

## Sparse representation and dictionary learning penalized image reconstruction for positron emission tomography

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2015 Phys. Med. Biol. 60 807

(<http://iopscience.iop.org/0031-9155/60/2/807>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.21.175.86

This content was downloaded on 29/02/2016 at 19:46

Please note that [terms and conditions apply](#).

# Sparse representation and dictionary learning penalized image reconstruction for positron emission tomography

Shuhang Chen<sup>1</sup>, Huafeng Liu<sup>1,4</sup>, Pengcheng Shi<sup>2</sup> and Yunmei Chen<sup>3</sup>

<sup>1</sup> State Key Laboratory of Modern Optical Instrumentation, Department of Optical Engineering, Zhejiang University, Hangzhou 310027, People's Republic of China

<sup>2</sup> B Thomas Golisano College of Computing and Information Sciences, Rochester Institute of Technology, Rochester, NY 14623, USA

<sup>3</sup> Department of Mathematics, University of Florida, 458 Little Hall Gainesville, FL 32611-8105, USA

E-mail: [liuhf@zju.edu.cn](mailto:liuhf@zju.edu.cn)

Received 19 July 2014, revised 29 October 2014

Accepted for publication 26 November 2014

Published 7 January 2015



CrossMark

## Abstract

Accurate and robust reconstruction of the radioactivity concentration is of great importance in positron emission tomography (PET) imaging. Given the Poisson nature of photo-counting measurements, we present a reconstruction framework that integrates sparsity penalty on a dictionary into a maximum likelihood estimator. Patch-sparsity on a dictionary provides the regularization for our effort, and iterative procedures are used to solve the maximum likelihood function formulated on Poisson statistics. Specifically, in our formulation, a dictionary could be trained on CT images, to provide intrinsic anatomical structures for the reconstructed images, or adaptively learned from the noisy measurements of PET. Accuracy of the strategy with very promising application results from Monte-Carlo simulations, and real data are demonstrated.

Keywords: positron emission tomography, image reconstruction, sparse representation

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

## 1. Introduction

Positron emission tomography (PET) uses the idea of injecting chemical compounds labeled with positron-emitting isotopes into a body to measure their spatial and temporal distribution

<sup>4</sup> Author to whom any correspondence should be addressed

externally. However, estimating the radioactivity concentration from measured data is a challenging problem and has been the focus of research of numerous researchers over the past several years.

The first attempt at reconstructing the emission activities occurred with the filtered backprojection (FBP) method using Radon transform. However, FBP method disregards the spatially-variant system response of PET, and statistical noise is neglected and often treated in a *post-hoc* manner. Hence the accuracy of FBP method is severely limited.

On the other hand, assuming data is Poisson distributed, iterative statistical reconstruction algorithms are able to model the physical detection process, and thus have been the primary focus of many recent efforts. However, from the perspective of statistical inference, such high dimensional maximum likelihood estimators are inevitably ill-conditioning, i.e. have the ‘checkerboard effect’ which is known for maximum likelihood-expectation maximization (ML-EM) in the PET community (Shepp and Vardi 1982). Common approaches to that are to add penalty constraints and compute the corresponding penalized ML solution, or the so called *maximum a posteriori* (MAP) estimation (Levitan and Herman 1987, Chang *et al* 2004). Typical penalty terms are quadratic smoothness, logcosh penalization, total variation and Huber potential function (Chang *et al* 2004, Qi and Leahy 2006). The main concern in designing these penalty functions is balancing between smoothing uniform regions and preserving edge sharpness. However, due to the low SNR (signal to noise ratio) nature of PET measurement, controlling such trade-off between smoothness and image details has been so far proven technically difficult.

Though some tissues are heterogeneous, in many situations, we can fairly assume that within patches of certain sizes a PET image can be considered as piece-wise constant. With this assumption in mind, intuitively, if we already knew the position of the boundary, where sharp variation is allowed and even encouraged, the rest of the image should readily be as smooth as possible. This idea forms the basis of these reconstruction methods that incorporate anatomical priors (Gindi *et al* 1993, Bouman and Sauer 1996, Tang *et al* 2008, Dewaraja *et al* 2010, Vunckx *et al* 2012). Since different anatomical structures have different physiological functions, we can expect to see differences in tracer uptake between anatomical structures. This general observation has also been borne out in high-resolution autoradiographic images in which functional images also clearly reveal the morphology of the underlying structures (Gindi *et al* 1993). While the anatomical modalities have superior resolution to PET, accurate estimates of anatomical boundaries can be formed from them to influence the PET reconstruction (Dewaraja *et al* 2010). However such anatomy based PET reconstruction methods rely heavily on the accuracy of image registration of PET and anatomical images like computed tomography (CT) and magnetic resonance imaging (MRI). Moreover, in some studies, the findings of anatomical and functional imaging modalities may disagree, and in some circumstances anatomical imaging modalities might miss some regions that appear suspicious on PET images.

This problem motivated us to decompose the boundary structures into local features. By decomposing, we mean to learn its intrinsic local structures such that anatomical image can be represented efficiently through the learned features. This idea utilizes the recently developed theory of sparse representation and dictionary learning. Aharon *et al* learns an over-complete dictionary from training set by K-SVD algorithm, and the elements of the dictionary are local image patches (Aharon *et al* 2006). They show that images with similar structures can then be sparsely represented by the learned elements. Moreover, such a sparse representation approach can effectively eliminate random noises. Given the fact that PET and anatomical images share similar boundary structures, we can effectively

represent the structures in PET images with the local features learned from anatomical images. Because we are dealing with small image patches that contain local edge elements, patches from PET can be reconstructed with reasonable accuracy by an appropriate linear combination of learned edge elements. For smooth regions, one patch is sufficient to represent them.

Several efforts related to sparse representation and dictionary learning have been made in the medical imaging community, Chen first demonstrated the effectiveness of sparse representation in MR image reconstruction using a pre-trained MR brain dictionary (Chen *et al* 2010). Ravishankar proposed to adapt the pre-trained dictionary to the input to reconstruct the MR images (Ravishankar and Bresler 2011). Xu also developed a dictionary learning based algorithm to reconstruct low-dose CT images (Xu *et al* 2012). However, the penalized weighted least-square function they used for data consistence have been proven negatively-biased (Fessler 1994). Furthermore, the validation, which trained the dictionary based on the ground truth or high-dose image to reconstruct the exactly corresponding low-dose image, is doubtful (Xu *et al* 2012).

In this paper, a novel algorithm that combines a maximum likelihood function and sparsity penalty is proposed. The resulting model is capable of representing the measured data with Poisson statistics, while the sparsity penalty term in the objective function encourages the reconstructed image patches being sparsely represented by the dictionary. An iterative procedure is then provided to optimize the resulting objective function. To demonstrate the applicability of the novel method, we have tested the proposed algorithm in terms of reconstruction accuracy and detectability based on Monte-Carlo generated data and real patient data.

## 2. Method

### 2.1. Problem formulation

**2.1.1. Likelihood estimate for data fidelity measure.** In the measurement process of PET, a coincidence event indicates that two gamma rays interact with the opposite detector pairs within a small coincidence timing window. During PET emission scans, one wishes to include only *true coincidences* that are related to gamma rays from the same annihilation and that have not scattered in the body prior to detection. However, in reality, the coincidence events include those records that originate from two unrelated positions, and those records that the annihilation photons have Compton scattering and lose their directions. The former is called *random coincidences* (RC) and the latter is called *scattered coincidences* (SC). RC events are a primary source of background noise in PET.

PET measured data, called sinogram  $y$ , is the sum of such events at each detector bin along the scanner ring, i.e.  $y = \{y_i, i = 1, \dots, M\}$ , where  $M$  is the number of detector bins. Given that detecting each decay event by the system is independent and can be modeled as a Bernoulli process, the sinogram data  $y$  can be assumed as a collection of independent Poisson random variables. Also, the measurement  $y$  is related to the unknown activity map  $x$  by the projective transform.

$$y \sim \text{Poisson} \{\bar{y}\} \quad \text{s. t.} \quad \bar{y} = Gx + r + s \quad (1)$$

where  $G$  is the system matrix representing the average probabilities of detecting an emission from voxel site  $j$  at detector  $i$ , which is mainly determined by physical property of PET

scanner;  $r$  and  $s$  are the RC and SC events respectively. Based on the independent Poisson assumption, we can write the likelihood function of  $y$  as

$$\Pr(y|x) = \prod_i^M e^{-\bar{y}_i} \frac{\bar{y}_i^{y_i}}{y_i!} \quad (2)$$

Instead of maximizing the above likelihood function, we often minimize its negative log-likelihood function to estimate  $x$  since the log of likelihood function is easier to deal with:

$$\min_x L(x) = \min_x -\log(\Pr(y|x)) = \min_x \sum_i^M \bar{y}_i - y_i \log(\bar{y}_i) \quad \text{s.t.} \quad \bar{y} = Gx + r + s \quad (3)$$

The constant term  $\log(y_i!)$  has been neglected. For penalized maximum likelihood reconstruction algorithms, one seeks to estimate of the image  $x$  that minimizes an objective function consisting of likelihood function and a penalty term.

$$\min_x \lambda L(x) + R(x) \quad \text{s.t.} \quad \bar{y} = Gx + r + s \quad (4)$$

where  $\lambda$  is weighting parameter and  $R(x)$  is a prior penalty term.

**2.1.2. Sparsity penalty on dictionary.** Numerous penalized methods have been proposed for PET reconstruction like quadratic penalization, Gibbs smoothing, Huber potential, etc (Qi and Leahy 2006). In this paper, we propose to regularize the likelihood function by encouraging sparsity with respect to dictionary. More specifically, we want each patch of the reconstructed image to be sparsely represented by the elements of the dictionary, which could be pre-trained on database or adaptive learned from the noisy measurements.

An image patch of size  $\sqrt{n} \times \sqrt{n}$  can be expressed as a  $n$  dimensional vector  $p \in \mathbb{R}^n$ . Given an emission image vector  $x \in \mathbb{R}^N$ , we decompose it into  $S$  overlapping patches  $p_s \in \mathbb{R}^n, s = 1, \dots, S$ , by matrix operator  $E_s$ , i.e.  $p_s = E_s x$ . The sparse penalty of  $x$  with respect to dictionary  $D$  is as follows (Aharon *et al* 2006)

$$\min_{\alpha} R_{\text{sparse}}(x, \alpha) = \min_{\alpha} \left( \sum_s^S \|E_s x - D\alpha_s\|_2^2 + \mu \|\alpha_s\|_0 \right) \quad (5)$$

where  $\alpha_s$  is the sparse coefficient for  $s$ th patch  $p_s, p_s = E_s x$ ,  $\|\alpha_s\|_0$  denotes the number of non-zero entries of  $\alpha_s$ , and  $\mu$  is weighting parameter of sparsity. Taking sparsity constraint (5) as penalty term for (4), we propose to minimize the following objective function to reconstruct PET images.

$$\min_{x, \alpha} \Psi(x, \alpha) = \lambda L(x) + R_{\text{sparse}}(x, \alpha) \quad \text{s.t.} \quad \bar{y} = Gx + r + s \quad (6)$$

where  $L(x)$  is likelihood function,  $R_{\text{sparse}}(x, \alpha)$  is sparsity constraint and  $\lambda$  is the weighting parameter.

## 2.2. Algorithms

The optimization algorithm alternates between minimizing the objective function (6) with respect to each variable  $x, D, \alpha$  with other variables fixed. We first demonstrate how to solve the  $x$ -subproblem by rewriting objective function (6) with the items that contain  $x$ :

$$\begin{aligned}
 x = \arg \min_x \Psi_x(x, \alpha) &= \lambda \sum_i^M \bar{y}_i - y_i \log(\bar{y}_i) + \sum_s^S \|E_s x - D\alpha_s\|_2^2 \\
 \text{s. t. } \bar{y} &= Gx + r + s
 \end{aligned} \tag{7}$$

We note that (7) is a typical MAP reconstruction problem with quadratic penalty (Levitan and Herman 1987). Various techniques have been developed to optimize it. Here we use the efficient expectation maximization (EM) based algorithm to minimize (7).

We now explain in detail the expectation maximization based scheme that we used for the  $x$  subproblem. Following the hidden variable formulation (Shepp and Vardi 1982), the  $x$  subproblem (7) is equivalent to minimize the following function  $\Psi_x(c_{ij}, x, \alpha)$

$$x = \arg \min_x \Psi_x(c_{ij}, x, \alpha) = \lambda \sum_j^N \sum_i^M (g_{ij}x_j - c_{ij} \log(g_{ij}x_j)) + \sum_s^S \|E_s x - D\alpha_s\|_2^2 \tag{8}$$

where  $c_{ij}$  is the hidden variable, representing the number of photons emitted from image pixel  $j$  and detected at line of response (LOR)  $i$ , and  $g_{ij}$  is the  $ij$ th entry of system matrix  $G$ . This new objective function is derived based on the independent Poisson statistical assumption of  $c_{ij}$ . The hidden variable is formulated here so that if they were observed, the maximum likelihood based function in (8) is readily solvable. Therefore, EM algorithm proceeds by two steps:

- E-step : Estimate the hidden variable  $c_{ij}$  given current estimate  $x^k$  in the  $k$ th iteration and sinogram  $y$ , i.e.  $\hat{c}_{ij} = \xi(c_{ij}|x^k, y)$ , where  $\xi$  denotes conditional expectation. Plugging in the estimated  $\hat{c}_{ij}$  into  $\Psi_x(c_{ij}, x, \alpha)$ , we have an intermediate function  $\hat{\Psi}_x(\hat{c}_{ij}, x, \alpha)$ .
- M-step: Minimize the intermediate function  $\hat{\Psi}_x(\hat{c}_{ij}, x, \alpha)$  by zeroing its derivative with respect to  $x$ .

For E-step, the expectation of  $c_{ij}$  conditioned on sinogram  $y_i$  and current estimate of  $x_j$  is given by (Shepp and Vardi 1982)

$$\hat{c}_{ij} = \xi(c_{ij}|y_i, x_j^k) = \frac{g_{ij}x_j^k}{\sum_j^N g_{ij}x_j^k + \hat{r}_i + \hat{s}_i} y_i \tag{9}$$

$\hat{r}_i, \hat{s}_i$  are the estimated random events and scatter events respectively (Qi and Leahy 2006). The resulting function  $\hat{\Psi}_x(\hat{c}_{ij}, x, \alpha)$  is

$$\hat{\Psi}_x(\hat{c}_{ij}, x, \alpha) = \lambda \sum_j^N \sum_i^M (g_{ij}x_j - \hat{c}_{ij} \log(g_{ij}x_j)) + \sum_s^S \|E_s x - D\alpha_s\|_2^2 \tag{10}$$

However, when minimizing the function  $\hat{\Psi}_x(\hat{c}_{ij}, x, \alpha)$ , we note that the second term  $\sum_s^S \|E_s x - D\alpha_s\|_2^2$  is not separable and direct solution by zeroing the derivative of (10) with respect to  $x_j$  is generally not possible. Instead, an approximate solution like iterated coordinate decent can be used here but it converges slowly unless given a good initial estimate of  $x$  (Bouman and Sauer 1996). Another approach is the one-step-late method, in which the partial derivative of  $\sum_s^S \|E_s x - D\alpha_s\|_2^2$  is evaluated at the current estimate  $x_j^k$  (Green 1990). In general, however, this method does not converge. Here, we adopt De Pierro’s convexity trick (De Pierro 1995) to replace the penalty term with its convex separable surrogate function. We first rewrite  $\|E_s x - D\alpha_s\|_2^2$  and  $[E_s x]_l$  as follows:

$$\|E_s x - D\alpha_s\|_2^2 = \sum_l^n ([E_s x]_l - [D\alpha_s]_l)^2 \tag{11}$$

$$[E_s x]_l = \sum_j^N \beta_{s,lj} \left( \frac{e_{s,lj}}{\beta_{s,lj}} (x_j - x_j^k) + [E_s x^k]_l \right) \quad \text{s. t.} \quad \sum_{j=1}^N \beta_{s,lj} = 1, \forall j, \beta_{s,lj} \geq 0 \tag{12}$$

$[E_s x]_l$  denotes the  $l$ -th entry of the vector  $E_s x$ ,  $e_{s,lj}$  is the  $lj$ -th entry of matrix  $E_s$ ,  $x_j^k$  is current estimate of the image vector  $x$  at pixel  $j$ . Since  $([E_s x]_l - [D\alpha_s]_l)^2$  is convex with respect to  $[E_s x]_l$ , we have

$$([E_s x]_l - [D\alpha_s]_l)^2 \leq \sum_j^N \beta_{s,lj} \left( \frac{e_{s,lj}}{\beta_{s,lj}} (x_j - x_j^k) + [E_s x^k]_l - [D\alpha_s]_l \right)^2 \tag{13}$$

Applying this inequality to equation (10) yields its convex separable surrogate function  $\phi(x; x^k)$ :

$$\begin{aligned} \phi(x; x^k) = & \lambda \sum_j^N \sum_i^M (g_{ij} x_j - \hat{c}_{ij} \log(g_{ij} x_j)) + \sum_s^S \sum_l^n \sum_j^N \beta_{s,lj} \left( \frac{e_{s,lj}}{\beta_{s,lj}} (x_j - x_j^k) \right. \\ & \left. + [E_s x^k]_l - [D\alpha_s]_l \right)^2 \end{aligned} \tag{14}$$

We choose  $\beta_{s,lj}$  to be  $\beta_{s,lj} = e_{s,lj} / \sum_{j=1}^N e_{s,lj}$  and minimize (14) by differentiating  $\phi(x; x^k)$  with respect to  $x$ .

$$\begin{aligned} & \frac{\partial \phi(x; x^k)}{\partial x_j} \\ = & \lambda \sum_i^M g_{ij} - \lambda \sum_i^M \hat{c}_{ij} \frac{1}{x_j} + \sum_s^S \sum_l^n 2e_{s,lj} \left( \frac{e_{s,lj}}{\beta_{s,lj}} x_j + [E_s x^k]_l - [D\alpha_s]_l - \frac{e_{s,lj}}{\beta_{s,lj}} x_j^k \right) = 0 \end{aligned} \tag{15}$$

Then we can find that  $x_j^{k+1}$  is the root of following second order polynomial equation

$$\begin{aligned} A_j x_j + B_j + C_j \frac{1}{x_j} = 0, \quad A_j = \sum_s^S \sum_l^n 2e_{s,lj} \sum_j^N e_{s,lj}, \quad C_j = -\lambda \sum_i^M \hat{c}_{ij} \\ B_j = \left( \sum_s^S \sum_l^n 2e_{s,lj} ([E_s x^k]_l - [D\alpha_s]_l) \right) - A_j x_j^k + \lambda \sum_i^M g_{ij} \end{aligned} \tag{16}$$

Because  $\phi(x; x^k)$  of (14) is a strictly convex problem (14),  $x_j^{k+1}$  is the only strictly positive solution of equation (16) (detailed proof is similar to De Pierro (1995)).

$$x_j^{k+1} = \frac{-B_j + \sqrt{B_j^2 - 4A_j C_j}}{2A_j} \tag{17}$$

Note that because there is no exact solution for  $x$ -subproblem, we iteratively take E-step and M-step until inner convergence to get an approximate solution. The alternating direction minimization method is expected to minimize each subproblem alternately. Therefore, such an inner converged solution is sufficient for achieving the minimum of  $\Psi(x, \alpha)$ .

**Algorithm 1.** Iteration procedures of proposed optimization algorithm.

**Require:** Sinogram  $y$  and system matrix  $G$ , weighting parameter  $\lambda$ .

1: **Initialize**  $\alpha = 0$ ,  $x = \text{FBP}(y)$ .

2: **repeat**

3:   **repeat** ▷Solving  $x$  subproblem

4:     E-step compute  $\widehat{\Psi}(\hat{c}_{ij}|x^k, y)$  using (9).

5:     M-step: update  $x_j^k, \forall j$  using (17).

6:   **until** Relative change of  $x < 10^{-3}$

7:    $D = K - \text{SVD}(x)$  ▷If we use adaptive learned dictionary

8:    $\alpha = \text{Sparse coding using OMP:}$

$$\forall_s \quad \alpha_s = \arg \min_{\alpha_s} \|\alpha_s\|_0 \quad \text{s. t.} \quad \|E_s x - D\alpha_s\|_2^2 \leq \epsilon \quad (18)$$

9:      $k \leftarrow k + 1$

10: **until** Relative change of  $x < 10^{-4}$ .

11: **Return**  $x^k$

The  $\alpha$ -subproblem is solved by orthogonal matching pursuit (OMP) (Tropp and Gilbert 2007). If we use a global dictionary, the dictionary  $D$  is fixed in the algorithm after pre-trained on CT images using K-SVD (Aharon *et al* 2006). For an adaptive dictionary, it is learned from the current estimate  $x^k$  in each iteration step using K-SVD algorithm. The whole iteration scheme stops when the relative change of  $x$  is smaller than the pre-specified tolerance. We have summarized the optimization procedures in algorithm 1.

### 2.3. Implementation

**2.3.1. Tuning parameters.** There are two parameters in the presented algorithm which needed to be manually tuned. The first one is  $\lambda$ , weighting parameter before likelihood function in equation (6). It depends on noise level, and it should be strengthened when SNR increases and *vice versa*. It is intuitive that data fidelity should be stressed when measurements are ‘clean’, while sparsity regularization comes into play when measurements cannot be trusted due to high level noise corruptions. Generally, it is hard to find simple methods to determine the appropriate value for  $\lambda$ . In all of our experiments, we empirically found it works well nearby 3. Another parameter is error tolerance  $\epsilon$  in the OMP sparse coding stage of equation (18) in algorithm 1. This value is related to patch size and noise level. In (Elad and Aharon 2006), the authors empirically choose it to be  $1.15 \cdot \sigma$ , where  $\sigma$  is standard variation of Gaussian noise. While PET measurements are assumed to be Poisson distributed, and as we also do not priorly know the statistics of noises, we empirically tune it within [0.05, 0.105].

**2.3.2. Dictionary.** Both global and adaptive dictionaries are trained from the extracted image patches by the K-SVD algorithm (Aharon *et al* 2006). In the literature, the size of the dictionary element ranges from  $6 \times 6$  to  $8 \times 8$ . A large element size may fail to capture fine local structures, while more computational time would be needed if we chose a small size. In our case, we find  $7 \times 7$  works well in all of our experiments. The number of dictionary elements is a trade-off between over-completeness and computational cost. Provided that PET images have a relative simple image structure and a strong similarity of the structures between the to-be-reconstructed PET image and the images whose patches are used for training the dictionary, we find 144 is enough in experiments. The global dictionary (GD) is pre-trained from eight

slices of thorax CT images, one of which is shown in figure 1(c). This learned dictionary would be used throughout the experiments as a GD, as shown in figure 1(b). The adaptive dictionary (AD) is learned from its corresponding measured data, as we explained in the previous section.

### 3. Experiments

To validate the proposed algorithm, we conduct two groups of thorax experiments based on: Monte Carlo simulated data and the real patient thorax data. Simulation studies include experiments based on Zubal thorax phantom to evaluate the reconstruction accuracy and the lesion phantom to test detectability of reconstruction algorithms. The real patient data include the data obtained from CTI ECAT PET scanner provided by Prof Fessler<sup>4</sup> and one data set of lung cancer scanned by SHR-22000. Finally, we present the experiments in terms of the two tuning parameters and dictionary settings.

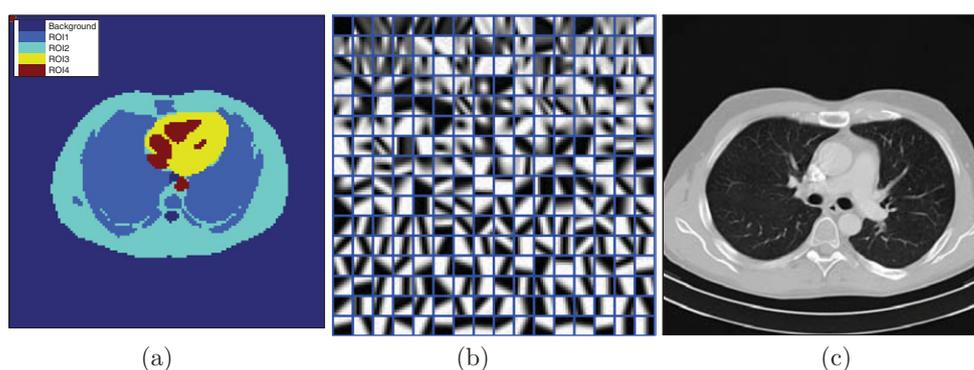
Reconstructions are evaluated by regional relative bias and relative variance, which are defined as  $\text{bias} = \frac{1}{N} \sum_j |x_j - \hat{x}_j|/\hat{x}_j$ ,  $\text{variance} = \frac{1}{N-1} \sum_j \left( \frac{x_j - \hat{x}_j}{\hat{x}_j} \right)^2$ .  $\hat{x}_j$  denotes  $j$ -th pixel value of ground truth.  $N$  is the overall number of pixels in the examined region. Since each region does not necessarily share the same number of pixels, the overall bias and variance may not add up to the whole bias and variance. The absolute difference between the reconstructed image and the ground true image is presented for visual judgments. As for quantitative evaluation of detectability, we use Jaccard index (JI) to evaluate the accuracy of lesion detection. The Jaccard index, defined as  $\text{JI} = \frac{A \cap B}{A \cup B}$ , measures the overlap of two regions, the true lesion region  $A$  and the recovered lesion  $B$ . The proposed methods with GD and AD were compared with the classical maximum likelihood-expectation maximization (ML-EM) algorithm and the state-of-the-art method separable paraboloidal surrogates with ordered-subsets (SPS-OS) regularized by the Huber potential function (Fessler and Erdogan 1998, Mehranian *et al* 2013).

#### 3.1. Monte Carlo simulations

**3.1.1. Zubal thorax phantom.** We first use the Monte Carlo simulated Zubal thorax phantom to validate the proposed algorithm. The Monte Carlo simulation is able to produce realistic sinogram data by simulating the imaging physical process. More importantly, it offers the opportunity to evaluate the reconstruction algorithms quantitatively against the ground truth. In this study, the simulated PET scanner was Hamamatsu SHR-22000 and all the settings were as same as the real situation, including dead time, energy resolution, etc. The Zubal thorax phantom, as indicated in figure 1(a), has four regions of interest. Sinogram had  $128 \times 128$  projections. Two sets of measured data under different counting rates ( $1 \times 10^6$  and  $5 \times 10^5$ ) were tested.

Figure 2 demonstrates the reconstruction results upon the Monte Carlo simulated data set. Due to Huber potential regularization, SPS-OS method produces a slightly smoother image than ML-EM method and the edges are also better preserved. On the other hand, by encouraging sparsity, the GD method could suppress much more noise and image details are also well recovered. In the third row of figure 2(d), there exists many artifacts in the AD reconstructed images. From the subtraction images in figure 2 (the second and fourth rows), we find that the

<sup>4</sup> <http://web.eecs.umich.edu/fessler/result/et/pet.emis/>



**Figure 1.** Zubal thorax phantom (left), trained global dictionary (middle), and one of eight CT thorax slices for training GD (right). (a) Zubal phantom. (b) Global dictionary. (c) One CT thorax slice.

most errors concentrate on the edges and region of interest (ROI) 4 for all four algorithms. Tables 1 and 2 further confirm this evidence that the proposed methods give better bias and variance performance in each ROI under both counting rates, especially within ROI4.

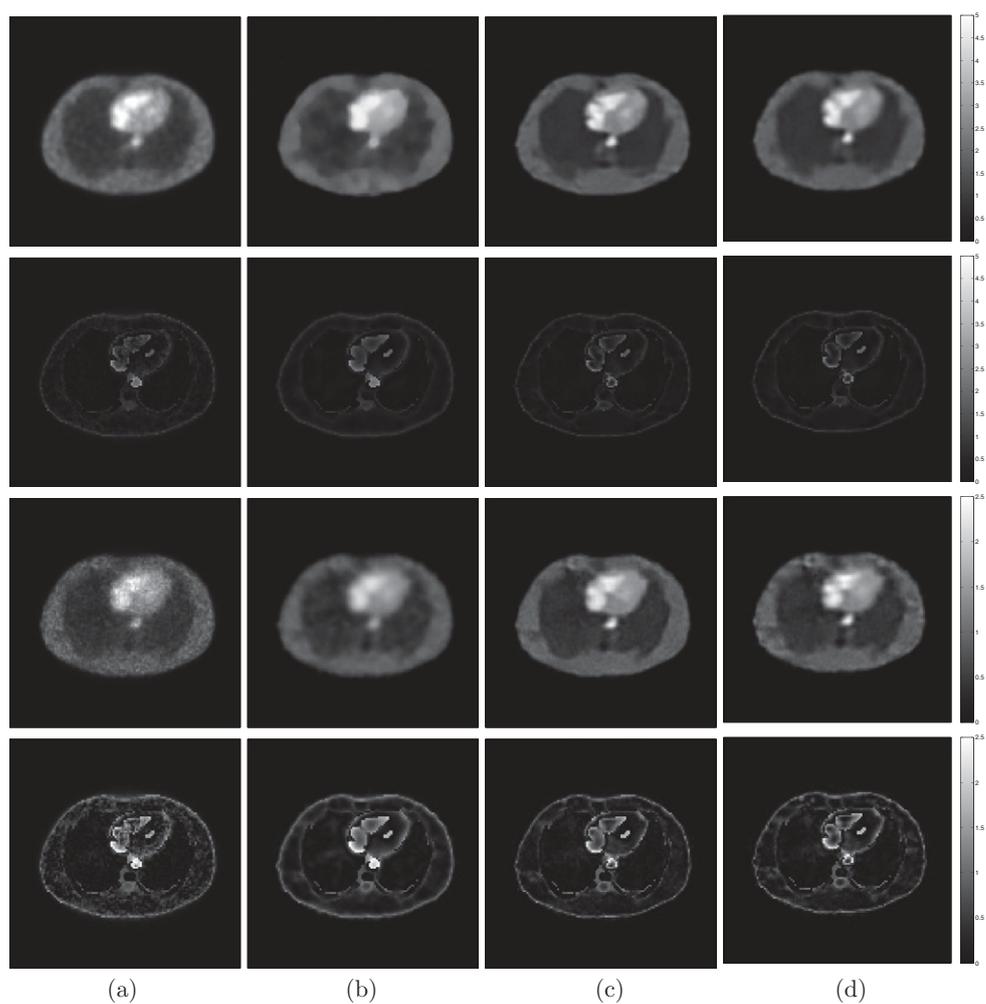
**3.1.2. Lesion detection.** In this study, we demonstrate the experiments concerning the detectability of proposed algorithms. As in the first row of figure 3(a), we design a simple thorax phantom ( $128 \times 128$ ) containing six circular lesion regions with different sizes. The six lesion regions are numbered from #1 ~ #6 with lesion radius of 6.5 to 1.5 pixels, decreasing one pixel at a time. Sinogram data of the designed lesion phantom were obtained by Monte Carlo simulation procedures with the same settings as in section 3.1.1.

After images are reconstructed by the presented algorithms, we first use K-means to cluster the intensity values of each pixel into three clusters, corresponding to dark background, thorax background and lesions respectively. Then, we further employ K-means algorithm to cluster the lesions based on their pixel coordinates. In this process, to present the essential performance of algorithm detectability, we minimize the artificial intervene by only predefining the number of clusters in the second stage clustering the lesions. We also replicate K-means algorithms for 10 times at each stage to eliminate the effect of random initializations.

In figure 3, we present the reconstruction and clustered lesion regions under two different counting rates of Monte Carlo simulations. First of all, we note that the proposed method produces clearer reconstructed images. Secondly, the clustered lesion of ML-EM suffers from random noises while SPS-OS and proposed methods seem to give comparable detecting results. In the  $5 \times 10^5$  studies, SPS-OS, GD and AD methods all fail to discern the 6th lesion. Though ML-EM image has the six clusters, the true lesion region has been submerged in high noises. Quantitative analysis given by Jaccard index of each region with respect to the ground truth are presented in table 3. From the table, we can find that the proposed algorithms give better performance in terms of lesion detectability. More importantly, when lesion size is comparable or smaller than patch size, AD method tends to preserve such lesion regions better than GD method.

### 3.2. Real patient data

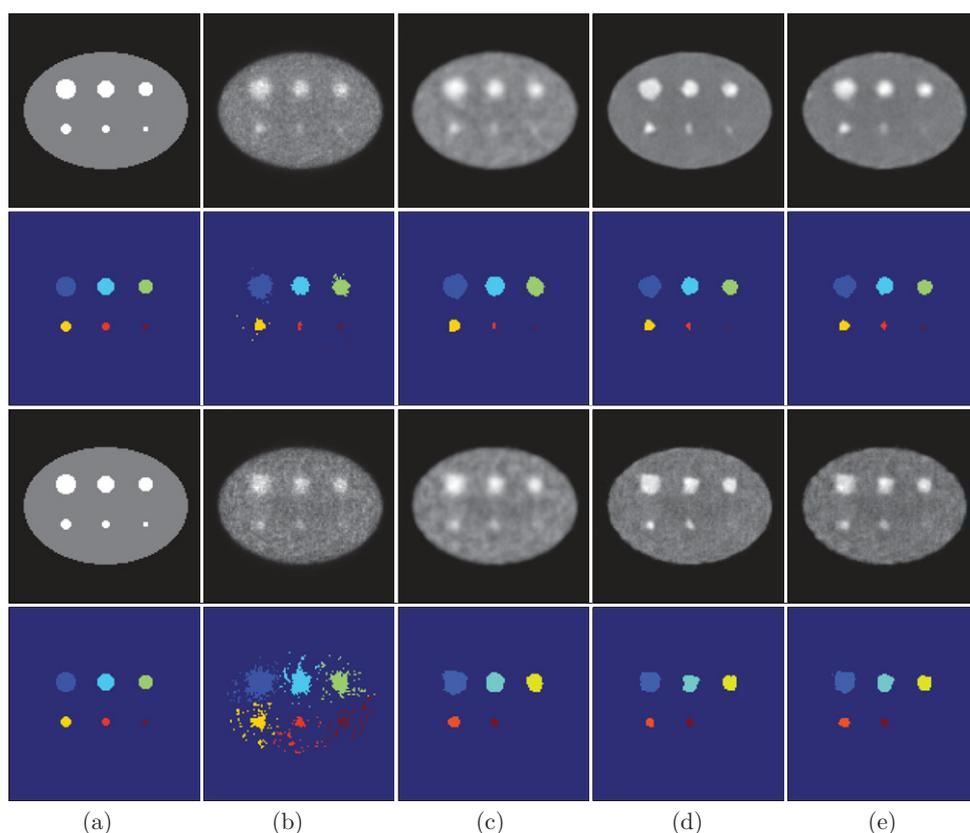
**3.2.1. CTI ECAT data.** In this data set, the 47 slices real patient data are obtained from subject who was scanned on an CTI ECAT PET scanner. The raw sinogram, along with the attenuation



**Figure 2.** Reconstruction of Monte Carlo simulated Zubal phantom data. The first two rows correspond to the reconstructed images and subtraction images under  $10^6$  counting rate. The last two rows correspond to the reconstructed images and subtraction images under  $5 \times 10^5$  counting rate. (a) ML-EM. (b) SPS-OS. (c) GD. (d) AD.

**Table 1.** Statistical analysis of reconstruction on Zubal phantom with  $1 \times 10^6$  counting rate.

	Bias					Variance				
	Whole	ROI1	ROI2	ROI3	ROI4	Whole	ROI1	ROI2	ROI3	ROI4
ML-EM	0.1812	0.1441	0.2143	0.1492	0.2650	0.0522	0.0393	0.0650	0.0331	0.0855
SPS-OS	0.1820	0.1587	0.2020	0.1388	0.2897	0.0524	0.0481	0.0570	0.0294	0.0952
GD	<b>0.1557</b>	0.1389	<b>0.1701</b>	<b>0.1345</b>	<b>0.2133</b>	<b>0.0387</b>	<b>0.0334</b>	<b>0.0445</b>	<b>0.0266</b>	<b>0.0560</b>
AD	0.1584	<b>0.1388</b>	0.1755	0.1347	0.2211	0.0413	0.0345	0.0484	0.0282	0.0608



**Figure 3.** Reconstruction and clustered lesion regions of the lesion phantom. The first two rows correspond to the reconstructed images and the final lesion clusters of experiments with  $1 \times 10^6$  counting rate. The last two rows correspond to the reconstructed images and the final lesion clusters of experiments with  $5 \times 10^5$  counting rate. (a) True. (b) ML-EM. (c) SPS-OS. (d) GD. (e) AD.

and efficiency factors, are radial samples by 192 angular samples (over  $180^\circ$ ). Sinogram have been pre-corrected for delayed coincidences, and four slices of reconstructed images are presented for visual judgment. During this studies, we also fix the tuning parameters in the four presented model through reconstructing all the four slices.

Figure 4 shows the reconstructed images using ML-EM, SPS-OS and the proposed algorithms with GD and AD. Similar to the simulation studies, ML-EM reconstructed images have severe noise corruptions and SPS-OS tends to over-smooth the images. Especially in the last two slices, the Huber penalty function produces severe artifacts. The proposed algorithms produce better image structures and more details are preserved.

**3.2.2. SHR-22000 data.** This data set is a real patient scan acquired by the Hamamatsu SHR-22000. The volunteer patient had a tumor in his left lung. During the scanning, 20 min transmission scan was performed after the 20 min whole body emission scanning. At last, 60 min blank scan was performed to calculate the normalization coefficients and attenuation coefficients. We choose the 32-nd slice out of the 63 overall slices. The raw sinogram data are  $384 \text{ bins} \times 384 \text{ angles}$ , which are reconstructed to image with size  $128 \times 128$ .

**Table 2.** Statistical analysis of reconstruction on Zubal phantom with  $5 \times 10^5$  counting rate.

	Bias					Variance				
	Whole	ROI1	ROI2	ROI3	ROI4	Whole	ROI1	ROI2	ROI3	ROI4
ML-EM	0.1918	0.1631	0.2186	0.1574	0.2653	0.0590	0.0497	0.0694	0.0372	0.0873
SPS-OS	0.1868	0.1707	0.2009	0.1436	0.2871	0.0572	0.0566	0.0596	0.0310	0.0947
GD	<b>0.1611</b>	<b>0.1496</b>	<b>0.1708</b>	<b>0.1362</b>	0.2241	<b>0.0445</b>	<b>0.0434</b>	<b>0.0471</b>	<b>0.0293</b>	<b>0.0611</b>
AD	0.1696	0.1604	0.1795	0.1392	<b>0.2202</b>	0.0493	0.0489	0.0524	0.0310	0.0615

Figure 5 shows the reconstructed images of the real patient data. All four methods are able to resolve the tumor in the patient's left lung whereas image reconstructed by EM method is very noisy and SPS-OS tends to over-smooth the image. The proposed algorithms can reduce the noise but also better delineate the tissues.

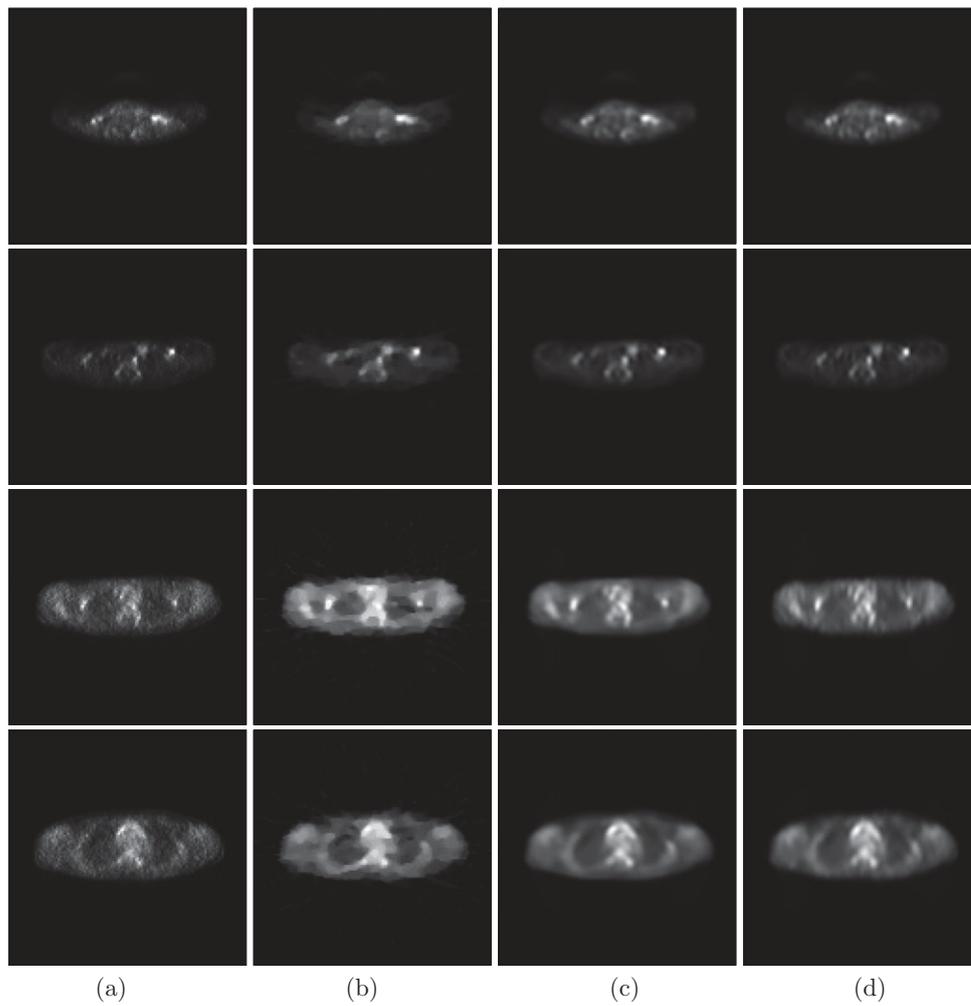
### 3.3. Implementation analysis

**3.3.1. Tuning parameters.** To determine parameter ranges and examine the robustness of the two tuning parameters within the range, we calculate the relative bias and variance based on different choices of  $\lambda$  and  $\epsilon$ .  $\lambda$  ranges from 0.01 to 10 and  $\epsilon$  is between [0.05,0.5]. The sinogram data in this study is the Zubal phantom with  $10^6$  counting rate in section 3.1.1. When different value of  $\lambda$  is tested,  $\epsilon$  is fixed at 0.055. While  $\lambda$  is set 3 for the experiments in terms of different choices of  $\epsilon$ .

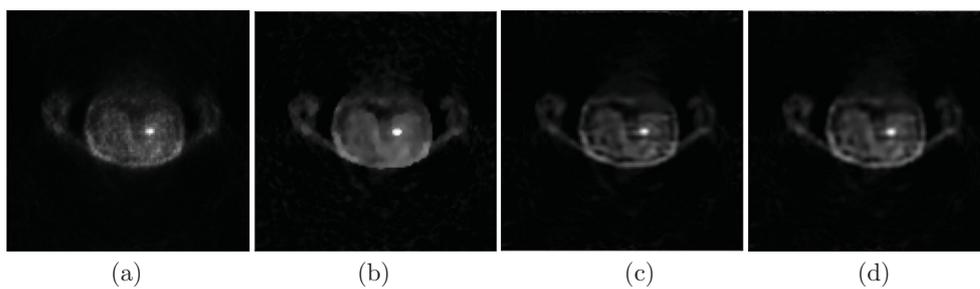
As in figure 6, we can find that bias and variance are relatively stable when  $\lambda$  is in [1, 3]. While the algorithm performance is more sensitive to  $\epsilon$ . Based on figures 6(c) and (d),  $\epsilon$  around 0.1 achieves lower bias and variance.

**3.3.2. Dictionary.** This study examines the impact on sparse representation under a dictionary trained using the patches extracted from images with different structures from the one to be reconstructed. In this experiment, the global dictionary that we have been using to reconstruct thorax PET images is used to reconstruct the Brain phantom in figure 7(a). The sinogram of Brain phantom ( $128 \times 128$ ) is obtained through Monte Carlo simulation with  $1 \times 10^6$  counting rate.

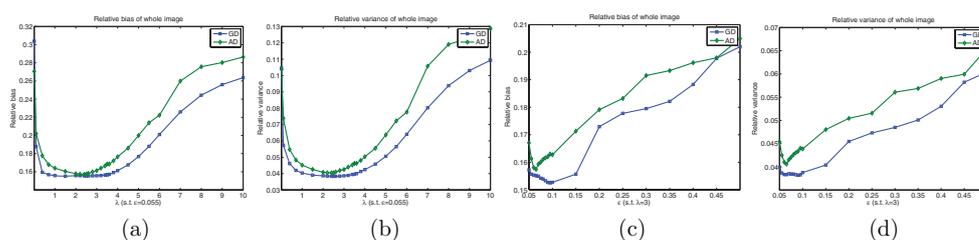
In figure 7, images reconstructed by ML-EM is blur. SPS-OS produces high contrast and clean images but many tiny features have also been over-smoothed. The proposed method with global dictionary produce satisfactory results though the global dictionary is trained from mismatched thorax images. This is because we decompose the anatomical structures into local features. Though the global structures are different between training data and the to be reconstructed image, these small patches can be reconstructed with reasonable accuracy by an appropriate linear combination of dictionary elements. Quantitative evidence is provided in table 4. We also notice that, in this case, AD method gives better estimates by learning local features directly from measured data. However, the sparsity of the representation for reconstructed PET image is affected by using the dictionary trained with the patches extracted from images that have different structures. In table 5, we present the average number of dictionary elements to reconstruct one patch of PET image. When the dictionary is trained using the patches extracted from thorax CT images, one to nine atoms in the dictionary are enough to represent a patch in the reconstructed PET thorax image. However, 22 atoms are needed for a PET brain image in this experiment. The accuracy of sparse representation method would be severely hampered if the sparsity assumption can not be guaranteed.



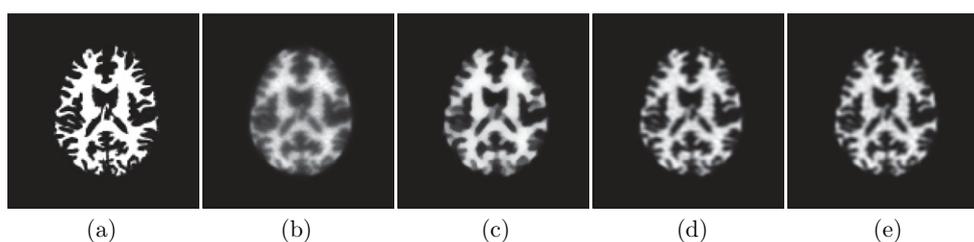
**Figure 4.** Reconstruction results of the real patient data obtained from CTI ECAT. From left to right: ML-EM, SPS-OS, GD, AD. From top to bottom, the 19th, 27th, 35th, 43th slices. (a) ML-EM. (b) SPS-OS. (c) GD. (d) AD.



**Figure 5.** Reconstruction results of the patient data obtained from SHR-22000. From left to right: (a) ML-EM, (b) SPS-OS, (c) GD, (d) AD.



**Figure 6.** Relative bias and variance of reconstructions under different  $\lambda$  and  $\epsilon$ ,  $\lambda \in [0.01, 10]$  and  $\epsilon \in [0.05, 0.5]$ . (a) Bias for  $\lambda$  range. (b) Variance for  $\lambda$  range. (c) Bias for  $\epsilon$  range. (d) Variance for  $\epsilon$  range.



**Figure 7.** True Brain phantom, reconstructed Brain phantom by ML-EM, SPS-OS, GD trained from thorax CT slices and AD (from left to right). (a) True. (b) ML-EM. (c) SPS-OS. (d) GD. (e) AD.

We then present the results based on different sizes of dictionary element in the proposed algorithm,  $7 \times 7$  and  $10 \times 10$ . The Zubal thorax phantom data with  $10^6$  counting rate in section 3.1.1 is used. In figure 8, we can see that images reconstructed by  $10 \times 10$  patch size are over-smoothed compared to  $7 \times 7$ . As the region pointed by the red arrow in the first row of figure 8, it is missed completely in GD method with  $10 \times 10$  patch. In the second row, region enclosed by the rectangle is zoomed in. We can also find that dictionary with smaller patch size preserve image details better.

#### 4. Discussion

We have developed a sparse representation and dictionary learning based framework for PET reconstruction. Given the similarity between functional image structures and anatomical image structures, we train a global dictionary from the anatomical thorax images. On the other hand, we also investigate the possibility of adaptively learn the dictionary from the noisy measured data. Combining the sparse representation using trained dictionary and maximum likelihood estimation of Poisson distribution, we formulate the reconstruction as an optimization problem. An iterative alternating procedure is then developed accordingly.

In the experiments section, we validate the proposed framework on Monte Carlo simulation data and real patient data. The results of classical ML-EM and SPS-OS with Huber potential function are also presented for comparisons. In Zubal phantom, the proposed algorithms produce more accurate estimates of the activity map in terms of their relative bias and variance. Moreover, the experiments show that the proposed method gave more robust reconstructions when the counting rate decreases. In the lesion phantom study, six lesion regions with decreasing radius are designed to demonstrate detectability of reconstruction algorithms. The

**Table 3.** Jaccard index of detected lesions.

	$1 \times 10^6$ counting rate				$5 \times 10^5$ counting rate			
	ML-EM	SPS-OS	GD	AD	ML-EM	SPS-OS	GD	AD
Lesion#1	0.6939	0.6990	<b>0.8563</b>	0.8447	0.4308	0.6850	<b>0.8947</b>	0.8509
Lesion#2	0.8393	0.7698	<b>0.9029</b>	0.8750	0.4524	0.7638	0.7455	<b>0.8165</b>
Lesion#3	0.6635	0.6273	<b>0.8519</b>	0.8481	0.3636	0.6832	<b>0.8608</b>	0.8481
Lesion#4	0.6600	0.7347	0.7442	<b>0.7805</b>	0.2628	0.8222	0.6579	<b>0.8250</b>
Lesion#5	0.4545	0.3810	0.5238	<b>0.5909</b>	0.2410	0.8261	0.8261	<b>0.8636</b>
Lesion#6	0.5455	0.2222	0.2222	<b>0.4444</b>	<b>0.0957</b>	0	0	0

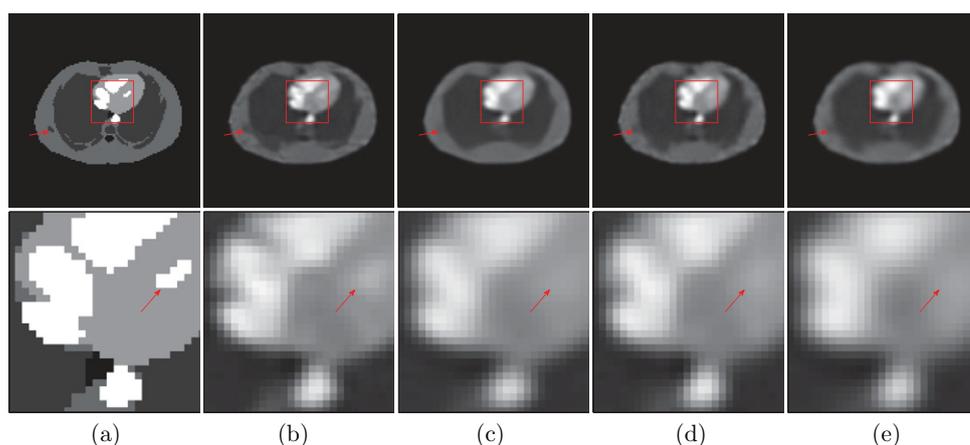
**Table 4.** Statistical analysis of reconstructions based on Brain phantom.

	ML-EM	SPS-OS	GD	AD
Bias	1.3239	1.0730	1.0597	<b>1.0459</b>
Variance	5.5813	3.2954	3.1065	<b>2.9672</b>

proposed method continues to produce cleaner image structures. Quantitative analysis based on Jaccard index shows that GD and AD method outperform the other two algorithms in terms of lesion detection. Moreover, the AD method that learns intrinsic local structures directly from measurements is better at preserving fine structures like the smaller lesions in the second row. In the real patient case, we demonstrate the results of four slices of real thorax data provided by Prof. Fessler and thorax patient data. In these cases, the proposed model also preserve better image structures. On the contrary, the classical ML-EM method reconstructed images are corrupted by noise and SPS-OS method tends to struggle in smoothness and sharp variations and produce artifacts.

One major concern is the possible loss of information in PET image that is missed in the anatomical priors. The global dictionary is trained from CT slices that have similar image structures with PET images. Most local patches from PET images can be reconstructed with reasonable accuracy by an appropriate linear combination of global dictionary elements. The experiments of both simulation and real data have verified this, especially in Brain phantom study, where anatomical priors and PET image have different image structures. However, due to different intensity of PET and CT images, complex and/or fine structures may fail to be recovered accurately. For instance, in lesion study 3.1.2, lesions in the bottom row are smaller than the patch size and they do not exist in the training set. You can find that their profiles are less accurately recovered by proposed model than the first row. ML-EM and SPS-OS method without anatomical priors also did not preserve them well due to high level noise corruptions in the reconstructed images. On the other hand, the AD method that learns local features directly from measurement gives more accurate estimates when anatomical prior and PET image structure mismatched. The fine structures, like small lesions, are also better recovered by AD method.

Due to the time consuming process of sparse coding, patch extraction and dictionary learning (only for AD), the computational time of proposed model is slower than ML-EM and SPS-OS method. As you can see in the method section 2, the proposed algorithm is designed to be parallel as each pixel being updated independently in the iteration scheme. Therefore, further accelerations based on ordered subsets (OS) and graphics processing unit computation can be readily utilized. We believe dual modality, like MRI-PET and PET-CT, can benefit from the proposed paradigms that incorporate anatomical priors into PET reconstruction by



**Figure 8.** Reconstructions with different dictionary element sizes in the proposed algorithm. (a) True. (b) GD  $7 \times 7$ . (c) GD  $10 \times 10$ . (d) AD  $7 \times 7$ . (e) AD  $10 \times 10$ .

**Table 5.** Average number of dictionary elements to reconstruct one patch for PET images.

	Zubal $10^6$	Zubal $5 \times 10^5$	Lesion $10^6$	Lesion $5 \times 10^5$	Brain
Average number	4.56	5.34	1.66	1.40	<b>22.48</b>
CRI ECAT data					SHR-22000 data
	Slice #19	Slice #27	Slice #35	Slice # 43	Lung
Average number	7.59	9.09	7.42	6.87	4.95

providing global dictionaries. The individual anatomical scan done just before the PET study could be better source for dictionary training. But such general global dictionary obtained without exact correspondence demonstrates the wider applicability of the proposed methods. However, experiments based on PET-CT or MRI-PET and investigations of its practical issues for clinical use are definitely the most urgent and important work in the future. By now, we only investigate the thorax images. Intuitively, similar results can be achieved for brain and cardiac PET, where PET have been used in research and clinical practice. Since pre-trained anatomical dictionary gives competitive results throughout the experiments. We think it would be interesting to construct a global dictionary for the whole human body, which consists of several categories for each part of human body. Such work is under investigation.

## Acknowledgments

We would like to thank Professor Fessler for sharing the real patient data on his website. H Liu is supported by the National Basic Research Program of China: 2010CB732500, National Natural Science Foundation of China: LR12F03001, 61271083, 61427807. Y Chen is supported by the Division of Mathematical Sciences: 1319050, Division Of Industrial Innovation and Partnerships: 1237814.

## References

- Aharon M, Elad M and Bruckstein A 2006 K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation *IEEE Trans. Signal Process.* **54** 4311–22
- Bouman C A and Sauer K 1996 A unified approach to statistical tomography using coordinate descent optimization *IEEE Trans. Image Process.* **5** 480–92
- Chang J-H, Anderson J M and Votaw J 2004 Regularized image reconstruction algorithms for positron emission tomography *IEEE Trans. Med. Imaging* **23** 1165–75
- Chen Y, Ye X and Huang F 2010 A novel method and fast algorithm for MR image reconstruction with significantly under-sampled data *Inverse Problems Imag.* **4** 223–40
- De Pierro A 1995 A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography *IEEE Trans. Med. Imaging* **14** 132–7
- Dewaraja Y K, Koral K F and Fessler J A 2010 Regularized reconstruction in quantitative SPECT using CT side information from hybrid imaging *Phys. Med. Biol.* **55** 2523
- Elad M and Aharon M 2006 Image denoising via sparse and redundant representations over learned dictionaries *IEEE Trans. Image Process.* **15** 3736–45
- Fessler J A 1994 Penalized weighted least-squares image reconstruction for positron emission tomography *IEEE Trans. Med. Imaging* **13** 290–300
- Fessler J A and Erdogan H 1998 A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction *Nuclear Science Symp. Conf. Record (Toronto, ON, 8–14 November 1998)* vol 2 (Piscataway, NJ: IEEE) pp 1132–5
- Gindi G, Lee M, Rangarajan A and Zubal I G 1993 Bayesian reconstruction of functional images using anatomical information as priors *IEEE Trans. Med. Imaging* **12** 670–80
- Green P J 1990 Bayesian reconstructions from emission tomography data using a modified EM algorithm *IEEE Trans. Med. Imaging* **9** 84–93
- 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
- Mehranian A, Rahmim A, Ay M R, Kotasidis F and Zaidi H 2013 An ordered-subsets proximal preconditioned gradient algorithm for edge-preserving PET image reconstruction *Med. Phys.* **40** 052503
- Qi J and Leahy R M 2006 Iterative reconstruction techniques in emission computed tomography *Phys. Med. Biol.* **51** R541
- Ravishankar S and Bresler Y 2011 MR image reconstruction from highly undersampled k-space data by dictionary learning *IEEE Trans. Med. Imaging* **30** 1028–41
- Shepp, L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography *IEEE Trans. Med. Imaging* **1** 113–22
- Tang J, Tsui B M W and Rahmim A 2008 Bayesian PET image reconstruction incorporating anatomical functional joint entropy *IEEE Int. Symp. on Biomedical Imaging (Paris, France, 14–17 May 2008)* pp 1043–6
- Tropp J A and Gilbert A C 2007 Signal recovery from random measurements via orthogonal matching pursuit *IEEE Trans. Inform. Theory* **53** 4655–66
- Vunckx K, Atre A, Baete K, Reilhac A, Deroose C M, Van Laere K and Nuyts J 2012 Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging *IEEE Trans. Med. Imaging* **31** 599–612
- Xu Q, Yu H, Mou X, Zhang L, Hsieh, J and Wang G 2012 Low-dose x-ray CT reconstruction via dictionary learning *IEEE Trans. Med. Imaging* **31** 1682–97