

Nonlocal Low-Rank and Total Variation Constrained PET Image Reconstruction

Nuobei Xie

Department of Optical Engineering
Zhejiang University
Hangzhou, China
xnb@zju.edu.cn

Yunmei Chen

Department of Mathematics
University of Florida
Gainesville, USA
yun@ufl.edu

Huafeng Liu

Department of Optical Engineering
Zhejiang University
Hangzhou, China
liuhf@zju.edu.cn

Abstract—Many efforts have been made for decades in order to improve the accuracy of radioactivity map in positron emission tomography (PET) images, which has important clinical implications for better diagnosis and understanding of diseases. However, there is still a challenging problem for reconstructing high resolution image with the limited acquired photon counts. In this paper, we present a nonlocal self-similar constraint for the purpose of exploiting structured sparsity within the PET reconstructed images. It is based on image patches and approached by low-rank approximation. Moreover, we adopt total variation regulation into our method to further denoise and compensate the demerits inherited in patch-based methods. These two regulation terms are firstly employed in the Poisson model, and are jointly solved in a distributed optimization framework. Experiments have presented that our proposed PNLTV method substantially outperforms existing state-of-the-art methods in PET reconstruction.

Keywords—positron emission tomography; Poisson model; image reconstruction; nonlocal; low-rank approximation; total variation

I. INTRODUCTION

As a typical type of Emission Computed Tomography (ECT), positron emission tomography (PET) is irreplaceably used in functional imaging. By injecting a certain isotope-labeled compound, positrons will be emitted and annihilated with nearby electrons in the body. The photons generated by annihilations can be collected and then be used to recover the tomographic radioactivity maps. How to recover high quality images is continuously studied and discussed by numerous researchers, given the increasing diagnosis requirement.

The emission image reconstruction algorithms largely fall into two groups: analytic strategies and iterative statistical methods. The first group is filtered back-projection (FBP) based algorithms, or modifications thereof, which is built on the Radon transform. Although this method is fast and easy to be implemented, it overlooks the statistical property and fails to produce high quality reconstructed images.

Iterative statistical approaches have become the primary focus of researches. Maximum likelihood-expectation maximization (ML-EM) [1][2] integrating the Poisson model, serves as a milestone in PET reconstruction algorithms. Nevertheless, it suffers from the ill-conditioning problem, i.e.

the solutions are sensitive to small changes in the data [3] and consequently lead to the ‘checkboard’ effect in recovered images. In order to overcome these drawbacks, image priors (or regulations) are introduced into the objective function, forming the *maximum a posteriori* (MAP) estimation [4], or penalized ML method [5]. Typical priors and regulations include quadratic regulation [6], Gaussian prior [7], Huber prior and Gibbs prior [8]. Though many algorithms are already developed, there is still a tradeoff between the noise depression and the resolution improvement in practice. What is more, the traditional priors may overlook the structural information and hence over-smooth the edges of different regions.

Like some other medical images are, PET images are fairly assumed to have the piece-wise constant property-- each pixel in the same region should have the same value, while the boundary of each region is allowed to be sharply-variant. Total variation (TV) [9] minimization is properly used based on this feature. From the other perspective, with the rapid rising of compressive sensing, sparsity is well employed in this piece-wise constant images. Several related efforts were made in this area. Xu developed a method based on dictionary learning to exploit structural sparsity in the low-dose CT reconstruction [10]. Simultaneous sparse coding was proposed in [11]. More recently, it was found that both natural and medical images have plenty of similar structures within themselves, and this originated the conception of ‘nonlocal self-similarity’, which has led to the famous nonlocal means methods [12], block-matching 3D denoising (BM3D)[13], *etc.* Also, nonlocal self-similarity was introduced in [14] and [15] to enhance the structural sparsity dramatically.

Although TV constraint is widely reckoned as eminent in providing edge-preserving guidance and proved to be quite efficient in denoising of medical image, it unfortunately causes the so-called ‘stair-case’ effect, which degrades the accuracy and the visibility of the recovered image. On the other hand, while the nonlocal patch-based methods are remarkably in keeping structural information, they are less effective in ruling out the noise, especially when the penal parameter is small. However, if the parameter were tuned to a large degree, the recovered image would suffer from the rectangular-shaped edge. In that sense, it is natural to develop an algorithm which utilizes the merits of both methods.

In this paper, we propose a novel algorithm which exploits both the local and nonlocal self-similarity of reconstructed

image to remarkably improve the resolution while keeping the noise at a low level. Nonlocal traits conducted by the low rank approach are firstly introduced into the reconstruction based on Poisson model, while total variation method is simultaneously implemented in this frame in order to compensate the shortcoming which inherited in patch-based methods. We implement these two regulations into Poisson model jointly, and optimize this function in an alternative and distributed way. Denoted as PNLTV, our proposed method shows its substantial superiority in reconstruction accuracy, denoising quality as well as robustness. Detailed implementation and experiments are presented in the following sections.

II. METHOD

A. PET imaging model

As a typical type of Emission Computed Tomography (ECT), the reconstruction of PET can be modeled as the affine transform,

$$\bar{y} = Gx + r + s, \quad (1)$$

where $\bar{y}$ denotes the expectation of sinogram y , and x denotes the radioactivity image which is supposed to be reconstructed. Here $y = \{y_q, q = 1, \dots, M\}$ denotes the sum of collected photons along each detector pair, and M is the number of detector pairs; $G \in \mathbb{R}^{M \times N}$ is the system matrix, whose entry G_{qj} is the average probability of detecting a *true coincidence* from voxel site j at *Line of Response (ROI) q* ; r and s refer to *random incidence* and *scattered coincidence* respectively.

Given the photon detecting nature mentioned above, y is modeled by independent Poisson distribution: $y \sim \text{Poisson}\{\bar{y}\}$, In that sense, the likelihood function of y can be written as:

$$\Pr(y|x) = \prod_q^M e^{-\bar{y}_q} \frac{\bar{y}_q^{y_q}}{y_q!}, \quad (2)$$

And for the sake of optimizational convenience, we often minimize the negative log-likelihood function instead:

$$\begin{aligned} \min_x P(x) &= \min_x -\log(\Pr(y|x)) = \min_x \sum_q^M \bar{y}_q - y_q \log(\bar{y}_q) \\ \text{s.t. } \bar{y} &= Gx + r + s \end{aligned} \quad (3)$$

The constant term $\log(y_q!)$ is left out in (3). It's also worth noticing that since the log function is monotonic, (3) is equal to maximizing the probability in (2).

B. Problem Formulation

Furthermore, based on (3), we can organize our regularized objective function in the form of:

$$\min_{x, L, \omega} P(x) + \alpha \text{NL}(x, L) + \beta \text{TV}(x, \omega). \quad (4)$$

Here the latter two terms jointly serve as the prior: $\text{NL}(x, L)$ denotes the nonlocal low-rank constraint which aiming at exploiting the nonlocal structured sparsity; $\text{TV}(x, \omega)$ denotes the total variation constraint which is used to enhance local structure sparsity and complement the rectangular-edge effect inherited in patch-based methods; $P(x)$ denotes the fidelity

term mentioned in (3); α, β serve as the weighting parameter respectively.

1) Nonlocal low-rank regularization

In this model, we use a low-rank constraint to exploit the nonlocal property in the reconstructed image. The implementation of nonlocal regularization includes 2 steps: grouping the self-similar characterized patches within the image and enforcing low-rank constraint on the grouped matrix [15]. The basic assumption underlying regularization term is that self-similarity is abundant in the restored image [15].

Practically, after each iteration, we can obtain lots of $\sqrt{n} \times \sqrt{n}$ sized patches $x_i \in \mathbb{C}^n$ at position i from the temporary estimated image. For each exemplar patch vector x_i , we search its m -nearest patches based on *Euclidean* distance within the image, i.e. $S_i = \{s | \|x_i - x_{i,s}\| < T_i\}$, where T_i is distance threshold which is determined by the m -th nearest patch to x_i , and S_i is the set of positions of patches related to x_i . Then we can get a matrix $X_i = [x_{i,0}, x_{i,1}, \dots, x_{i,s}, \dots, x_{i,m-1}]$, $X_i \in \mathbb{C}^{n \times m}$ based on each exemplar patch vector x_i , where $x_{i,0}$ is denotes the patch vector x_i .

The low-rank constraint comes after the grouping procedure. Since the image has the nonlocal self-similar structures, each matrix X_i is meant to possess low-rank property. In practice, each matrix X_i can be decomposed into $L_i + W_i$, where L_i is the low-rank matrix while W_i is the Poisson noise matrix which is meant to be eliminated by the constraint. Therefore, L_i can be recovered by the following simplified model:

$$L_i = \arg \text{rank}(L_i), \text{ s.t. } \|X_i - L_i\|_F^2 \leq \sigma_\omega^2, \quad (5)$$

where $\|\cdot\|_F^2$ is the Frobenious norm and σ_ω^2 denotes the variance of Poisson noise.

2) Total variation regulation

Rudin, Osher, and Fatemi (ROF) [16] first introduced total variation (TV) method into the image denoising. From then on, TV method had become a classic and widely-discussed topic in this field. The general version of TV objective function in image reconstruction is:

$$\min_x \sum_j \|D_j x\|_p \quad \text{s.t. } \bar{y} = Gx + r + s, \quad (6)$$

where the operating matrix D_j to represent the discrete differential operation at position j ; $p=1$ or 2 .

However, in order to implement the augment Lagrangian method into global optimization, we consider an equivalent l_2 form of (6):

$$\begin{aligned} \min_x \sum_j \|\omega_j\|_2 \\ \text{s.t. } \bar{y} = Gx + r + s \text{ and } D_j x = \omega_j \text{ for all } j \end{aligned} \quad (7)$$

C. The Optimization of PNLTV

Solving the objective function (4) directly is nearly an impossible mission. However, note that the objective function (4) is composed of three subproblems related to x, L, ω , thus we would minimize (4) by each variable while make other variables fixed. In this section we demonstrate our distributed optimization method with respect to each subproblem.

1) L-subproblem

During the optimization procedure, suppose a estimation of image x^k is given after k -th iteration, we can get a set of $X_i^{k+1} = B_i^{k+1} x^k = [x_{i,0}^k, x_{i,1}^k, \dots, x_{i,m-1}^k] \in C^{n \times m}$, $i = 1, 2, \dots, n_i$, where B_i^{k+1} denotes the operative matrix which can extract X_i^{k+1} from x^k . While the original low-rank approximation problem is given by (4), we practically rewrite the problem into a Lagrangian form by choosing a proper λ , namely:

$$\begin{aligned} L_i^{k+1} &= \arg \min_{L_i} \|X_i^{k+1} - L_i\|_F^2 + \lambda \|L_i\|_* \\ &= \arg \min_{L_i} \|B_i^{k+1} x^k - L_i\|_F^2 + \lambda \sum_u \sigma_u(L_i) \end{aligned} \quad (8)$$

where $\|\cdot\|_*$ denotes nuclear norm, or sum of the singular value in other words; σ_u denotes the u -th largest singular value.

By approximating the rank to nuclear norm, the low-rank model can be turned into convex optimization from NP-hard problem. In this algorithm, we adopt the singular value thresholding (SVT) method [17] to solve the rank-minimization problem efficiently and effectively. In that sense, L_i^{k+1} can be estimated by:

$$L_i^{k+1} = U(\tilde{\Sigma} - \lambda)_+ V^T \quad (9)$$

where $U \tilde{\Sigma} V^T$ denotes the singular value decomposing (SVD) of matrix X_i^{k+1} , and $(x)_+ = \max(x, 0)$

2) ω -subproblem

Based on Augmented Lagrangian method, we formulate the ω -subproblem as:

$$\omega^{k+1} = \arg \min_{\omega} \|\omega\|_2 + \frac{\eta}{2} \|\omega - Dx^k\|_2^2 - \nu^k (\omega - Dx^k) \quad (10)$$

where η is called penalty parameter and ν^k is called Lagrange multiplier. Equation (10) share the same purpose with (7). Applying shrinkage operator to (10) yields the algorithm:

$$\omega^{k+1} = \max \left\{ \left\| Dx^k - \frac{\nu^k}{\eta} \right\|_2 - \frac{1}{\eta}, 0 \right\} \frac{Dx^k - \nu^k / \eta}{\left\| Dx^k - \nu^k / \eta \right\|_2}, \quad (11)$$

and the multiplier is updated as: $\nu^{k+1} \leftarrow \nu^k - \eta(Dx^k - \omega^{k+1})$.

3) x -subproblem

After L, ω being updated, current estimation of reconstructed image x can be solved by our proposed algorithm with other variable fixed. Actually previous subproblems both serve as the penalty terms of x -subproblem, thus this part plays a key role in our proposed framework.

Algorithm 1: PET reconstruction via joint regulations (PNLTV)

Input: Sinogram y and system matrix G , weighting parameter $\alpha, \beta, \lambda, \eta$.

1: **Initialization:** $k = 0, x^0 = FBP(y)$.

2: **Repeat:**

3: Compute matrix $X_i^k, \forall i$ using (9) $\triangleright L$

4: Update differential vector ω^{k+1} using (11) $\triangleright \omega$

5: Update multiplier $\nu^{k+1} \leftarrow \nu^k - \eta(Dx^k - \omega^{k+1})$.

6: **Repeat:** $\triangleright x$

7: E-step: compute $\Omega(x; x^k)$ using (13).

8: M-step: update $x_j^{k+1}, \forall j$ using (15).

9: **Until:** Inner Relative change $(x^{k+1} - x^k)/x^{k+1} < 10^{-4}$

10: $k \leftarrow k + 1$

11: **Until:** Relative change $(x^{k+1} - x^k)/x^{k+1} < 10^{-5}$

12: **Output:** x^k

Initially, based on (4), our objective function can be written in a joint way:

$$\begin{aligned} \Omega(x, L, \omega) &= \sum_q \bar{y}_q - y_q \log(\bar{y}_q) + \alpha \left(\sum_i \|B_i x - L_i\|_F^2 + \lambda \|L_i\|_* \right) \\ &\quad + \beta (\|\omega\|_2 + \frac{\eta}{2} \|Dx - \omega\|_2^2) \\ \text{s.t. } \bar{y} &= Gx + r + s. \end{aligned} \quad (12)$$

Here we adopt L to represent L^{k+1} and ω to represent ω^{k+1} for illustrative simplicity. Thus, in this subproblem, the aim of the optimization is to solve:

$$\begin{aligned} x &= \arg \min_x \Omega(x; x^k) = \arg \min_x \sum_j \sum_q (g_{qj} x_j - \hat{c}_{qj} \log(g_{qj} x_j)) + \\ &\quad \alpha \sum_i \sum_j (B_{ij} x_j - L_{ij})^2 + \frac{\beta \eta}{2} \sum_j (D_j x_j - \omega_j)^2 \\ \text{s.t. } \hat{c}_{qj} &= \frac{g_{qj} x_j^k}{\sum_j g_{qj} x_j^k + r_q + s_q} y_i, \end{aligned} \quad (13)$$

where g_{qj} denotes the qj -th entry of system matrix G ; B_{ij} denotes the j -th entry of operative matrix B ; D_j represents the discrete differential operation at position j . It is worth mentioning that we adopt expectation maximization (EM) in our algorithm by introducing c_{qj} , whose expectation $\hat{c}_{qj}$ represents the possibility of the j -th voxel detected by the q -th detector pair.

By differentiating $\Omega(x; x^k)$ with respect to x , we get a quadric function:

$$\frac{\partial \Omega(x; x^k)}{\partial x_j} = 0 \Leftrightarrow A_j x_j + B_j + C_j \frac{1}{x_j} = 0, \quad (14)$$

where the solution x_j^{k+1} is the positive root of (14):

$$x_j^{k+1} = \frac{-B_j + \sqrt{B_j^2 - 4A_j C_j}}{2A_j}, \quad A_j = 2\alpha \sum_i B_{ij} + \beta \eta D_j$$

$$B_j = \sum_q g_{qj} - 2\alpha \sum_i L_{ij} - \beta \eta \omega_j, \quad C_j = -\sum_q \hat{c}_{qj}^k \quad (15)$$

Note that since the exact solution of x is no available, we conduct EM step iteratively until approaching a inner convergence within x -subproblem.

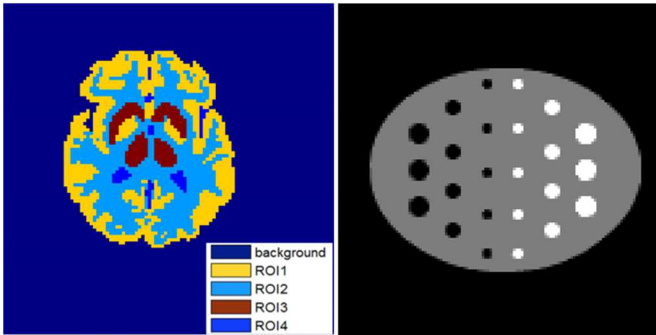
III. EXPERIMENTS

In this section, we conduct groups of experiments in order to validate the merits of our proposed method. The data we adopt are mainly based on the Monte Carlo simulation, which is widely-used in research of emission tomography. The simulation data consist of a set of Zubal brain phantoms aimed at testing the reconstruction accuracy and robustness, and lesion phantom aimed at evaluating the detection property. Furthermore, analysis respect to denoising and structural enhancement are conducted based on the Monte Carlo simulated data. Also, we employ the real PET thorax data obtained from CTI ECAT EXACT System provided by Fessler's group, in order to validate the proposed method in real case.

For quantitative accuracy evaluation, we employed the relative bias as well as relative variance, which are defined as:

$$bias = \frac{1}{N} \sum_j |x_j - \hat{x}_j| / \hat{x}_j, \quad variance = \frac{1}{N-1} \sum_j (|x_j - \bar{x}| / \hat{x}_j)^2 \quad [20],$$

where N denotes the pixel number in the tested region; $\bar{x}$ denotes the mean value within this region; $\hat{x}_j$ denotes the value of ground truth at position j . Given the fact that the emission tomography images are of piece-wise constancy, the better the reconstruction outcome is, the smaller the bias and variance are within each region. Besides, the peak signal to noise ratio (PSNR) is introduced in validating the convergent performance. Also, the qualitative evaluation is visually presented by contrasting the ground truth image with images reconstructed by different algorithms, including the famous maximum likelihood-expectation maximization (ML-EM) algorithm [19], penalized likelihood incremental optimization (denoted as PL-IO) method[20], and total variation minimization by augmented Lagrangian (TV-AL) method [21].



(a) Zubal brain phantom (b) Lesion phantom
Fig.1. Ground truth phantom

TABLE I
RELATIVE BIAS AND VARIANCE OF RECONSTRUCTED BRAIN PHANTOM UNDER DIFFERENT COUNTING RATES

Algorithm	Bias			Variance		
	1×10^6	5×10^6	2×10^7	1×10^6	5×10^6	2×10^7
ML-EM	0.2870	0.2199	0.1984	0.1517	0.0773	0.0572
TV-AL	0.2182	0.1323	0.0771	0.1106	0.0524	0.0229
PL-IO	0.2706	0.1403	0.0826	0.1109	0.0408	0.0263
PNLTV	0.1753	0.1086	0.0764	0.0917	0.0294	0.0149

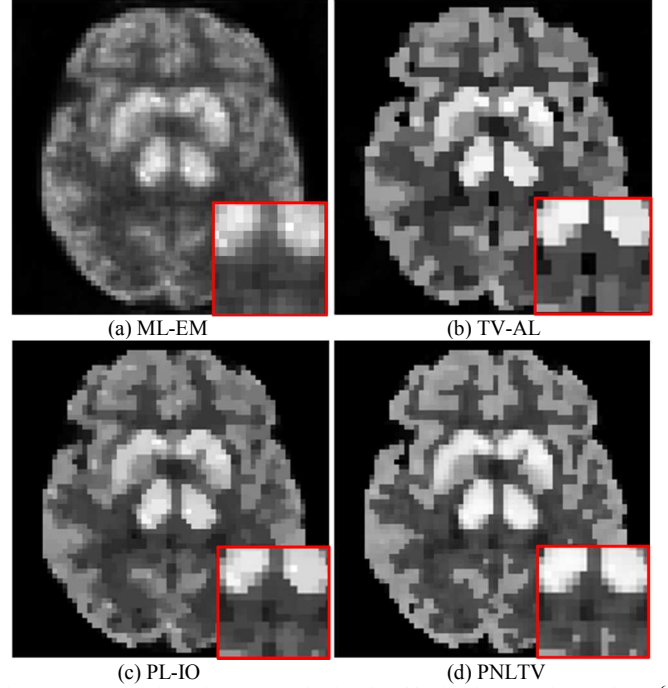


Fig. 2. Reconstruction of Monte Carlo simulated brain phantom data under 10^6 counting rate.

A. Zubal brain phantom

Initially, we use the Zubal brain phantom to examine the accuracy of the reconstructed image. As the ground truth shown in Fig.1(a), the image includes 4 regions of interest (ROI), and pixel shares same value within the same ROI according to the piece-wise constant property. On the other hand, the Monte Carlo simulation is based on Geant4 Application for Emission Tomography (GATE) platform. In this experiment, all the settings are simulated completely corresponding to real case, including the photon energy, scanner structure and so on. Data measured under different counting rates (1×10^6 , 5×10^6 , 2×10^7) are collected in the form of sinogram, and each one has 64×64 projections.

Table 1 demonstrates the relative bias and variance of the reconstructed results of different algorithms under diversified counting rates: 1×10^6 , 5×10^6 , 2×10^7 , and minimums of each column are in bold. According to the table, our proposed method PNLTV has substantially lower biases and variances than that of others, which means less differences compared with ground truth and lower level of noise. Besides, the

TABLE II
RELATIVE BIAS AND VARIANCE OF DIFFERENT ROI
UNDER 5×10^6 COUNTING RATE

Algorithm	Bias					Variance				
	Whole	ROI1	ROI2	ROI3	ROI4	Whole	ROI1	ROI2	ROI3	ROI4
ML-EM	0.2199	0.2514	0.1830	0.2348	0.3843	0.0773	0.0168	0.0399	0.0099	0.1936
TV-AL	0.1323	0.0916	0.1504	0.0734	0.5472	0.0524	0.0131	0.0551	0.0035	0.5386
PL-IO	0.1403	0.1187	0.1488	0.1033	0.3895	0.0408	0.0087	0.0442	0.0025	0.2532
PNLTV	0.1086	0.0879	0.1131	0.0877	0.3583	0.0294	0.0074	0.0325	0.0026	0.1586

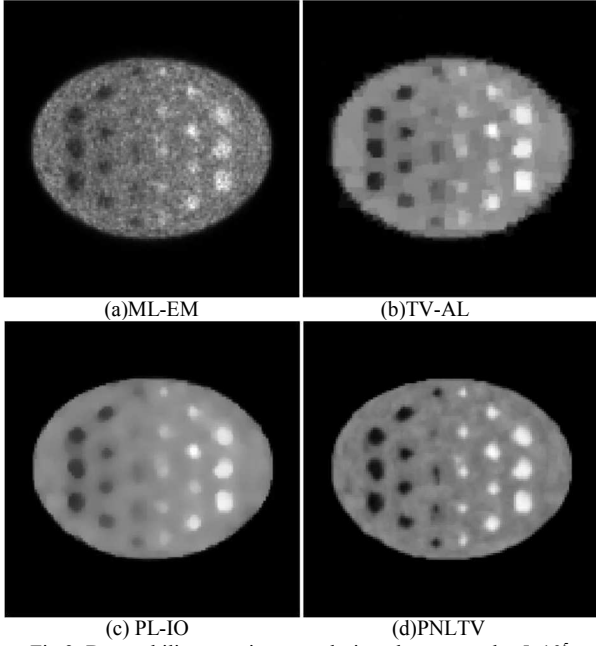


Fig.3. Detectability experiment on lesion phantom under 5×10^5 counting rate.

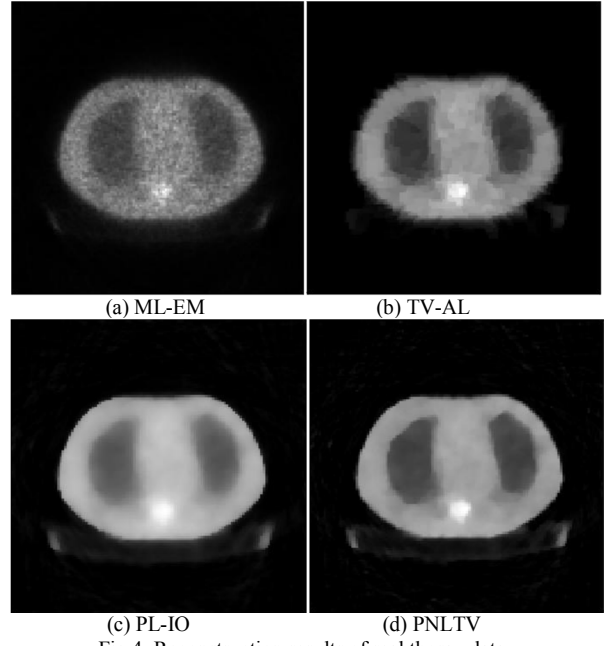


Fig.4. Reconstruction results of real thorax data.

robustness of PNLTV is validated under different counting rates.

To analyze each algorithm further, relative bias and variance of each ROI are collected under 5×10^6 counting rate, as shown in Table 2. Similarly, optimal statistical results are shown in bold. It can be seen that the results of proposed method are overall superior to its counterparts of other algorithms. Nevertheless, there is only one exception that the bias of TVAL in ROI3 is slightly lower than that of PNLTV, it is partially owing to TVAL's good edge preserving property.

Also, visual analyses are introduced in this trail of experiments. In order to simulate the low-dose cases, we adopt 1×10^6 counting rate data to present the reconstructed images, as shown in fig.2. According to this figure, image reconstructed by ML-EM suffers from noises severely; image reconstructed by TVAL has a stair-case effect; over-smoothness is spotted in image reconstructed by PL-IO. In contrast, our proposed method not only suppresses the Poisson noise excellently but also has an eminent performance in preserving structural information.

In addition, the convergence performance of mentioned algorithms is shown in Fig.5. As we can see, our PNLTV and

PL-IO have demonstrated faster convergence while PNLTV finally converged at the highest PSNR.

B. Lesion phantom

One major goal of emission tomography is to detect lesions or tumors, especially in PET and SPECT image analysis. In this regard, a 128×128 sized lesion phantom simulated by Monte Carlo is adopted in this experiment. The lesion phantom, as indicated in Fig.1.(b), has two groups of lesion regions, and each group has three sizes of circular lesion regions.

Lesion images with 5×10^5 counting rate reconstructed by different methods are present in Fig.3. In this study, different algorithms perform variedly. Without image prior, ML-EM fails to discern the smallest lesion regions from the background noise. Moreover, although TVAL and PL-IO both have acceptable results when recover the white regions, they fail to recover the smallest black lesion regions. In comparison, our method recognizes each of the regions and shows a better performance in detection.

C. Real patient data

The data set is provided by Prof. Fessler. It was collected in 160 radial samples (varies fastest) by 192 angular samples (over 180 degrees), obtained from CTI 921 ECAT EXACT

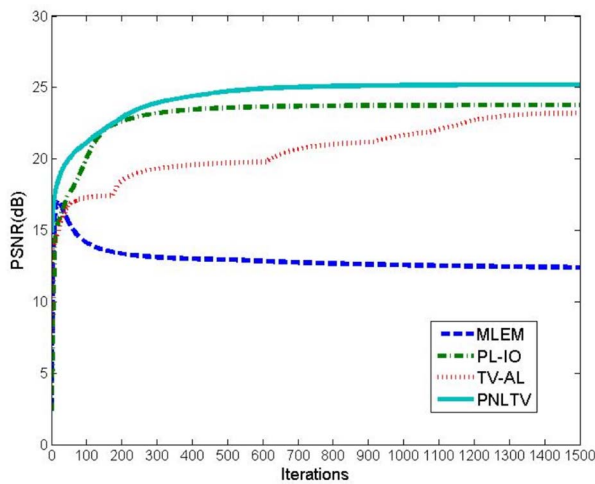


Fig.5 PSNR curves of the four algorithms along the iteration on the brain phantom.

Scanner. The sum of the photons is 920653. The reconstructed results are presented in Fig. 4. It can be witnessed that PNLTV performs better while images recovered by TVAL and PL-IO suffer from stair-case effect and over-smoothness respectively.

IV. CONCLUSION

In this paper, we have designed a novel algorithm for PET reconstruction. On one hand, by a low-rank approach, our method introduces the nonlocal self-similarities within the recovered image to itself, thus the structural information is enhanced while the sparsity is also strongly kept. On the other hand, total variation constraint is adopted in order to further depressed the noise and compensate the rectangular-edge effect inherited in the patch-based methods. These two penalty terms, which serve as the image prior, are firstly and jointly optimized in the Poisson reconstruction frame. The proposed method shows remarkable performance in resolution improvement, denoising, lesion detection and convergence. Moreover, since the structural information is acquired from the recovered image itself during the iterations, the method is of outstanding adaptability as well as robustness.

ACKNOWLEDGMENT

The authors would like to thank Prof. Fessler for sharing code resources and the real PET data on his website. This work is supported in part by the National Natural Science Foundation of China (No: 61525106, 61427807), by Shenzhen Innovation Funding (No: JCYJ20170818164343304, JCYJ20170816172431715), by National Key Technology Research and Development Program of China (No: 2017YFE0104000, 2016YFC1300302) and by the Division of Mathematical Science: NSF/DMS-1719932.

REFERENCES

[1] Rockmore A J, Macovski A. A maximum likelihood approach to emission image reconstruction from projections[J]. IEEE transactions on nuclear science, 1976, 23(4): 1428-1432.

[2] Shepp L A, Vardi Y. Maximum likelihood reconstruction for emission tomography[J]. IEEE transactions on medical imaging, 1982, 1(2): 113-122.

[3] Qi J, Leahy R M. Iterative reconstruction techniques in emission computed tomography[J]. Physics in medicine and biology, 2006, 51(15): R541.

[4] Levitan E, Herman G T. A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography[J]. IEEE Transactions on Medical Imaging, 1987, 6(3): 185-192.

[5] Fessler J A, Rogers W L. Spatial resolution properties of penalized-likelihood image reconstruction: space-invariant tomographs[J]. IEEE Transactions on Image processing, 1996, 5(9): 1346-1358.

[6] J. A. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *Medical Imaging, IEEE Transactions on*, vol. 13, no. 2, pp. 290-300, 1994.

[7] Huesman R H, Klein G J, Moses W W, et al. List-mode maximum-likelihood reconstruction applied to positron emission mammography (PEM) with irregular sampling[J]. IEEE transactions on medical imaging, 2000, 19(5): 532-537.

[8] Geman S, Geman D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images[J]. Journal of Applied Statistics, 1993, 20(5-6): 25-62.

[9] Chambolle A, Lions P L. Image recovery via total variation minimization and related problems[J]. Numerische Mathematik, 1997, 76(2): 167-188.

[10] Xu Q, Yu H, Mou X, et al. Low-dose X-ray CT reconstruction via dictionary learning[J]. IEEE Transactions on Medical Imaging, 2012, 31(9): 1682-1697.

[11] Mairal J, Bach F, Ponce J, et al. Non-local sparse models for image restoration[C]//Computer Vision, 2009 IEEE 12th International Conference on. IEEE, 2009: 2272-2279.

[12] Buades A, Coll B, Morel J M. A review of image denoising algorithms, with a new one[J]. Multiscale Modeling & Simulation, 2005, 4(2): 490-530.

[13] Dabov K, Foi A, Katkovnik V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering[J]. IEEE Transactions on image processing, 2007, 16(8): 2080-2095.

[14] Dong W, Li X, Zhang L, et al. Sparsity-based image denoising via dictionary learning and structural clustering[C]//Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011: 457-464.

[15] Dong W, Shi G, Li X, et al. Compressive sensing via nonlocal low-rank regularization[J]. IEEE Transactions on Image Processing, 2014, 23(8): 3618-3632.

[16] Rudin L I, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms[J]. Physica D: Nonlinear Phenomena, 1992, 60(1-4): 259-268.

[17] Cai J F, Candès E J, Shen Z. A singular value thresholding algorithm for matrix completion[J]. SIAM Journal on Optimization, 2010, 20(4): 1956-1982.

[18] Chen S, Liu H, Shi P, et al. Sparse representation and dictionary learning penalized image reconstruction for positron emission tomography[J]. Physics in medicine and biology, 2015, 60(2): 807.

[19] Shepp L A, Vardi Y. Maximum likelihood reconstruction for emission tomography[J]. IEEE transactions on medical imaging, 1982, 1(2): 113-122.

[20] Ahn S, Fessler J A, Blatt D, et al. Incremental optimization transfer algorithms: Application to transmission tomography[C]//Nuclear Science Symposium Conference Record, 2004 IEEE. IEEE, 2004, 5: 2835-2839.

[21] Li C, Yin W, Jiang H, et al. An efficient augmented Lagrangian method with applications to total variation minimization[J]. Computational Optimization and Applications, 2013, 56(3): 507-530.