Simultaneous Reconstruction and Segmentation of Dynamic PET via Low-Rank and Sparse Matrix Decomposition

Shuhang Chen, Huafeng Liu*, *Member, IEEE*, Zhenghui Hu, Heye Zhang*, Pengcheng Shi, and Yunmei Chen

***Abstract*—Although of great clinical value, accurate and robust reconstruction and segmentation of dynamic positron emission tomography (PET) images are great challenges due to low spatial resolution and high noise. In this paper, we propose a unified framework that exploits temporal correlations and variations within image sequences based on low-rank and sparse matrix decomposition. Thus, the two separate inverse problems, PET image reconstruction and segmentation, are accomplished in a simultaneous fashion. Considering low signal to noise ratio and piece-wise constant assumption of PET images, we also propose to regularize low-rank and sparse matrices with vectorial total variation norm. The resulting optimization problem is solved by augmented Lagrangian multiplier method with variable splitting. The effectiveness of proposed approach is validated on realistic Monte Carlo simulation datasets and the real patient data.**

***Index Terms*—Augmented Lagrangian multiplier, convex optimization, dynamic PET reconstruction, low-rank/sparse decomposition, poisson likelihood function, segmentation.**

I. Introduction

With the increasingly wider availability of radiotracers, dynamic positron emission tomography (PET) imaging enables to provide quantitative and noninvasive information of different biologic and physiologic processes through reconstructing the spatiotemporal distribution of radiopharmaceutical (tracer) labeled biological substrates in living tissue. In practice, dynamic PET images are often segmented into regions of interest (ROI) to obtain time activity curves (TAC) which can be further analyzed to estimate physiological parameters, such as blood flow, metabolism, and receptor concentration depending on the characteristics of the tracer. However, even with the recently developed technology PET/computed tomography (CT) or PET/magnetic resonance imaging (MRI), segmenting PET images with the aid of anatomical information still has many challenges. This is mainly because, in many studies, the findings of anatomical and functional imaging modalities have disagreed, and in some circumstances anatomical imaging modalities (i.e., CT or MRI) might miss some regions that appear suspicious on PET images. Therefore, reliable and automated detection and characterization of the tissue elements from dynamic PET images have been so far proven technically difficult because PET has relatively low spatial resolution, high noise, and lacks of anatomical information.

Conventionally, reconstruction and segmentation are viewed as two separate steps. Segmentation requires the high quality reconstruction algorithm as an essentially introductory step. In the past decade, there were abundant efforts in medical imaging community devoted to the reconstruction for PET. Prior to dynamic PET, various algorithms have been developed for reconstructing static PET images, including analytic filtered backprojection (FBP) [1], maximum likelihood expectation maximization estimate (ML-EM) [2], maximum *a posteriori* (MAP) [3], state space analysis [4]. Especially for MAP reconstruction, designing penalty terms that preserve regional smoothness and sharp variations of edges has been the focus of researches on PET reconstruction [5]. These investigated penalty terms include Gibbs priors, logcosh, total variation, Huber potential [6], [7], [8]. Some researchers also incorporated anatomical priors from anatomy modalities, like CT and MRI, to help reconstruct PET images [9]. For dynamic PET imaging, temporal priors including spatiotemporal spline modeling [10], [11], [12], tracer kinetics [13], [14], signal subspaces [15], wavelets [16] are also exploited extensively in the literature.

Although of great progresses achieved, all aforementioned reconstruction techniques incorporate statistical assumptions on the measurement distribution (i.e., Poisson, shifted Poisson [17]). In general, there is a tradeoff in spatial and temporal resolution, i.e., if the number of photon counts collected for each time frame is large, then the number of time frames that can be collected in a unit of time is small, and *vice versa*. In order to permit a reasonable temporal resolution, the signal to noise ratio (SNR) of the measured dynamic PET data is relatively low. Such noisy images make the following segmentation a very challenging issue.

In the segmentation efforts, many strategies from computer vision community has been applied to PET images, notably the geometric level set methods [18], [19], [20], graph-cut-constrained framework [21], [22], [23]. These models, in

Manuscript received October 29, 2014; revised January 2, 2015; accepted February 2, 2015. Date of publication February 18, 2015; date of current version June 16, 2015. This work was supported in part by the National Natural Science Foundation of China (61427807, 61271083), Shenzhen Innovation Funding (SGLH20131010110119871 and GJHZ20140415152115754), Natural Science Foundation of Zhejiang (LR12F03001), NSF Grants IIP-1237814 and DMS-1319050. *Asterisks indicate corresponding author.*

*H. Liu is with the State Key Laboratory of Modern Optical Instrumentation, Department of Optical Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: liuhf@zju.edu.cn).

S. Chen and Z. Hu are with the Zhejiang University.

*H. Zhang is with the Shenzhen Institute of Advanced Technology, Shen-zhen, China (e-mail: hy.zhang@siat.ac.cn).

P. Shi is with the Rochester Institute of Technology.

Y. Chen is with the University of Florida.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2015.2404296

general, are based on piecewise smooth assumptions and often need manual initialization. Besides, many machine learning-based methods are also investigated like active appearance models [24] and spectral clustering [25], [26], [27]. The active appearance model learns shape priors from existing database, and spectral method takes TAC as a feature point for each voxel and apply spectral clustering to segment the image into ROIs that group voxels with a similar TAC.

Unlike other state-of-the-art methods which focus on either activity map reconstruction or segmentation as two sequential processes, our approach treats the spatiotemporal reconstruction and spatial boundary finding problem as a coherent and unified process. Since accurate segmentation requires knowledge and modeling of the noise distribution in images, which is extremely difficult for dynamic PET because the noise is spatially variant and correlated between voxels in image domains. If these two problems are tackled in a joint or simultaneous fashion, raw sinogram data can be well modeled by Poisson statistics while segmentation also enforce homogeneity for each region. Thus, more consistent and probably more appropriate results can be expected.

In this paper, we propose to simultaneously reconstruct and segment dynamic PET image sequence. Dealing directly with sinogram data, we are able to accurately model the noise statistics by Poisson likelihood function. Through low-rank and sparse decomposition of image sequence, the segmented tissues are extracted in the sparse matrix while the low-rank constraint also eliminates the noise by intrinsic averaging within tissues. Considering the low SNR issue of PET measurement data, we also incorporate vectorial total variation (VTV) norm of both low-rank and sparse matrices into the model. An iterative procedure is then provided for optimizing the resulting objective function. The proposed model is validated with realistic Monte Carlo simulated data and real patient data.

A. Related Work

We realized that some of the recent research efforts upon simultaneous reconstruction and segmentation may be relevant to our study. Kamasak proposed to segment dynamic PET image directly from sinogram data [31], [32]. Given the number of ROIs, their model alternates between reconstructing activity maps and clustering TACs. In [33], a level-set-based method for simultaneously reconstructing and segmenting CT image was proposed, which assumed that intensity value within each region was bounded. Since stronger smooth and boundary priors can be incorporated into reconstruction after segmentation, the reported model is able to produce images with higher quality. However, first of all, how to perform the initialization curve for each region remains a challenging issue. Second, the assumption of bounded regions is prone to fail in low SNR cases.

As for low-rank and sparse decomposition, Gao employed such paradigm to reconstruct 4-D-CT images with undersampled cardiac data in [34], where the sparse term was further sparsified by framelet transform to extract the moving objects. Moreover, the measured sinogram of CT were assumed to have i.i.d Gaussian noise, which is far from realistic in practice. In [35], similar strategy was applied to reconstruct undersampled dynamic MRI imaging, where framelet transform was replaced by Fourier transform. While it turns out that the stationary background is also sparse in the framelet/Fourier domain. This violates the incoherence condition in robust principle component analysis (PCA), which requires the low-rank matrix to be non-sparse [36]. Since when both matrices are sparse, the identifiability of the proposed method is questionable. As a result, the sparse components in [34], [35] contain both the stationary background and moving objects. More importantly, these drawbacks may heavily hamper their applications to handle the complicated nature of PET images.

B. Contribution of This Paper

The contribution of this paper may be summarized as follows:

1) *Low-Rank and Sparse Modeling for Dynamic PET*: Despite the similar form of objective function, we have a completely different formulation for decomposing dynamic PET. Based on the metabolism imaging nature of dynamic PET, in the first place, we model the image sequence by low-rank matrix given the fact that TACs within each tissue should be approximately identical. We then take one more step forward to low-rank/sparse decomposition by noting that the rank of image sequence can be further minimized by extracting the tissue having different metabolism rates relative to the background. Thus, in this way, more appropriate segmentation can be obtained for PET without framelet or Fourier transform.
2) *Poisson Likelihood and VTV Regularization:* We model sinogram with Poisson statistics, not the straightforward Gaussian as previous study did [34]. Hence, more accurate decomposition and reconstruction can be expected. Given the intensity-varying nature of dynamic PET, we incorporate vectorial TV norm to regularize the low-rank and sparse components in both temporal and spatial domains [37]. Because vectorial TV norm is able to combine information from other frames, more consistent image structure could be preserved by our model.

II. Method

In this section, we present the framework of our low-rank and sparse decomposition model. We first recap the imaging model and Poisson likelihood function of PET measurement. Then we elaborate the low-rank/sparse decomposition of dynamic PET image sequence and introduce the VTV norm. Finally, an iterative algorithm is provided to minimize the resulting objective function.

A. Dynamic PET Imaging Model

Dynamic PET imaging involves a sequence of contiguous acquisitions with certain frame rate in order to capture the temporal information of radio-tracer. Within each frame, the detected events can be assumed to follow the Poisson statistics assumption. For each individual frame, sinogram $\mathbf{y}$ denote the sum of detected coincidence events at each detector bin within the time period, i.e., $\mathbf{y} = \{y_i, i = 1, \ldots, I\}$, where I is the number of detector bins. Assuming $\mathbf{y}$ as a collection of independent Poisson

random variables, the relationship between measurement $\mathbf{y}$ and the unknown activity map $\mathbf{x}$ is as follows:

$$\mathbf{y} \sim \text{Poisson}\{\overline{\mathbf{y}}\}, \quad \overline{\mathbf{y}} = G\mathbf{x} + \mathbf{r} + \mathbf{s} \tag{1}$$

where $x \in \mathbb{R}^J$ is the unknown vectorized activity map and J is the number of voxels, G is the system matrix with g_{ij} representing the probabilities of detecting an emission from voxel site j at detector i. It is mainly determined by physical property of PET scanner and the effects of attenuation by the patient's body. $\mathbf{r}$ and $\mathbf{s}$ denote the random coincidence and scattered coincidence events, respectively. Based on the independent Poisson assumption, we can write the likelihood function of $\mathbf{y}$ by

$$\Pr(\mathbf{y}|\mathbf{x}) = \prod_i^I e^{-\overline{y}_i} \frac{\overline{y}_i^{y_i}}{y_i!}, \quad \overline{\mathbf{y}} = G\mathbf{x} + \mathbf{r} + \mathbf{s}. \tag{2}$$

Instead of maximizing above likelihood function, we often minimize the negative log-likelihood function of (2) since it is often easier to deal with.

$$\min_{\mathbf{x}} \Psi(\mathbf{y}|\mathbf{x}) = -\log(\Pr(\mathbf{y}|\mathbf{x})) = \sum_i^I \overline{y}_i - y_i \log(\overline{y}_i). \tag{3}$$

(3) constitutes the data fidelity term of many single frame PET reconstruction algorithms.

For dynamic PET, we can combine multiple frame sinogram into sinogram matrix by taking the mth sinogram vector $\mathbf{y_m} \in \mathbb{R}^I$ as the mth column of matrix Y.

$$Y = [\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_m}, \ldots, \mathbf{y_M}], \quad Y \in \mathbb{R}^{I \times M}. \tag{4}$$

M is overall number of frames. Since the unknown activity map $\mathbf{x} \in \mathbb{R}^J$ for each frame is also vector, the same procedure can be applied to get the image matrix X.

$$X = [\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_m}, \ldots, \mathbf{x_M}], \quad X \in \mathbb{R}^{J \times M}. \tag{5}$$

With these definitions, we obtain $\Psi(Y|X)$ by summing up the likelihood function (3) for each frames.

$$\min_X \Psi(Y|X) = \sum_{m=1}^M \sum_{i=1}^I \overline{y}_{im} - y_{im} \log \overline{y}_{im} \tag{6}$$

where y_{im} is the ith bin value of the mth frame sinogram $\mathbf{y_m}$, and $\overline{y}_{im}$ is the ith entry of mth frame $\overline{\mathbf{y}}_m$.

B. Low-Rank and Sparse Decomposition

Given the definition of $X = [\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_m}, \ldots, \mathbf{x_M}]$, its jth row is the TAC of the jth voxel. In many situations, it is fair to assume that PET images are locally piecewise constant. Each piece corresponds to a certain type of tissue that the TACs of voxels within this tissue are almost identical. In other words, row vectors that belong to voxels from the same tissue should be linear dependent. Hence we can reconstruct dynamic PET sequence by enforcing X being low-rank to eliminate random noise of TACs. However, it is interesting that if we take one more step forward by further minimizing the rank of X, we are able to segment the tissues in dynamic PET image sequence.

To be more specific, the intuition behind the decomposition is that it segments the tissues according to the magnitude of TAC into background components (L) and foreground components (S). The L components have average TAC so that the tissue TACs within L are similar to one another but are dissimilar to those drawn from the S components. Therefore, the image matrix can be decomposed as X = L + S. Moreover, matrix L should be low-rank since it merely contains the average TACs for each voxel. The heterogeneous tissues that have different metabolism rates should be sparse relative to background. This leads to the following decomposition model:

$$\min_{L,S} \text{rank}(L) + \lambda \|S\|_0 + \mu \Psi(Y|X) \quad \text{s.t.} \quad L + S = X \tag{7}$$

where L is the background with shared metabolism rates for each pixel, S is composed of the heterogeneous tissues, $\|S\|_0$ denotes the number of nonzero entries of S, $\Psi(Y|X)$ is the negative log-likelihood function in (6) and μ is weighting parameter. Unfortunately, this is a highly nonconvex problem and no efficient solution is known. Therefore, we resort to solve its convex surrogate function [36]

$$\min_{L,S} \|L\|_* + \lambda \|S\|_1 + \mu \Psi(Y|X) \quad \text{s.t.} \quad L + S = X \tag{8}$$

where $\|L\|_* := \sum_j \sigma_j(L)$ denotes the nuclear norm of matrix L, i.e., the sum of singular values of L, and $\|S\|_1 = \sum_{jm} |S_{jm}|$ denotes the ℓ_1-norm of S seen as a long vector in $\mathbb{R}^{J \cdot M}$, λ is weighting parameter of sparsity. The optimum solution of (8) would give us the low-rank induced background L and the sparse matrix S that consists of the heterogeneous tissues that vary with respect to the background.

C. VTV Norm

The above low-rank/sparse decomposition model in Section II-B implicitly assumes the image matrix X, reconstructed by likelihood function $\Psi(Y|X)$, is free of noise for decomposition, which is not the case in practice. In fact, such maximum likelihood based estimate of X are ill-conditioning and noisy, i.e., the "checkerboard" effect known for ML-EM [2]. Based on our heuristic experiments, both L and S components suffer from noises.

Dynamic PET are used to capture temporal information of the biological processes in living tissues, i.e., the intensity variations for each pixel. Therefore, frames in the dynamic image sequence share similar image structures and voxels in neighboring frames have close values. Many spatial smoothing techniques like Huber function and total variation norms have been proven effective for static PET reconstruction. For kinetic parameter estimation in dynamic PET, several spatiotemporal smoothing models have also been proposed based on wavelet and spline [10]. Here we adopt the simple and efficient VTV norm. VTV norm is an extension of scalar TV to vector-valued image [38]. It keeps the piecewise smoothing property of scalar TV norm but also introduces coupling between frames that interpolates and smooths data in different frames throughout the image sequence. Therefore, we propose to regularize both low-rank matrix L and sparse matrix S with VTV norms [37].

Let $U = [\mathbf{u_1}, \ldots, \mathbf{u_m}, \ldots, \mathbf{u_M}]$ denote a M-frames image sequence, where $\mathbf{u_m} \in \mathbb{R}^J$ is the mth frame image vector. The definition of VTV norm is as follows [39]:

$$R_{\mathbf{vtv}}(\overline{\nabla}U)=\sum_{j}^{J}\sqrt{\sum_{m}^{M}\|(\overline{\nabla}\mathbf{u_m})_j\|_2^2} \tag{9}$$

<?vsp -3pt?>where $\overline{\nabla}$ is the gradient operator on vectorized image $\mathbf{u_m}$, and $(\overline{\nabla}\mathbf{u_m})_j$ denotes the gradient vector at jth pixel of the mth frame. Incorporating VTV norms of L and S into (8), we propose to minimize the following objective function.

$$\begin{aligned}\min_{L,S,X}\ & \|L\|_*+\lambda\|S\|_1+\mu\Psi(Y|X)+\nu_L R_{\mathbf{vtv}}(\overline{\nabla}L)\\ &+\nu_S R_{\mathbf{vtv}}(\overline{\nabla}S)\\ \text{s.t.}\quad & L+S=X\end{aligned} \tag{10}$$

where ν_L,ν_S are weighting parameters of VTV norms.

D. Optimization Algorithm

To optimize problem (10), we first introduce two auxiliary variables U,Q to split the minimization problem (10) into the equivalent constrained optimization problem

$$\begin{aligned}\min_{L,S,X}\ & \|L\|_*+\lambda\|S\|_1+\mu\Psi(Y|X)+\nu_L R_{\mathbf{vtv}}(\overline{\nabla}U)\\ &+\nu_S R_{\mathbf{vtv}}(\overline{\nabla}Q)\\ \text{s.t.}\quad & L+S=X,\quad L=U,\quad S=Q.\end{aligned} \tag{11}$$

We then define the augmented Lagrangian function for (11) [42]

$$\begin{aligned}\mathcal{L}(L,S,U,Q,X)=\ & \|L\|_*+\lambda\|S\|_1-\langle Z,X-(L+S)\rangle\\ &+\frac{\beta}{2}\|X-(L+S)\|_F^2+\mu\Psi(Y|X)\\ &+\nu_L R_{\mathbf{vtv}}(\overline{\nabla}U)-\langle Z_L,L-U\rangle\\ &+\frac{\beta_L}{2}\|L-U\|_F^2+\nu_S R_{\mathbf{vtv}}(\overline{\nabla}Q)\\ &-\langle Z_S,S-Q\rangle+\frac{\beta_S}{2}\|S-Q\|_F^2\end{aligned} \tag{12}$$

where Z,Z_L,Z_S are Lagrangian multipliers and β,β_L,β_S are penalty parameters. The definition of matrix inner product and Frobenius norm can be found in [40]. Because it is difficult to solve problem (12) exactly, we divide it into five subproblems, each of which minimizes the above objective function w.r.t. each variable with others fixed. Some of them have exact one-step solution, some do not. Based on the characteristic of each subproblem, we group them into three categories including:

(1) L,Ssubproblems
(2) X subproblem
(3) U,Q subproblems

and derive the solution for each of them in the following sections.

1) L,S Subproblem: Given other variables fixed, we have the following L subproblem:

$$\begin{aligned}\min_{L}\|L\|_*+\frac{\beta}{2}\|X-(L+S)-Z/\beta\|_F^2+\frac{\beta_L}{2}\|L\\ -U-Z_L/\beta_L\|_F^2.\end{aligned} \tag{13}$$

By completing the squares, the minimization problem (13) is equivalent to

$$\begin{aligned}\min_{L}\|L\|_*+\left(\frac{\beta+\beta_L}{2}\right)\|L\\ -\frac{\beta\left(X-S-\frac{Z}{\beta}\right)+\beta_L\left(U+\frac{Z_L}{\beta_L}\right)}{\beta+\beta_L}\|_F^2.\end{aligned} \tag{14}$$

This form can be solved directly with singular value thresholding (SVT) [43]. Providing soft shrinkage defined as $\mathcal{S}_\delta(x)=\operatorname{sgn}(x)\max(|x|-\delta,0)$, SVT is $\mathcal{D}_\delta(\Omega)=U\mathcal{S}_\delta(\Sigma)V^T$, where $U\Sigma V^T=\Omega$ is singular value decomposition of Ω. The solution of (14) is

$$\begin{aligned}L^{k+1}=\ & \mathcal{D}_{1/(\beta+\beta_L)}(\Omega_L)\\ \text{s.t.}\quad \Omega_L=\ & \frac{\beta(X-S-Z/\beta)+\beta_L(U+Z_L/\beta_L)}{\beta+\beta_L}.\end{aligned} \tag{15}$$

S subproblem has a similar form

$$\begin{aligned}\min_{S}\lambda\|S\|_1+\ & \frac{\beta}{2}\|S-(X-L-Z/\beta)\|_F^2\\ &+\frac{\beta_S}{2}\|S-(Q+Z_S/\beta_S)\|_F^2\end{aligned} \tag{16}$$

whose solution can be directly obtained by soft shrinkage [36]

$$\begin{aligned}S^{k+1}=\ & \mathcal{S}_{\lambda/(\beta+\beta_S)}(\Omega_S)\\ \text{s.t.}\quad \Omega_S=\ & \frac{\beta(X-L-Z/\beta)+\beta_S(Q+Z_S/\beta_S)}{\beta+\beta_S}.\end{aligned} \tag{17}$$

2) X Subproblem: To compute X^{k+1}, we first rewrite the negative log-likelihood function $\Psi(Y|X)$ with respect to hidden variable $\omega=\left\{\left\{\{\omega_{ij}^m\}_j^J\right\}_i^I\right\}_m^M$, where ω_{ij}^m denotes the emissions from voxel j detected at LOR i in the mth frame [2].

$$\Psi(\omega|X)=\sum_{m=1}^{M}\left(\sum_{j=1}^{J}\left[\sum_{i=1}^{I}\left(g_{ij}x_{jm}-\omega_{ij}^m\log(g_{ij}x_{jm})\right)\right]\right). \tag{18}$$

Replacing the likelihood function $\Psi(Y|X)$ in (12) with $\Psi(\omega|X)$, we find that X subproblem consists of minimizing a combination of $\Psi(\omega|X)$ and a quadratic term. Such form arises in PET reconstruction community as MAP reconstruction [3]. It does not has an exact solution, and we use expectation maximization (EM)-based algorithm to iteratively optimize it. The two steps of our EM-based algorithm are:

1) *E-Step*: Take conditional expectation of ω, $\hat{\omega}=E(\omega|Y;X^k)$, and plug it into function $\Psi(\omega|X)$. Then we have the following surrogate function $\phi(X;X^k)$ of X subproblem:

$$\begin{aligned}\phi(X;X^k)=\ & \Psi(\hat{\omega}|X)\ =\mu\sum_{m=1}^{M}\left(\sum_{j=1}^{J}\left[\sum_{i=1}^{I}(g_{ij}x_{jm}\right.\right.\\ &\left.\left.-\hat{\omega}_{ij}^m\log(g_{ij}x_{jm}))\right]\right)\\ &+\frac{\beta}{2}\|X-(L+S)-Z/\beta\|_F^2\\ \text{s.t.}\quad \hat{\omega}_{ij}^m=\ & \frac{g_{ij}y_{im}}{\left(\sum_{l=1}^{J}g_{il}x_{lm}^{(k)}+r_{im}+s_{im}\right)}x_{jm}^k.\end{aligned} \tag{19}$$

2) *M-Step*: Minimize $\phi(X;X^k)$ by zeroing its derivative with respect to x_{jm}

$$\frac{\partial\phi(X;X^k)}{\partial x_{jm}} = 0 = \mu\sum_{i=1}^{I} g_{ij} - \frac{\mu}{x_{jm}}\sum_{i=1}^{I}\hat{\omega}_{ij}^{m} + \beta x_{jm} - \beta[L+S+Z/\beta]_{jm} \tag{20}$$

where $[\cdot]_{jm}$ denote the jmth entry of the matrix. The solution of x_{jm} is then the root of the following second order polynomial equation:

$$\beta x_{jm}^2 + (\mu\sum_{i=1}^{I} g_{ij} - \beta[L+S+Z/\beta]_{jm})x_{jm} - \mu\sum_{i=1}^{I}\hat{\omega}_{ij}^{m} = 0. \tag{21}$$

Noting that objective function in (19) is strictly convex, whose minimum upon the positive cone is unique, we update x_{jm} by choosing the larger root of (21) [3]

$$x_{jm}^{k+1} = \frac{-b_{jm}+\sqrt{b_{jm}^2-4a_{jm}c_{jm}}}{2a_{jm}}$$

$$\text{s.t.} \quad a_{jm} = \beta, b_{jm} = \left(\mu\sum_{i=1}^{I} g_{ij} - \beta[L+S+Z/\beta]_{jm}\right),$$

$$c_{jm} = -\mu\sum_{i=1}^{I}\hat{\omega}_{ij}^{m}. \tag{22}$$

The EM algorithm is guaranteed to converge monotonically to the minimum. Hence, we optimize X subproblem by alternatively taking E-step and M-step until inner convergence. In fact, the augmented Lagrangian function (12) is expected to be minimized by solving each subproblem alternately. Therefore, such inner converged solution at each iteration is sufficient for achieving the minimum of augmented Lagrangian function in (12).

3) U,Q Subproblem: Before deriving the solution of U,Q subproblem, we first give some basic notations related to vectorial TV norm. Recall the matrix $U=[\mathbf{u}_1,\ldots,u_m,\ldots,u_M], U\in\mathbb{R}^{J\times M}$, each column of U is the vectorized image. We transform each column $\mathbf{u_m}\in\mathbb{R}^J$ into its matrix form $\widetilde{\mathbf{u}}\in\mathbb{R}^{B\times D}$, where $B\cdot D=J$. The transformed $\widetilde{U}=[\widetilde{\mathbf{u}}_\mathbf{1},\ldots,\widetilde{\mathbf{u}}_\mathbf{m},\ldots,\widetilde{\mathbf{u}}_\mathbf{M}]$ is now in the space of $\mathbb{R}^{B\times D\times M}$ and let $\mathbf{\Lambda}$ denote this new space. The vectorial TV term of U defined in (9) can then be equivalently reformulated for $\widetilde{U}$ as

$$R_{\text{vtv}}(\nabla\widetilde{U}) = \sum_{b,d}^{B,D}\sqrt{\sum_{m}^{M}\|(\nabla\widetilde{\mathbf{u}}_\mathbf{m})_{b,d}\|_2^2} \tag{23}$$

where ∇ is now the ordinary gradient operator of $\widetilde{\mathbf{u}}_\mathbf{m}$ in matrix domain. Operator ∇ maps $\widetilde{U}$ to $\nabla\widetilde{U}=[\nabla\widetilde{\mathbf{u}}_\mathbf{1},\ldots,\nabla\widetilde{\mathbf{u}}_\mathbf{m},\ldots,\nabla\widetilde{\mathbf{u}}_\mathbf{M}]$, and we denote the space of $\nabla\widetilde{U}$ as $\mathbf{\Gamma}$ here. The usual inner product and induced norm of $\mathbf{\Lambda}$ and $\mathbf{\Gamma}$ are defined as follows [39]:

$$\langle\widetilde{T},\widetilde{P}\rangle_{\mathbf{\Lambda}} = \sum_{m}^{M}\text{trace}(\widetilde{\mathbf{t}}_\mathbf{m}\widetilde{\mathbf{p}}_\mathbf{m}^*)$$

$$\|\widetilde{T}\|_{\mathbf{\Lambda}} = \sqrt{\langle\widetilde{T},\widetilde{T}\rangle_{\mathbf{\Lambda}}} = \sqrt{\sum_{m}^{M}\|\widetilde{\mathbf{t}}_\mathbf{m}\|_F^2} \tag{24}$$

$$\langle\nabla\widetilde{T},\nabla\widetilde{P}\rangle_{\mathbf{\Gamma}} = \langle\nabla^1\widetilde{\mathbf{T}},\nabla^1\widetilde{\mathbf{P}}\rangle_{\mathbf{\Lambda}} + \langle\nabla^2\widetilde{\mathbf{T}},\nabla^2\widehat{\mathbf{P}}\rangle_{\mathbf{\Lambda}}$$

$$\|\nabla\widetilde{T}\|_{\mathbf{\Gamma}} = \sqrt{\langle\nabla\widetilde{T},\nabla\widetilde{T}\rangle_{\mathbf{\Gamma}}} \tag{25}$$

where $\widetilde{\mathbf{p}}_\mathbf{m}^*$ denotes the conjugate transpose of $\widetilde{\mathbf{p}}_\mathbf{m}$, $\nabla^1\widetilde{T}$ and $\nabla^2\widetilde{T}$ denote the gradient image of $\widetilde{T}$ along horizontal and vertical directions, respectively. For an element $\widetilde{U}\in\mathbb{R}^{B\times D\times M}$ in $\mathbf{\Lambda}$ and its corresponding point U in space $\mathbb{R}^{J\times M}$, we show that $\|U\|_F^2$ is equivalent to $\|\widetilde{U}\|_{\mathbf{\Lambda}}^2$ defined above.

$$\|U\|_F^2 = \sum_{m}^{M}\sum_{j}^{J}|u_{jm}|^2 = \sum_{m}^{M}\sum_{b}^{B}\sum_{d}^{D}|\widetilde{u}_{bd,m}|^2 = \sum_{m}^{M}\|\widetilde{\mathbf{u}}_\mathbf{m}\|_F^2 = \|\widetilde{U}\|_{\mathbf{\Lambda}}^2. \tag{26}$$

With the notations defined above, we are able to demonstrate the solution of U,Q subproblem now. Noting that U and Q subproblems are almost identical, we first demonstrate how to solve the U problem and then the solution of Q could be obtained in the same way.

The objective function for U subproblem has the following form:

$$\min_U \nu_L R_{\text{vtv}}(\overline{\nabla}U) + \frac{\beta_L}{2}\|U-H\|_F^2 \quad \text{s.t.} \quad H = L - Z_L/\beta_L. \tag{27}$$

To keep notations clear, we denote $H=L-Z_L/\beta_L$ as an intermediate variable representing $L-Z_L/\beta_L$. Such vectorial TV model in (27) has been extensively studied in the literature [38]. We use the augmented Lagrangian multiplier (ALM) scheme to solve (27), which has been shown one of the fastest methods for such objectives [39]. We next explain in detail the ALM scheme that we use for the VTV subproblem.

Given the equivalence relationship between Frobenius norm in $\mathbb{R}^{J\times M}$ and the usual norm $\|\cdot\|_{\mathbf{\Lambda}}$ in $\mathbf{\Lambda}$, we rewrite the objective function of (27) with transformed $\widetilde{U},\widetilde{H}$.

$$\min_U \nu_L R_{\text{vtv}}(\nabla\widetilde{U}) + \frac{\beta_L}{2}\|\widetilde{U}-\widetilde{H}\|_{\mathbf{\Lambda}}^2 \quad \text{s.t.} \quad \widetilde{H} = \widetilde{L} - \widetilde{Z}_L/\beta_L. \tag{28}$$

By introducing an auxiliary variable $\widehat{E}\in\mathbf{\Gamma}$, the minimization problem of (28) is equivalent to the following constrained optimization problem:

$$\min_U \nu_L R_{vtv}(\widehat{E}) + \frac{\beta_L}{2}\|\widetilde{U}-\widetilde{H}\|_{\mathbf{\Lambda}}^2 \quad \text{s.t.} \quad \widehat{E} = \nabla\widetilde{U}, \widetilde{H} = \widetilde{L} - \widetilde{Z}_L/\beta_L. \tag{29}$$

We then define the augmented Lagrangian function of (29)

$$\mathcal{L}_{vtv}(\widetilde{U},\widehat{E}) = \nu_L R_{\mathbf{vtv}}(\widehat{E}) + \frac{\beta_L}{2}\|\widetilde{U}-\widetilde{H}\|^2_{\mathbf{\Lambda}} - \langle Z_{vtv}, \widehat{E}-\nabla\widetilde{U}\rangle_{\mathbf{\Gamma}} + \frac{\beta_{vtv}}{2}\|\widehat{E}-\nabla\widetilde{U}\|^2_{\mathbf{\Gamma}} \tag{30}$$

where $Z_{vtv} \in \mathbf{\Gamma}$ is Lagrangian multiplier, $\langle\cdot\rangle_{\mathbf{\Gamma}}$ and $\|\cdot\|_{\mathbf{\Gamma}}$ are inner product and norm in $\mathbf{\Gamma}$ defined in (25). The ALM scheme iteratively minimizes (30) with respect to each variable $\widetilde{U}, \widehat{E}$

$$\widetilde{U}^{k+1} = \min_{\widetilde{U}} \frac{\beta_L}{2}\|\widetilde{U}-\widetilde{H}\|^2_{\mathbf{\Lambda}} - \langle Z_{vtv}, \widehat{E}-\nabla\widetilde{U}\rangle_{\mathbf{\Gamma}} + \frac{\beta_{vtv}}{2}\|\widehat{E}-\nabla\widetilde{U}\|^2_{\mathbf{\Gamma}} \tag{31}$$

$$\widehat{E}^{k+1} = \min_{\widehat{E}} \nu_L R_{\mathbf{vtv}}(\widehat{E}) - \langle Z_{vtv}, \widehat{E}-\nabla\widetilde{U}\rangle_{\mathbf{\Gamma}} + \frac{\beta_{vtv}}{2}\|\widehat{E}-\nabla\widetilde{U}\|^2_{\mathbf{\Gamma}} \tag{32}$$

$$Z^{k+1}_{vtv} = Z_{vtv} - \beta_{vtv}(\widehat{E}-\nabla\widetilde{U}). \tag{33}$$

Given periodic boundary condition, we can use Fourier transform on the optimal condition of (31) to transform $\nabla^T\nabla$ into diagonal matrix. While such feasibility is not available for the vectorized image U, which is the reason we transform vectorized image U into its matrix form $\widetilde{U}$. The solution of $\widetilde{U}^{k+1}$ is [39]

$$\widetilde{U}^{k+1} = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{\beta_L\widehat{H} - \nabla^T Z_{vtv} + \beta_{vtv}\nabla^T\widehat{E}\}}{\beta_L\widehat{I} + \beta_{vtv}\mathcal{F}\{\nabla^T\nabla\}}\right\} \tag{34}$$

where $\widehat{I}$ is the identity element of $\mathbf{\Lambda}$. Here, applying Fourier transform to an element $\widetilde{U} = [\widetilde{\mathbf{u}}_\mathbf{1}, \ldots, \widetilde{\mathbf{u}}_\mathbf{m}, \ldots, \widetilde{\mathbf{u}}_\mathbf{M}]$ in $\mathbf{\Lambda}$ means applying Fourier transform to each $\widetilde{\mathbf{u}}_\mathbf{m} \forall m, m = 1, \ldots, M$. $\widehat{E}^{k+1}$ in (32) has closed form solution by soft shrinkage. Let $\widehat{\mathbf{e}}_{bd}$ denote the vector $(\widehat{e}_{bd,1}, \ldots, \widehat{e}_{bd,m}, \ldots, \widehat{e}_{BD,M})$, where $\widehat{e}_{bd,m}$ denotes the bdth entry of the mth frame image in $\widehat{E}$. Updating $\widehat{E}$ is accomplished by updating each $\widehat{\mathbf{e}}_{bd}$ as

$$\text{for } \forall(b,d) \quad \widehat{\mathbf{e}}_{bd} = \max\left\{1 - \frac{\nu_L}{\beta_{vtv}}/\sqrt{\sum_m^M \|\widehat{w}_{bd,m}\|_2^2}, 0\right\}\cdot\widehat{\mathbf{w}}_{bd}$$
$$\text{s.t.} \quad \widehat{W} = \nabla\widetilde{U} + Z_{vtv}/\beta_{vtv}. \tag{35}$$

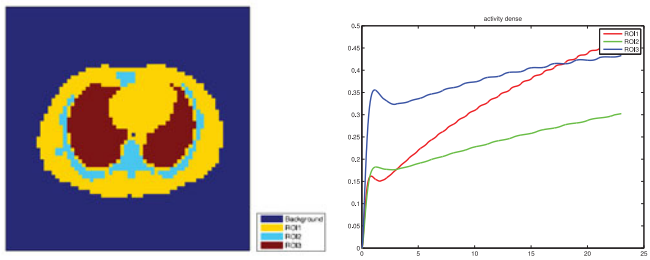
We then have an inner iterative procedure to alternatively update $\widetilde{U}$ and $\widehat{E}$ according to (34) and (35).

Subproblem of Q can be solved in the same way since they share the same form. The Lagrangian multipliers Z, Z_L, Z_S are updated as usual. By now, we have the update rules for all variables in our objective function (12). To wrap up the optimization algorithm, we summarize the iterative procedure in Algorithm 1.

Algorithm 1 Iteration Procedure of Optimization Algorithm

Require: Sinogram Y and System matrix G, weighting parameters $\lambda, \mu, \nu_L, \nu_S$.
Initialize $L^0 = S^0 = 0, U^0 = Q^0 = 0, Z = Z_L = Z_S = 0, X = \text{FBP}(Y)$.
repeat
  Compute L^{k+1} and S^{k+1} using (15) and (17).
  ▷ *Solving L, S subproblem*
  repeat ▷ *Solving X subproblem*
    Take E-step using (19).
    Take M-step using (22).
  until Relative change of $X < 10^{-3}$
    $U^{k+1} \leftarrow \text{VTV}U^k$, $Q^{k+1} \leftarrow \text{VTV}Q^k$. ▷*Solving U, Q subproblem*
  Update multipliers Z, Z_L, Z_S.
  $k \leftarrow k+1$
until Relative change of $L, S, X < 10^{-4}$.
Return L^k, S^k, X^k

function VTVU
  repeat
    Compute $\widehat{U}$ using (34).
    Update $\widehat{E}$ using (35).
    Update multipliers Z_{vtv}.
  until Relative change of $\widehat{U}, \widehat{E} < 10^{-3}$
  Return U^k
end function



(a) (b)

Fig. 1. (a) Zubal thorax phantom with three ROIs indicated by different colors. (b) TAC for three ROIs respectively.

E. Parameter Setting

During experiments, μ, ν_L and ν_S, these three regularization parameters are selected by evaluating reconstruction performance from a range of values. Specifically, μ is the weighting parameter of the likelihood function $\Psi(Y|X)$. It depends on noise level, and it should be strengthened when SNR increases. In the following experiments, it ranges from 0.001 to 0.005. ν_L and ν_S are coefficients before VTV norms. Since both of them serve as piecewise smoothers, we set the two variables with the same value, which lies in $[1\times10^{-4}, 1\times10^{-3}]$ in the experiments. Other parameters are fixed thorough out this paper, λ follows the rule from robust PCA theory [36], i.e., $\lambda = 1/\sqrt{\max(J,M)}$. All the Lagrangian penalty parameters $\beta, \beta_S, \beta_L, \beta_{vtv}$ are fixed at 0.1. As to the parameters of

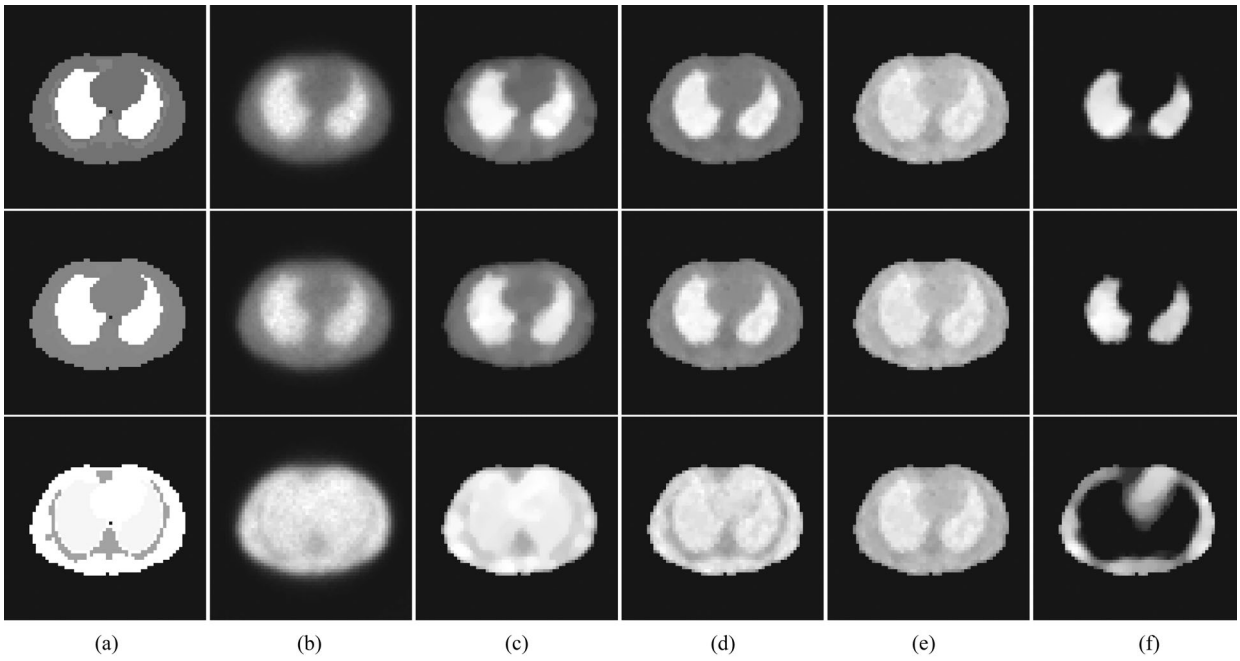


Fig. 2. Reconstructed activity images for Monte Carlo simulated Zubal phantom. From top row to bottom row: fifth, eighth and 17th frame.

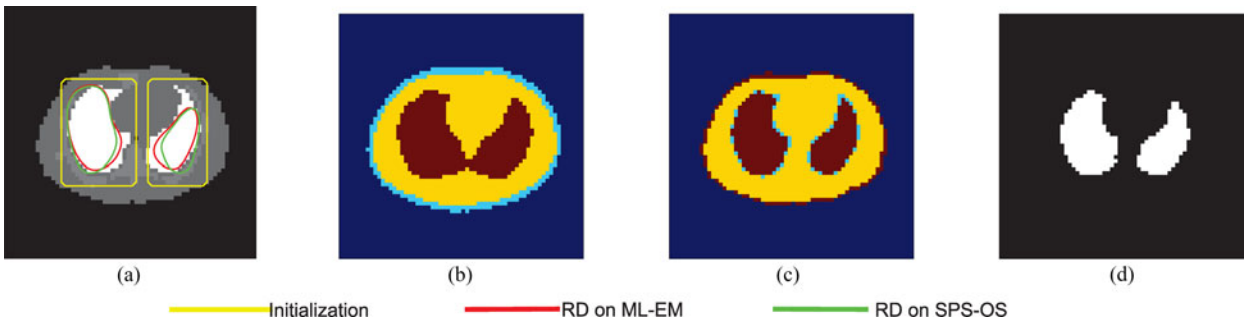


Fig. 3. Segmentation on the fifth frame of Monte Carlo simulated Zubal thorax phantom. (a) RD-based level-set on ML-EM and SPS-OS reconstructed image. (b) KSC on ML-EM reconstructed images. (c) KSC on SPS-OS reconstructed images. (d) Sparse component in the proposed model.

TABLE I
STATISTICAL STUDIES OF RECONSTRUCTED ZUBAL PHANTOM

Method	Bias	Variance	RMSE
ML-EM	0.0645	0.0206	0.1431
SPS-OS	0.0521	0.0189	0.1276
Ours	**0.0338**	**0.0074**	**0.0844**

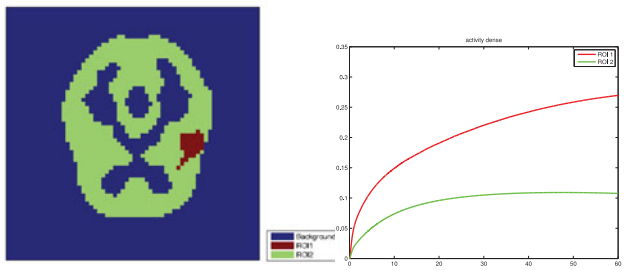
TABLE II
JACCARD INDEX OF SEGMENTATION FOR THE FIFTH FRAME OF ZUBAL PHANTOM

Method	RD on ML-EM	RD on SPS-OS	KSC on ML-EM	KSC on SPS-OS	Ours
Jaccard index	0.6832	0.7476	0.7863	0.7211	**0.9373**

compared algorithms in the following experiments, we tried our best to manually tune them regarding to reconstruction and segmentation accuracy statistics.

III. RESULTS

In this section, we validate the proposed model on both simulated phantom data and real patient data. The reconstructed images were compared to the classical ML-EM [2] and separable paraboloidal surrogates with ordered-subsets (SPS-OS) [28], [29]. The regularization term for SPS-OS method is the Huber potential function. On the other hand, two different methods for segmenting PET image sequence were examined along with the proposed model. One is the reaction diffusion (RD) model based on level-set with its initialization curve set manually as a rectangle surrounding the target [30]. Another is the kinetic spectral clustering (KSC) method [27], which views the TAC as data points and segment dynamic PET image sequence by clustering TACs using spectral clustering.



(a) (b)

Fig. 4. (a) Brain phantom with two ROIs indicated by different colors. (b) TAC for two ROIs.

In simulation studies, we are able to evaluate the algorithms by ground truth value. Quantitative results of reconstruction are given by relative bias, relative variance, and relative root mean square error (RMSE). They are calculated by bias $= \frac{1}{n}\sum_j^n |x_j - \hat{x_j}|/\hat{x_j}$, variance $= \frac{1}{n-1}\sum_j^n (\frac{x_j - \hat{x_j}}{\hat{x_j}})^2$ and RMSE $= (\frac{1}{n}\sum_j^n (\frac{x_j - \hat{x_j}}{\hat{x_j}})^2)^{\frac{1}{2}}$, where x_j denotes the reconstructed value at voxel j and $\hat{x_j}$ denotes the true value at voxel j. The bias, variance and RMSE presented in the following are averaged over all the frames of image sequence, whose intensity value have been normalized into [0,1]. We adopt the Jaccard index to measure the overlapping of the segmented region A and the true region B

$$\text{Jaccard index} = \frac{(A \cap B)}{(A \cup B)}. \tag{36}$$

Jaccard index ranges from 0, indicating no spatial overlap, to 1, indicating complete overlap.

The organization of this section is as follows: We first demonstrate the results of Monte Carlo simulated datasets based on Zubal phantom. In Monte Carlo simulated brain phantom, three datasets with different counting rates, i.e., the number of coincidence events, are examined to test the robustness of algorithms. At last, we show the results of the real patient data.

A. Monte Carlo Simulations

Monte Carlo simulation procedure is able to produce realistic sinogram data for PET by simulating the physical imaging process. More importantly, given ground truth, it enables us to evaluate the reconstruction algorithms quantitatively via bias-variance analysis. In this section, we present the results based on Monte Carlo simulation datasets including Zubal thorax phantom and brain phantom.

1) Zubal Thorax Phantom: The first experiment was conducted on Monte Carlo simulated data based on Zubal phantom, as shown in Fig. 1(a). TAC for three ROIs are shown in Fig. 1(b). The dynamic PET image sequence ($64 \times 64 \times 18$) was generated by two-compartment-model. The simulated tracer was ^{18}F-FDG and kinetic parameters were based on previous studies [13]. The PET scanner in simulation procedure is the Hamamatsu SHR74000. The final generated sinogram has 64×64 projections for every slice.

Fig. 2 shows the results of Monte Carlo simulated Zubal thorax phantom. Without spatial regularization, ML-EM method reconstructs noisy and blur images. As you can see in the first two rows of Fig. 2, SPS-OS method produces slight staircase effects. The proposed model successfully reconstructs the images and decomposes the sequence into low-rank and sparse components, as presented in Fig. 2(d)–(f). As we explained in previous sections, the low-rank constraint naturally enforce regional smoothness and vectorial TV norms also spatially smooth the TACs to achieve better reconstructions. Rank of the recovered low-rank matrix is one here that activity time curves for each region share the same curve in the low-rank matrix. The statistical analysis of reconstruction in Table I shows that our method gives the best results.

Before discussing the results of segmentation, we first give some details of presented algorithms and data postprocessing procedures. The initialization curves for level set-based RD method are manually set based on visual inspection. The KSC method operates on the whole image sequence. As a result, we show different clusters of voxels by different colors. For ease of quantitative comparison, we first threshold the sparse component in the proposed model by $0.05*$ *maximal pixel value* through out this experiment and the following to eliminate the random noise and binarize the resulting image.

Fig. 3 shows the results of three segmentation algorithms, the RD method and our sparse component of the fifth frame, and the results of KSC method for the whole image sequence. We can find that, due to low spatial resolution and noise, RD method does not describe the boundaries of objects well. The KSC method does a much better job. But both KSC and proposed method fail to identify the second ROI, as indicated with cyan color in Fig. 1(a). Fig. 3(d) is the binarized sparse component of proposed model of the fifth frame. The quantitative comparison of segmentation is given by the Jaccard index in Table II.

2) Brain Phantom: We next validate the proposed model on the Monte Carlo simulated brain phantom. The template of brain phantom and the corresponding TACs are shown in Fig. 4. A primary tumor is inserted into the phantom indicated as ROI1. We think the proposed model would be helpful in extracting tumors if the metabolism rate of tumor has significant difference from other normal tissues.

The dynamic image sequence is generated by the two-compartment-model. The simulated tracer is ^{18}F-FDG, and the corresponding kinetic parameters are based on previous work [41]. The raw sinogram data are produced by the Monte Carlo simulation procedure with the same settings as in last experiment. To validate the robustness of reconstruction algorithms, three groups of datasets with different counting rates $(10^4, 10^5, 10^6)$ are tested.

Fig. 5 shows the 17th reconstructed frame of brain image sequence for three level of counting rates. Three rows correspond to $10^4, 10^5, 10^6$ counting rates. SPS-OS method could preserve image structure well in $10^5, 10^6$ cases. When it comes to 10^4, both ML-EM and SPS-OS reconstructed images are severely corrupted that the tumor region becomes hard to visually identify. While the proposed method is able to produce more accurate and stable results. The quantitative evidence is in Table III. Our method has much better bias and variance performance.

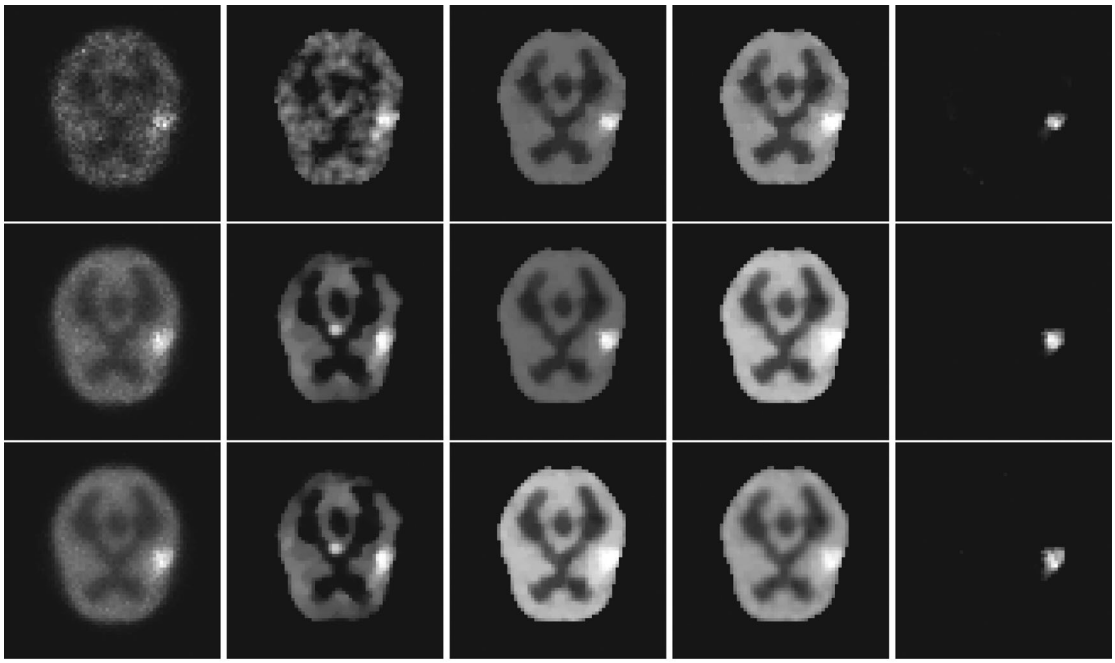


Fig. 5. 17th frame of reconstructed activity images for Monte Carlo simulated brain phantom. From top row to bottom row: 10^4, 10^5, 10^6 counting rates.

TABLE III
STATISTICAL RESULTS OF RECONSTRUCTED BRAIN PHANTOM IMAGES

Method	10^4 counts			10^5 counts			10^6 counts		
	Bias	Variance	RMSE	Bias	Variance	RMSE	Bias	Variance	RMSE
ML-EM	0.1199	0.0738	0.2611	0.0716	0.0350	0.1649	0.0650	0.0308	0.1519
SPS-OS	0.1039	0.0662	0.2446	0.1011	0.0679	0.2394	0.0758	0.0421	0.1830
Ours	**0.0530**	**0.0250**	**0.1264**	**0.0409**	**0.0117**	**0.1027**	**0.0363**	**0.0100**	**0.0930**

As for segmenting the brain phantom, we are interested in accurately extracting the tumor regions. The results of the 17th frame for three counting rates are demonstrated in Fig. 6 along with Jaccard index of segmented tumors in Table IV. For the level-set-based RD method, we find it does no converge in low counting rates data. In such cases, we have to manually set the maximum iteration number to early stop the algorithm. Anyway, such gradient and regional homogeneity-based method is challenged on such low resolution and noisy images. We can also find that the KSC method is sensitive to the increasing noise in dynamic PET images. On the other hand, using temporal change and the coupling between frames by VTV norm, the proposed method gives accurate and stable results, as indicated in Table IV. In 10^4 counting case that KSC method failed to identify the tumor and background, we use ∞ to denote failure.

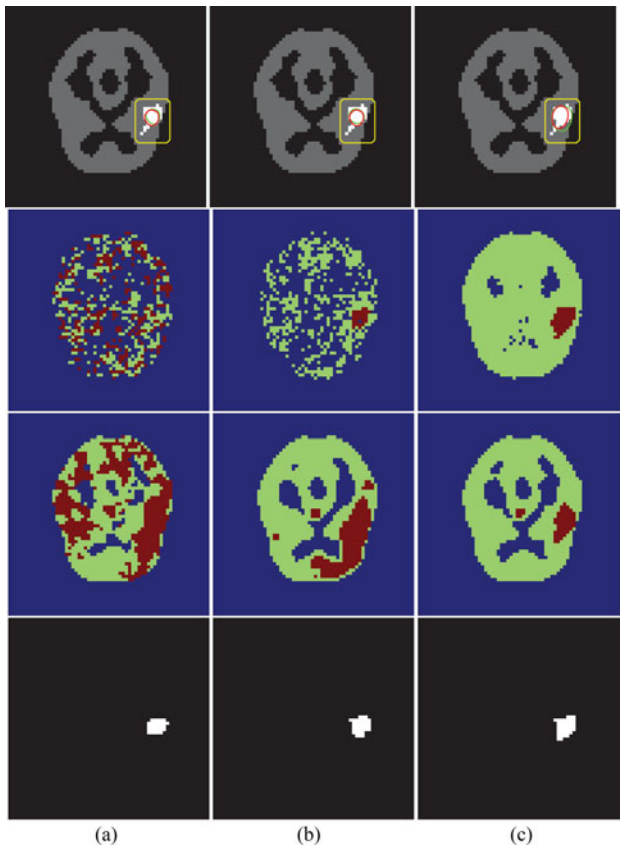


Fig. 6. Segmentation on the 17th frame of Monte Carlo simulated brain phantom. First row: RD-based level-set on ML-EM and SPS-OS reconstructed images. Second row: KSC on ML-EM reconstructed image sequence. Third row: KSC on SPS-OS reconstructed image sequence. Fourth row: sparse component of the proposed model.

B. Real Patient Data

The real data in this study is a dynamic PET scan acquired from a cardiac patient at a local hospital. The scanner is Hamamatsu SHR-22000 whole body PET scanner. It has 32

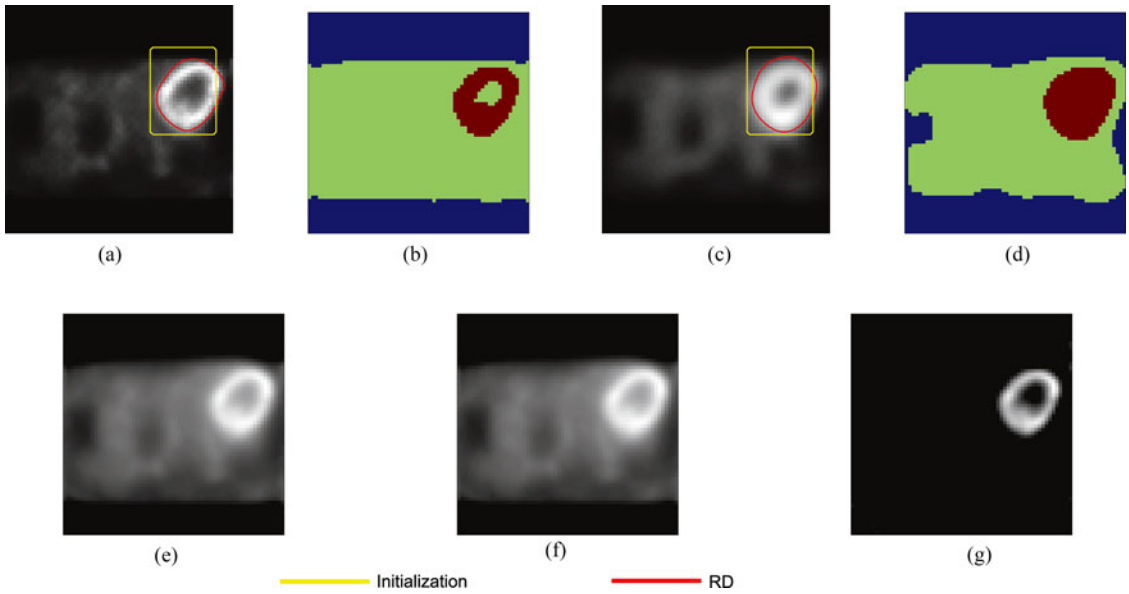


Fig. 7. Results for the fourth frame reconstructions and segmentation of the real cardiac patient data. (a) RD om ML-EM, (b) KSC on ML-EM, (c) RD on SPS-OS, (d) KSC on SPS-OS, (e) Sparse + Low-rank, (f) Low-rank, (g) Sparse.

TABLE IV
JACCARD INDEX OF SEGMENTED 17th FRAME FOR MONTE CARLO SIMULATED BRAIN PHANTOM DATA

Method	RD on ML-EM	RD on SPS-OS	KSC on ML-EM	KSC on SPS-OS	Ours
10^4 counts	0.6667	0.6401	∞	∞	**0.7105**
10^5 counts	0.5278	0.5128	0.6401	0.1636	**0.7750**
10^6 counts	0.4444	0.3056	0.6481	0.5373	**0.8293**

crystal rings and is operated in 2-D acquisition mode during the scan process. The transaxial resolution of the central FOV is 3.7 mm/pixel. A 10-min transmission scan using rotating ^{68}Ge line sources was performed then 10mCi ^{18}F-FDG was injected to test metabolism of patient. A dynamic scan was performed consisting of 19 frames with variable frame length for a total duration of 60 min.

The reconstruction and segmentation results are presented together in Fig. 7. The endocardial boundary was not detected by RD method on the ML-EM and SPS-OS reconstructed results. KSC on ML-EM reconstructed image successfully identifies the myocardium region. While KSC on SPS-OS fails to segment the endocardial boundary, as indicated in Fig. 7(d). In this experiment, we observed that both the reconstructed image quality and the segmentation result have been improved through using our proposed method.

C. Implementation Analysis

1) Computation Time: We summarize the computation time of each experiment in Table V. All algorithms are implemented in the Matlab environment (Version 2009a) on 3.40-GHz CPU, 16 RAM personal computer. For reconstruction and segmentation that are obtained by separate algorithms, like ML-EM and RD method, we provide the the overall computation time along with the time consumption of each individual procedure. For instance, the computation time for ML-EM + RD method is represented as 14.7 s (6.7 + 8.0). It means the ML-EM reconstruction takes 6.7 s and RD segmentation takes 8.0 s. So the sum of the two procedures is 14.7 s.

In Table V, the proposed method is slower compared to other algorithms. However, RD method can only delineate one object in a single frame after manual initialization. As indicated in the previous experiments, the proposed method obtains more accurate reconstruction and segmentation since these two procedures are tackled coherently. The computation time of proposed algorithm is currently limited by the singular value decomposition procedure and the subproblem of vectorial TV norm. We will explore the possibility of SVD-free method to accelerate the algorithms in the future [44].

2) Parameter Analysis: To examine the robustness of the tuning parameters within the given range, we calculate the relative bias and Jaccard index based on different choices of μ, ν_L and ν_S. The definition of Jaccard index is the same as (36). μ is tested within $[0.001, 0.01]$, while ν_L and ν_S share the same value, which is examined between 1×10^{-4} and 1.5×10^{-3}. The sinogram data in this study is the Monte Carlo simulated Zubal phantom in Section III-A1. When different values of μ is tested, ν_L and ν_S are fixed at 5×10^{-4}. μ is set 0.003 for the experiments in terms of different choices of ν_L and ν_S.

Fig. 8(a) and (b) shows that bias is less than 0.04 with $\mu \in [0.001, 0.01]$ and the segmentation accuracy is relatively stable when $\mu > 0.003$. The reconstruction accuracy of the proposed method is sensitive to ν_L, ν_S when they are less than 4×10^{-4} as shown in Fig. 8(c). Moreover, the proposed method may produce relatively stable segmentation results in terms of the Jaccard index even using different ν_L and ν_S as seen in Fig. 8(d).

TABLE V
COMPUTATION TIME OF PRESENTED ALGORITHMS

Method	Zubal	Brain 10^4	Brain 10^5	Brain 10^6	Patient
ML-EM + RD	**14.7 s** (6.7 + 8.0)	**9.7 s** (6.5 + 3.2)	**9.6 s** (6.5 + 3.1)	**9.7 s** (6.5 + 3.2)	**9.0 s** (5.6 + 3.4)
ML-EM + KSC	32.1 s (6.7 + 25.4)	30.3 s (6.5 + 23.8)	29.5 s (6.5 + 23.0)	31.2 s (6.5 + 24.7)	51.7 s (5.6 + 46.1)
SPS-OS + RD	26.4 s (18.6 + 8.4)	19.6 s (16.6 + 3.0)	20.8 s (17.8 + 3.0)	20.8 s (17.7 + 3.1)	18.6 s (15.1 + 3.5)
SPS-OS + KSC	42.7 s (18.6 + 24.1)	42.6 s (16.6 + 26.0)	42.8 s (17.8 + 25.0)	41.5 s (17.7 + 23.8)	38.6 s (15.1 + 23.8)
Ours	38.3 s	92.8 s	90.4 s	106.3 s	65.7 s

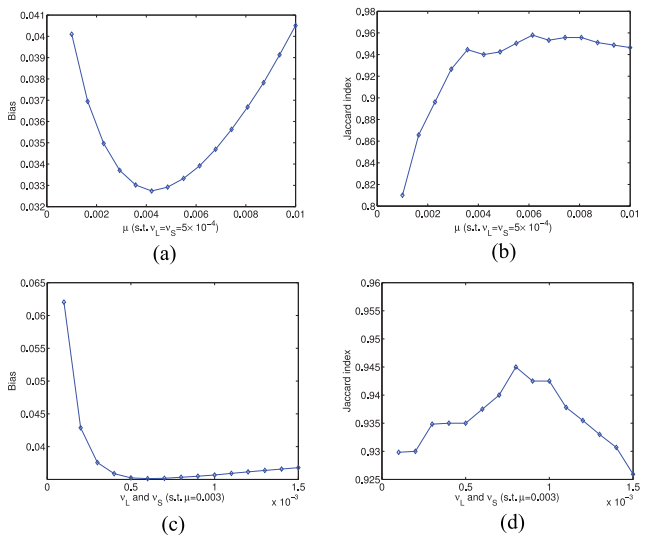


Fig. 8. Relative bias and Jaccard index under different choices of μ and ν_L, ν_S. $\mu \in [0.001, 0.01]$ and $\nu_L, \nu_S \in [1 \times 10^{-4}, 1.5 \times 10^{-3}]$. (a) Bias for μ, (b) Jaccard index for μ, (c) Bias for v_L, v_S, (d) Jaccard index for v_L, v_S.

IV. CONCLUSION

We have developed a low-rank and sparse decomposition framework that dynamic PET reconstruction and tissue segmentation can be achieved simultaneously. To suppress the noises in the measurement, we incorporate Poisson likelihood function for data fidelity and vectorial TV norm into the objective function. Since segmentation is often posed as a piecewise constant approximation while low-rank matrix is also composed of several homogeneous regions. The vectorial TV norm is applied to both low-rank and sparse matrices. Both experiments on realistic simulation data and real patient data show that the proposed model gives competitive results. We believe the reported promising results on reconstruction and segmentation could be of help for parametric image calculation and ROI-based quantification.

REFERENCES

[1] S. Basu and Y. Bresler, "O($N^2 log_2 N$) filtered backprojection reconstruction algorithm for tomography," *IEEE Trans. Imag. Process.*, vol. 9, no. 10, pp. 1760–1773, Oct. 2000.

[2] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imag.*, vol. 1, no. 2, pp. 113–122, Oct. 1982.

[3] E. Levitan and G. T. Herman, "A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography," *IEEE Trans. Med. Imag.*, vol. 6, no. 3, pp. 185–192, Sep. 1987.

[4] H. Liu *et al.*, "PET image reconstruction: A robust state space approach," in *Information Processing in Medical Imaging*. New York, NY, USA: Springer, 2005, pp. 197–209.

[5] J.-H. Chang *et al.*, "Regularized image reconstruction algorithms for positron emission tomography," *IEEE Trans. Med. Imag.*, vol. 23, no. 9, pp. 1165–1175, Sep. 2004.

[6] R. M. Leahy and J. Qi, "Statistical approaches in quantitative positron emission tomography," *Statist. Comput.*, vol. 10, no. 2, pp. 147–165, 2000.

[7] Z. Harmany *et al.*, "This is SPIRAL-TAP: Sparse poisson intensity reconstruction algorithms: Theory and practice," *IEEE Trans. Imag. Process.*, vol. 21, no. 3, pp. 1084–1096, Mar. 2012.

[8] J. Qi and R. M. Leahy, "Iterative reconstruction techniques in emission computed tomography," *Phys. Med. Biol.*, vol. 51, no. 15, pp. R541–R578, 2006.

[9] J. Bowsher *et al.*, "Bayesian reconstruction and use of anatomical a priori information for emission tomography," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 673–686, Oct. 1996.

[10] A. Rahmim *et al.*, "Four-dimensional (4D) image reconstruction strategies in dynamic PET: Beyond conventional independent frame reconstruction," *Med. Phys.*, vol. 36, pp. 3654–3670, 2009.

[11] T. Nichols *et al.*, "Spatiotemporal reconstruction of list mode PET data," *IEEE Trans. Med. Imag.*, vol. 21, no. 4, pp. 396–404, Apr. 2002.

[12] B. W. Reutter *et al.*, "Direct least-squares estimation of spatiotemporal distributions from dynamic SPECT projections using a spatial segmentation and temporal B-splines," *IEEE Trans. Med. Imag.*, vol. 19, no. 5, pp. 434–450, May 2000.

[13] S. Tong and P. Shi, "Tracer kinetics guided dynamic PET reconstruction," in *Information Processing in Medical Imaging*. New York, NY, USA: Springer, 2007, pp. 421–433.

[14] M. E. Kamasak *et al.*, "Direct reconstruction of kinetic parameter images from dynamic PET data," *IEEE Trans. Med. Imag.*, vol. 24, no. 5, pp. 636–650, May 2005.

[15] J. S. Maltz, "Optimal time-activity basis selection for exponential spectral analysis: Application to the solution of large dynamic emission tomographic reconstruction problems," *IEEE Trans. Nucl. Sci.*, vol. 48, no. 4, pp. 1452–1464, Aug. 2001.

[16] J. Verhaeghe *et al.*, "Dynamic PET reconstruction using wavelet regularization with adapted basis functions," *IEEE Trans. Med. Imag.*, vol. 27, no. 7, pp. 943–959, Jul. 2008.

[17] M. Yavuz and J. A. Fessler, "Statistical image reconstruction methods for randoms-precorrected PET scans," *Med. Imag. Anal.*, vol. 2, pp. 369–378, 1998.

[18] S. Osher and J. A. Sethian, "Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.*, vol. 79, pp. 12–49, 1988.

[19] J. Cheng-Liao and J. Qi, "Segmentation of mouse dynamic PET images using a multiphase level set method," *Phys. Med. Biol.*, vol. 55, pp. 6549–6569, 2010.

[20] C. Li *et al.*, "A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI," *IEEE Trans. Imag. Process.*, vol. 20, no. 7, pp. 2007–2016, Jul. 2011.

[21] Y. Boykov *et al.*, "Fast approximate energy minimization via graph cuts," in *Proc. Int. Conf. Comput. Vision*, 1999, pp. 377–384.

[22] J. L. Chen *et al.*, "Markov random field models for segmentation of PET images," in *Information Processing in Medical Imaging*. New York, NY, USA: Springer, 2001, pp. 468–474.

[23] X. Chen *et al.*, "Medical image segmentation by combining graph cuts and oriented active appearance models," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2035–2046, Apr. 2012.

[24] S. C. Mitchell *et al.*, "Segmentation of cardiac MR images: An active appearance model approach," in *Proc. Med. Imag. Int. Soc. Opt. Photon.*, 2000, pp. 224–234.
[25] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.*, vol. 15, pp. 1373–1396, 2003.
[26] A. Y. Ng *et al.*, "On spectral clustering: Analysis and an algorithm," in *Proc. Neural Inform. Process. Syst.*, 2001, pp. 849–856.
[27] S. Mouysset *et al.*., (2013). Segmentation of dynamic PET images with kinetic spectral clustering. *Phys. Med. Biol.* [Online]. *58(19)*, pp. 6931–6944. Available: http://stacks.iop.org/0031-9155/58/i=19/a=6931
[28] J. A. Fessler and H. Erdogan, "A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction," in *Proc. Nuclear Sci. Symp. Conf. Record*, 1998, pp. 1132–1135.
[29] A. Mehranian *et al.*, "An ordered-subsets proximal preconditioned gradient algorithm for edge-preserving pet image reconstruction," *J. Med. Phys.*, vol. 40, no. 5, p. 052503, 2013. doi: 10.1118/1.4801898
[30] K. Zhang *et al.*, "Reinitialization-free level set evolution via reaction diffusion," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 258–271, Jan. 2013.
[31] M. E. Kamasak and B. Bayraktar, "Clustering dynamic PET images on the projection domain," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 3, pp. 496–503, Jun. 2007.
[32] M. E. Kamasak, "Clustering dynamic PET images on the Gaussian distributed sinogram domain," *Comput. Methods Programs Biomed.*, vol. 93, pp. 217–227, 2009.
[33] S. Yoon, A. R. Pineda, R. Fahrig, "Simultaneous segmentation and reconstruction: A level set method approach for limited view computed tomography," *J. Med. Phys.*, vol. 37, no. 5, pp. 2329–2340, 2010.
[34] H. Gao, J. F. Cai, Z. Shen, and H. Zhao, "Robust principal component analysis-based four-dimensional computed tomography," *Phys. Med. Biol.*, vol. 56, no. 11, pp. 3181–3198, 2011.
[35] R. Otazo *et al.*, "Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components," *Magn. Resonance Med.*, vol. 73, pp. 1125–1136, 2014.
[36] E. J. Candès, X. Li, M. Yi, and J. Wright, "Robust principal component analysis?," *J. ACM*, vol. 58, no. 3, art. no. 11, May 2011.
[37] P. Blomgren and T. F. Chan, "Color tv: Total variation methods for restoration of vector-valued images," *IEEE Trans. Imag. Process.*, vol. 7, no. 3, pp. 304–309, Mar. 1998.
[38] X. Bresson and T. Chan, "Fast dual minimization of the vectorial total variation norm and applications to color image processing," *Inverse Problems Imag.*, vol. 2, pp. 455–484, 2008.
[39] C. Wu and X. Tai. (2010). Augmented lagrangian method, dual methods, and split bregman iteration for rof, vectorial tv, and high order models. *SIAM J. Imag. Sci.* [Online]. *3(3)*, pp. 300–339. Available: http://dx.doi.org/10.1137/090767558
[40] J. Yang and X. Yuan, "Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization," *Math. Comput.*, vol. 82, no. 281, pp. 301–329, 2013.
[41] A. M. Spence *et al.*, "18f-FDG PET of gliomas at delayed intervals: Improved distinction between tumor and normal gray matter," *J. Nuclear Med.*, vol. 45, no. 10, pp. 1653–1659, 2004.
[42] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," 2010.
[43] J.-F. Cai *et al.*, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, pp. 1956–1982, 2010.
[44] J.-F. Cai and S. Osher, "Fast singular value thresholding without singular value decomposition," Dept. Math., Univ. California, Los Angeles, CA, USA, Rep. no. cam10–24, 2010.

Authors' photographs and biographies not available at the time of publication.