3. Relative bias-variance results of each ROI for Monte Carlo simulated brain phantom under different counts.

| Counts  | Method  | Whole | ROI1  | ROI2  | ROI3  | Whole | ROI1  | ROI2  | ROI3  |
|---------|---------|-------|-------|-------|-------|-------|-------|-------|-------|
| $5 \times 10^4$ | joint + TV | 0.1751 | 0.1292 | 0.1935 | 0.1473 | 0.0464 | 0.0236 | 0.0556 | 0.0337 |
|         | ML-EM   | 0.3992 | 0.2683 | 0.4547 | 0.2462 | 0.1895 | 0.0809 | 0.2352 | 0.0764 |
| $1 \times 10^5$ | joint + TV | 0.1503 | 0.1308 | 0.1580 | 0.1462 | 0.0373 | 0.0228 | 0.0431 | 0.0310 |
|         | ML-EM   | 0.3157 | 0.1890 | 0.3691 | 0.1782 | 0.1344 | 0.0440 | 0.1725 | 0.0388 |
| $5 \times 10^5$ | joint + TV | 0.1462 | 0.1193 | 0.1573 | 0.1231 | 0.0345 | 0.0187 | 0.0412 | 0.0178 |
|         | ML-EM   | 0.2936 | 0.1623 | 0.3487 | 0.1572 | 0.1224 | 0.0325 | 0.1602 | 0.0300 |
| $1 \times 10^6$ | joint + TV | 0.1411 | 0.1104 | 0.1542 | 0.1024 | 0.0326 | 0.0164 | 0.0395 | 0.0132 |
|         | ML-EM   | 0.2817 | 0.1500 | 0.3368 | 0.1485 | 0.1156 | 0.0286 | 0.1519 | 0.0298 |

### Figure 6. Reconstruction of $V_d$ for Monte Carlo simulated brain phantom data under different counts. (a) Groundtruth. (b) $5 \times 10^4$ counts. (c) $1 \times 10^5$ counts. (d) $5 \times 10^5$ counts. (e) $1 \times 10^6$ counts.

### Table 4. RMSE results compared with DEPICT for Monte Carlo simulated brain phantom under different counts.

| Counts  | $5 \times 10^4$ | $1 \times 10^5$ | $5 \times 10^5$ | $1 \times 10^6$ |
|---------|----------------|----------------|----------------|----------------|
| joint + TV | MSE   | 0.0512 | 0.0498 | 0.0361 | 0.0319 |
| joint     | MSE   | 0.0546 | 0.0539 | 0.0436 | 0.0392 |
| DEPICT    | MSE   | 0.0807 | 0.0701 | 0.0442 | 0.0366 |

### Figure 7. Template of cardiac phantom with ROI information.
Figure 5 shows the 9th and 12th frames of the reconstructed image sequence of four different counts. As the counts become higher, the noise is less involved and the reconstructed images preserve the structure of the ground truth better. To be more specific, the quantitative results for relative bias and variance exhibiting significant improvement compared to the ML-EM method are shown in table 3.

As for the parametric images, reconstructed $V_d$ images are shown in figure 6. The inserted tumor is visually hard to identify even with counts of $1 \times 10^6$. The improvement for $V_d$ in higher count cases is quite obvious compared to that in lower cases, whereas they all present a promising result against the other reconstruction algorithms shown in table 4.

3.2.2. Cardiac phantom. Next, we evaluate the proposed framework on a Monte Carlo simulated cardiac phantom. The template of the phantom is shown in figure 7. Since its ROI structure is simple, this time we mainly focus on the tunable parameters. The dynamic image sequence is also generated using the two-compartment model and the relative parameters (kinetic parameters and template distributions) are fixed as described in [45]. As for Monte Carlo simulations, the PET scanner and the tracer are the same as in the previous section. The total number of coincidence events is set as $1 \times 10^6$.

The quantitative results are shown in table 5, and it is clear that our algorithm can provide more accurate results both in activity and parametric images.

3.3. Real patient data

The real patient data presented in this section is a sequence of a dynamic PET scan offered by a volunteer from a local hospital. The scanner is a Hamamatsu SHE-22000 whole body PET scanner with 32 crystal rings, and can be operated in 2D or 3D modes. The transaxial

|                | Bias  | Variance | RMSE  |
|----------------|-------|----------|-------|
| joint + TV     | 0.1432| 0.0354   | 0.0564|
| ML-EM          | 0.2787| 0.1431   | —     |
| DEPICT         | —     | —        | 0.7929|

Table 5. Relative bias, variance (for activity map $X$), and MSE (for $V_d$) results compared with ML-EM and DEPICT for Monte Carlo simulated cardiac phantom data.

Figure 8. Reconstructed image results for real patient data. The first two are the results of reconstructed activity distribution through different algorithm. The last two are the results of parametric images of $V_d$ for different algorithms. (a) 4th: joint + TV. (b) 4th: ML-EM. (c) $V_d$: joint + TV. (d) $V_d$: DEPICT.
resolution of the central FOV is about 3.7 mm. $^{18}$F-FDG is injected as the tracer and a dynamic acquisition sequence starts just after injection. The acquisition sequence consists of 19 frames for 60 min. As for the dictionary and other tunable parameters: these stay the same as in the previous section, since they both use $^{18}$F-FDG as injected tracer.

The reconstruction results for both activity and parametric images are presented together in figure 8. A bright ring considered as the myocardium (like that in the Monte Carlo simulated cardiac phantom) can be clearly identified from all these images. However, as there is no ground truth to compare, we can only say that the reconstruction results are visually close to that of the ML-EM method (as in the activity map case) and the DEPICT method (as in the $V_d$ case). Furthermore, it is quite clear that the contrast between myocardium and other tissues is more explicit in the joint + TV case.

### 3.4. Tunable parameters

Four parameters need to be manually tuned to achieve a promising estimation. The parameter $\alpha$, which controls the relative weight of the TV regularization, and the penalty parameter $\beta$ in the optimizing scheme are preferred to be in the range of $[2^4, 2^{13}]$ as mentioned in [40]. In this paper, we empirically find that the algorithm performs well when $\alpha$ is set at about $2^8$ and $\beta$ at around $2^7$.

---

**Figure 9.** RMSE, relative bias, and variance results for different ranges of parameters. (a) RMSE for $\gamma$ range. (b) Bias and variance for $\gamma$ range. (c) RMSE for $\mu$ range. (d) Bias and variance for $\mu$ range.
To determine the ranges of the other two tunable parameters $\gamma$, $\mu$ and examine their robustness within the determined range, relative bias, variance, and MSE are calculated based on different values of these parameters. In our experiment, the ranges for $\gamma$ and $\mu$ are pre-determined as [0.5, 3.5] and [0.005, 0.17] respectively. When one of the parameters is tested, the other is fixed at a suitable value so that relatively promising results can be achieved. In practice, the fixed value for $\gamma$ is 1.5, and $\mu$ is set to 0.01.

As shown in figure 9, the values of relative bias and variance stay in stable and proper regions across the whole testing range for all these parameters, while the relative MSE is more sensitive to these parameters. The lowest point for relative MSE can be reached when $\gamma$ and $\mu$ are selected with the fixed values mentioned before.

Beyond that, the stop criteria for the algorithm are satisfied when it reaches the max iteration time, which is fixed at 300 in the experiments, or the changing ratio for both $X$ and $\Phi$ is lower than $10^{-3}$.

4. Discussion

In this paper, we have proposed a TV constrained dictionary sparse coding framework that allows us to reconstruct the activity concentration and parametric images simultaneously. Combining TV regularization and biological based regularization (two-compartment model), the reconstruction process can be viewed as an optimization problem solved through an iterative process, and the generated images can be both statistically and biologically meaningful.

In our experiments, the reconstruction algorithm is verified on synthetic Zubal thorax phantom data, Monte Carlo simulated brain and cardiac data, and real patient data. For comparisons, a TV eliminated version of the proposal framework, i.e. reconstruction algorithm only based on DSC, as well as the ML-EM method and DEPICT are also presented. First, we test the robustness to noise on the synthetic Zubal phantom data and the joint + TV method presents a much better image of activity distribution under all noise levels. As for $V_d$, the additional TV regularization makes the whole system less sensitive to noise. Next, in Monte Carlo simulations, we present the results on different counts (brain phantom) to again evaluate the robustness of the algorithm and examine the proper range for tunable parameters ($\gamma$, $\mu$). The results demonstrated in the experiments section show that as the counts get lower, TV regularization keeps the error of reconstructed images in a low and stable range compared to other methods. The Monte Carlo simulated cardiac phantom, with simple structure and truly reflecting the imaging physical process of PET, is set as the training data for determining the tunable parameters. As shown in figure 9, the planned $\gamma$ and $\mu$ did not achieve the optimal results for both activity and parametric images. However, by balancing the error on both categories of the results, it is believed that they have reached a relatively suitable value for the whole framework. For the final experiments, in which the data comes from a real patient in a local hospital, we demonstrate the results for a cardiac case; the reconstruction images preserve the structure of both the activity and parametric image well, as compared to the ML-EM method and DEPICT method.

Acknowledgments

This work is partially supported by the National Natural Science Foundation of China (No: 61427807,61271083, 61525106), by National Key Research and Development Program of China (No: 2016YFC1300302), and by Shenzhen Innovation Funding (SGLH2013101011019871 and GJHZ20140415152115754). In addition, it is supported by the NSF grant DMS-131905 in part.
References

[1] Gravel P and Reader A J 2015 Direct 4D PET (MLEM) reconstruction of parametric images using the simplified reference tissue model with the basis function method for \([1^1C]\) raclopride Phys. Med. Biol. 60 R4533

[2] Rahim M, Tang J and Zaidi H 2009 Four-dimensional (4D) image reconstruction strategies in dynamic PET: beyond conventional independent frame reconstruction Med. Phys. 36 3654–70

[3] Qi J and Leahy R M 2006 Iterative reconstruction techniques in emission computed tomography Phys. Med. Biol. 51 R541

[4] Chinn G and Huang S C 1997 A general class of preconditioners for statistical iterative reconstruction of emission computed tomography IEEE Trans. Med. Imaging 16 1–10

[5] Levitan E and Herman G T 1987 A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography IEEE Trans. Med. Imaging 6 185–92

[6] Qi J and Leahy R M 2000 Resolution and noise properties of MAP reconstruction for fully 3D PET IEEE Trans. Med. Imaging 19 493–506

[7] Nyúts J, Bequère D, Dupont P and Mortelmans L 2002 A concave prior penalizing relative differences for maximum-a posteriori reconstruction in emission tomography IEEE Trans. Nucl. Sci. 49 56–60

[8] Locke L W, Berr S S and Kundu B K 2011 Image-derived input function from cardiac gated maximum a posteriori reconstructed PET images in mice Mol. Imaging Biol. 13 342–7

[9] Bai B, Lin Y, Zhu W, Ren R, Li Q, Dahlbom M, DiFilippo F and Leahy R M 2014 MAP reconstruction for fourier rebinned TOF-PET data Phys. Med. Biol. 59 925

[10] Lewitt R M and Muehllehner G 1986 Accelerated iterative reconstruction for positron emission tomography based on the EM algorithm for maximum likelihood estimation IEEE Trans. Med. Imaging 5 16–22

[11] Hudson J I, Goldenberg D L, Pope H G, Keck P E and Schlesinger L 1992 Comorbidity of fibromyalgia with medical and psychiatric disorders Am. J. Med. 92 363–7

[12] Gregor J and Huff D A 1997 A focus-of-attention preprocessing scheme for EM-ML PET reconstruction IEEE Trans. Med. Imaging 16 218–23

[13] Chapeau-Blondeau F and Jeanguillaume C 2014 Characterization of maximum likelihood solutions to image reconstruction in photon emission tomography J. Math. Imaging Vision 49 467–80

[14] Beister M, Kolditz D and Kalender W A 2012 Iterative reconstruction methods in x-ray CT Phys. Med.: Eur. J. Med. Phys. 28 94–108

[15] Das M, Gifford H C, Michael O’Connor J and Glick S J 2011 Penalized maximum likelihood reconstruction for improved microcalcification detection in breast tomosynthesis IEEE Trans. Med. Imaging 30 904–14

[16] Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography IEEE Trans. Med. Imaging 1 113–22

[17] Burger M and Lucka F 2014 Maximum-a posteriori estimates in linear inverse problems with concave priors are proper bayes estimators Inverse Problems 30 114004

[18] Calvetti D, Kairem J P and Somersalo E 2014 Inverse problems in the bayesian framework Inverse Problems 30 110301

[19] Persson M, Bone D and Elmqvist H 2001 Three-dimensional total variation norm for spect reconstruction Nucl. Instrum. Methods Phys. Res. A 471 98–102

[20] Burger M, Muller J, Papoutsellis E and Schonlieb C 2014 Total variation regularisation in measurement and image space for pet reconstruction Inverse Problems 30 105003

[21] Bardsley J M, Calvetti D and Somersalo E 2010 Hierarchical regularisation for edge-preserving reconstruction of pet images Inverse Problems 26 035010

[22] Reader A J, Sureau F C, Comtat C, Trebossen R and Buvat I 2006 Joint estimation of dynamic pet images and temporal basis functions using fully 4D ml-em Phys. Med. Biol. 51 5455

[23] Chen S, Liu H, Hu Z, Zhang H, Shi P and Chen Y 2015 Simultaneous reconstruction and segmentation of dynamic pet via low-rank and sparse matrix decomposition IEEE Trans. Biomed. Eng. 62 1784–95

[24] O’Sullivan F and Saha A 1999 Use of ridge regression for improved estimation of kinetic constants from PET data IEEE Trans. Med. Imaging 18 115–25

[25] Wang G and Qi J 2012 An optimization transfer algorithm for nonlinear parametric image reconstruction from dynamic PET data IEEE Trans. Med. Imaging 31 1977–88
