Physics in Medicine & Biology

Anatomy-guided PET reconstruction using l_1 bowsher prior

To cite this article before publication: Seung Kwan Kang *et al* 2021 *Phys. Med. Biol.* in press https://doi.org/10.1088/1361-6560/abf2f7

IOP Publishing *Phys. Med. Biol.*
Phys. Med. Biol. **XX** (XXXX) XXXXXX https://doi.org/XXXX/XXXX

Anatomy-Guided PET Reconstruction Using l_1 Bowsher Prior

Seung Kwan Kang [1,2] and Jae Sung Lee[1,2,3,4]

[1] Department of Nuclear Medicine, Seoul National University College of Medicine, Seoul 03080, Korea
[2] Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul 03080, Korea
[3] Institute of Radiation Medicine, Medical Research Center, Seoul National University College of Medicine, Seoul 03080, Korea
[4] Brightonix Imaging Inc., Seoul 03080, Korea

E-mail: jaes@snu.ac.kr

Received xxxxxx
Accepted for publication xxxxxx
Published xxxxxx

Abstract

Advances in simultaneous positron emission tomography/magnetic resonance imaging (PET/MRI) technology have led to an active investigation of the anatomy-guided regularized PET image reconstruction algorithm based on MR images. Among the various priors proposed for anatomy-guided regularized PET image reconstruction, Bowsher's method based on second-order smoothing priors sometimes suffers from over-smoothing of detailed structures. Therefore, in this study, we propose a Bowsher prior based on the l_1-norm and an iteratively reweighting scheme to overcome the limitation of the original Bowsher method. In addition, we have derived a closed solution for iterative image reconstruction based on this non-smooth prior. A comparison study between the original l_2 and proposed l_1 Bowsher priors was conducted using computer simulation and real human data. In the simulation and real data application, small lesions with abnormal PET uptake were better detected by the proposed l_1 Bowsher prior methods than the original Bowsher prior. The original l_2 Bowsher leads to a decreased PET intensity in small lesions when there is no clear separation between the lesions and surrounding tissue in the anatomical prior. However, the proposed l_1 Bowsher prior methods showed better contrast between the tumors and surrounding tissues owing to the intrinsic edge-preserving property of the prior which is attributed to the sparseness induced by l_1-norm, especially in the iterative reweighting scheme. Besides, the proposed methods demonstrated lower bias and less hyper-parameter dependency on PET intensity estimation in the regions with matched anatomical boundaries in PET and MRI. Therefore, these methods will be useful for improving the PET image quality based on the anatomical side information.

Keywords: image reconstruction, positron emission tomography, anatomical prior, regularization

1. Introduction

Positron emission tomography (PET) is a medical imaging device that is highly sensitive in identifying functional and molecular abnormalities in patients with various diseases. However, PET has relatively poor spatial resolution and higher noise levels compared to anatomical imaging modalities, such as computed tomography (CT) and magnetic resonance imaging (MRI). To improve the image quality and quantitative accuracy of PET, various iterative reconstruction algorithms that account for the noise properties of measured data have been widely investigated and adopted (Shepp and Vardi 1982, Lange and Carson 1984, Qi and Leahy 2006). However, the formulation of optimization problems based on Poisson statistics for iterative PET image reconstruction is generally ill-posed at high noise levels (Louis and Natterer 1983, Tikhonov 1963, Gourion and Noll 2002).

To mitigate such problems, we use maximum a posteriori (MAP) reconstruction algorithms, also known as penalized likelihood reconstruction methods, that stabilize the solution by incorporating prior information into the formulation of the optimization problem (Artzy *et al.* 1979, Hebert and Leahy 1989, Kaufman 1993, Fessler and Hero 1995, Qi and Leahy 2000). However, smoothing priors (e.g., the quadratic prior) used for reducing noise in reconstructed images also eliminate some essential high-frequency features such as edge and small lesions. To preserve these features, we can use edge-preserving priors, such as the non-local means prior (Wang and Qi 2012). Alternatively, anatomical information provided by CT or MRI can be utilized as a prior (anatomy-guided regularized PET image reconstruction) (Bowsher *et al.* 1996, Vunckx *et al.* 2012, Bai *et al.* 2013, Tang and Rahmim 2015, Ehrhardt *et al.* 2016, Novosad and Reader 2016, Mehranian *et al.* 2017, Schramm *et al.* 2018, Knoll *et al.* 2017).

In recent years, advances in simultaneous PET/MRI technology (Judenhofer *et al.* 2008, Yoon *et al.* 2012, Delso *et al.* 2011, Ko *et al.* 2016, Levin *et al.* 2016) have led to an active investigation of such anatomy-guided regularized PET image reconstruction algorithms based on MR images. MRI with higher soft-tissue contrast compared to CT would be a useful anatomic prior, particularly for brain and head/neck regions. Either raw MR images or segmentation outcomes can be used as the priors for PET image reconstruction (Baete *et al.* 2004a, Baete *et al.* 2004b, Nuyts *et al.* 2005, Goffin *et al.* 2010, Hutchcroft *et al.* 2016). In this study, we focused on the former method because the segmentation-based method is vulnerable to segmentation error.

Among the various priors proposed for anatomy-guided regularized PET image reconstruction, Bowsher's method is one of the best performing anatomical priors (Bowsher *et al.* 2004, Schramm *et al.* 2018). However, the original Bowsher's method that is based on l_2-norm prior sometimes suffers from over-smoothing of detailed structures. Therefore, in this study, we propose a Bowsher prior based on the l_1-norm to overcome the limitation of the original Bowsher method. An interesting property of newly derived prior is that it induces sparseness of the image like total-variation (TV) prior (Chambolle *et al.* 2010, Esser 2009). Accordingly, we could improve the performance of the proposed prior by applying an iterative reweighting scheme introduced in (Candes *et al.* 2008). A modified proximal gradient algorithm was used to solve the optimization problem of Poisson log-likelihood and non-smooth prior. Computer simulation studies under different noise conditions were conducted to compare the performance of the original and proposed l_1 Bowsher priors. We also analyzed both priors using clinical [^{18}F]FDG PET images.

2. Methods

2.1 PET Data Model

The Poisson log-likelihood model is used for PET image reconstruction to account for the statistical properties of PET image acquisition (Lange and Carson 1984, Qi and Leahy 2006). However, the maximum log-likelihood solution for unknown images usually yields noisy results because the problem is fundamentally ill-posed. Thus, regularization is considered to recover better images by imposing some appropriate assumptions. The penalized negative log-likelihood estimate of the unknown image $\boldsymbol{x}$ is expressed as

$$\arg\min_{x\geq 0} \sum_i \hat{y}_i(\boldsymbol{x}) - \boldsymbol{y}_i \log \hat{y}_i(\boldsymbol{x}) + \beta R(\boldsymbol{x}) \tag{1}$$

where y_i is the observed data for the ith line of response, $R(\cdot)$ is the penalty function, β is a weighting parameter of the penalty function, and $\hat{y}_i(\cdot)$ is a forward projection of the image to the *i*-th line of response. The expected count distribution $\hat{y}_i(\boldsymbol{x})$ for image $\boldsymbol{x}$ is expressed as $\hat{y}_i(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{s}$, where A is a system matrix and s denotes the expected distribution of random and scatter events. We can provide anatomical information available in the MR image to the penalty function $R(\cdot)$. As mentioned earlier, one of the popular choices for the penalty function $R(\cdot)$ is the Bowsher prior (Bowsher *et al.* 2004), which will be discussed in the following sections.

2.2 Original Bowsher Prior

The original Bowsher prior is expressed as (Bowsher *et al.* 2004, Schramm *et al.* 2018)

$$R_{l_2}(\boldsymbol{x}|\boldsymbol{z}) = \sum_j \sum_{l\in N_j} w_{lj}\left(\boldsymbol{x}_l - \boldsymbol{x}_j\right)^2, \tag{2}$$

$$w_{lj} = \begin{cases} 1 & \forall l \in N_j, if\ \ \exists z_k \in B_j \\ & where\ |z_j - z_l| \leq |z_j - z_k| \\ 0 & \text{else} \end{cases} \tag{3}$$

where $\boldsymbol{z}$ is a prior MR image, $\boldsymbol{z}_j$ is a *j*-th voxel of the MR image and N_j is the neighbor voxel of the *j*-th voxel. The weight $w_{\cdot j}$ uses the difference between the center of the MR image patch and its surrounding voxels to determine the smoothness in the homogenous region. If the difference is large, the boundary of the given image is preserved. B_j consists of the b most similar voxels in the anatomical image around the *j*-th voxel. In the previous study, authors showed that the modifying quadratic term in (2) to relative difference yielded better performance (Vunckx and Nuyts 2010, Vunckx *et al.* 2012, Schramm *et al.* 2018).

$$R_{l_2}^{rel} = \sum_j \sum_{l \in N_j} w_{lj} \frac{(\boldsymbol{x}_l - \boldsymbol{x}_j)^2}{\boldsymbol{x}_l + \boldsymbol{x}_j} \tag{4}$$

To reconstruct the image using this prior, we utilized the asymmetric Bowsher prior and one-step-late algorithm developed in (Nuyts *et al.* 2002, Vunckx *et al.* 2012). The update of each voxel $\boldsymbol{x}_j$ is expressed as the following equation:

$$\boldsymbol{x}_j^{n+1} = \boldsymbol{x}_j^n + \left(\frac{\partial L}{\partial x_j} + \frac{\partial R_{l_2}^{rel}}{\partial x_j}\right) / \left(\frac{a_j}{x_j^n} - \frac{\partial^2 R_{l_2}^{rel}}{\partial x_j^2}\right), \tag{5}$$

where L is the negative log-likelihood, and $a_j = \sum_j a_{ij}$ is the sum of the system matrix. This original Bowsher prior is a l_2-norm prior; therefore, it sometimes suffers from over-smoothing of detailed structures.

2.3 Proposed l_1 Bowsher Prior

Our proposed l_1 Bowsher prior is defined as follows:

$$R_{l_1}(\boldsymbol{x}|\boldsymbol{z}) = \sum_j \sum_{l \in N_j} w_{lj} |\boldsymbol{x}_l - \boldsymbol{x}_j| \,. \tag{6}$$

Instead of using a squared function between the center voxel and its neighbors, the l_1-norm was exploited. This prior is convex but not smooth. Therefore, we devised a modified proximal gradient algorithm because the reconstruction scheme from the original Bowsher prior was not applicable. At first, the EM update equation can also be described as (Sangtae and Fessler 2003)

$$\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - D(\boldsymbol{x}^n)\nabla L(\boldsymbol{x}^n), \tag{7}$$

where $D(\boldsymbol{x}^n) = diag(\boldsymbol{x}^n / A^T \mathbf{1})$, and $A^T \mathbf{1}$ is the backprojection of a vector whose elements are equal to 1. Thus, $\boldsymbol{x}^{n+1}$ is the solution of the following problem:

$$\begin{aligned} &\text{aargmin}_x \, L(\boldsymbol{x}^n) + \nabla L(\boldsymbol{x}^n)^T (\boldsymbol{x} - \boldsymbol{x}^n) \\ &\quad + \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{x}^n\|_{D(\boldsymbol{x}^n)^{-1}}^2 \end{aligned} \tag{8}$$

$$\begin{aligned} &= \text{argmin}_x \, \frac{1}{2} \|\nabla L(\boldsymbol{x}^n)\|_{D(\boldsymbol{x}^n)^{-1}}^2 + \nabla L(\boldsymbol{x}^n)^T (\boldsymbol{x} - \boldsymbol{x}^n) \\ &\quad + \frac{1}{2} D \|\boldsymbol{x} - \boldsymbol{x}^n\|_{D(\boldsymbol{x}^n)^{-1}}^2 \\ &= \text{argmin}_x \, \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{x}^n + D(\boldsymbol{x}^n)\nabla L(\boldsymbol{x}^n)\|_{D(\boldsymbol{x}^n)^{-1}}^2 .. \end{aligned}$$

In the (8), the $L(\boldsymbol{x}^n)$ can be removed because it does not depend on $\boldsymbol{x}$ and $1/2\|\nabla L(\boldsymbol{x}^n)\|_{D(\boldsymbol{x}^n)^{-1}}^2$ can be inserted into the first equation for the same reason. This equation can be regarded as the second-order Taylor approximation of Poisson log-likelihood where the a Hessian is substituted for $D(\boldsymbol{x}^n)^{-1}$. Consequently, we can rewrite the original problem (1) in the following approximated form after combining the proposed regularization term:

$$\text{argmin}_{x \geq 0} \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{x}^{n+1}\|_{D(\boldsymbol{x}^n)^{-1}}^2 + \beta R_{l_1}(\boldsymbol{x}|\boldsymbol{z}). \tag{9}$$

This formula is the modified proximal mapping for a penalty function R_{l_1}, where $D(\boldsymbol{x}^n)^{-1}$ plays the role of diagonal weighting. The proximal gradient algorithm is efficient when a closed expression of the proximal mapping is provided. We were able to determine the proximal mapping for the individual voxel $\boldsymbol{x}_j$ by applying the subgradient optimality condition (Parikh and Boyd 2014) (see Figure 1.).

$$\begin{aligned} &0 \in \partial \left(\frac{1}{2d_j} \left(\boldsymbol{x}_j - \boldsymbol{x}_j^{n+1}\right)^2 + \beta \left(\sum_{l \in N_j} w_{lj} |\boldsymbol{x}_l - \boldsymbol{x}_j| + \right.\right. \\ &\qquad \left.\left. \sum_{m \in N_j} w_{jm} |\boldsymbol{x}_j - \boldsymbol{x}_m| \right) \right) \end{aligned} \tag{10}$$

The final term $\sum_{m \in N_j} w_{jm} |\boldsymbol{x}_j - \boldsymbol{x}_m|$ is included for the symmetricity of the proposed prior. If we remove this term, it becomes asymmetric l_1 Bowsher prior. We used the asymmetric l_1 Bowsher in this study. The solution of the subgradient is given by if we set $\boldsymbol{x}_i \neq \boldsymbol{x}_j$ for $i, j \in N_j$ and $\boldsymbol{x}_i < \boldsymbol{x}_j$ for $\forall\, i < j$:

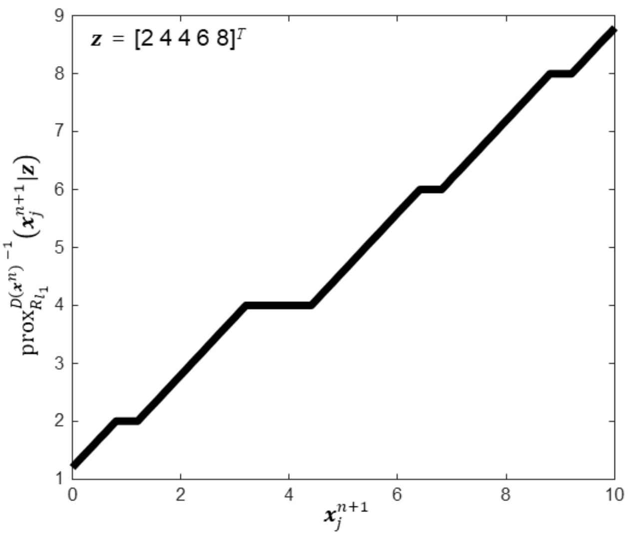

Figure 1. Example of the proposed proximal operator.

$$\partial\left(\sum_{l\in N_j} w_{lj}|\boldsymbol{x}_l-\boldsymbol{x}_j|\right)= \begin{cases} -\sum_{l\in N_j} w_{lj} & \text{if } \boldsymbol{x}_j<\boldsymbol{x}_1 \\ \vdots & \\ w_i[-1,1]+\sum_{l=1}^{i-1} w_{lj}-\sum_{l=i+1}^{n_l} w_{lj} & \text{if } \boldsymbol{x}_j=\boldsymbol{x}_i \\ \sum_{l=1}^{i} w_{lj}-\sum_{l=i+1}^{n_l} w_{lj} & \text{if } \boldsymbol{x}_i<\boldsymbol{x}_j<\boldsymbol{x}_{i+1} \\ \vdots & \\ \sum_{l\in N_j} w_{lj} & \text{if } \boldsymbol{x}_j>\boldsymbol{x}_{n_l} \end{cases}. \quad (11)$$

Rearranging (11) yields the following solution:

$$\mathrm{prox}_{R_{l_1}}^{D(\boldsymbol{x}^n)^{-1}}(\boldsymbol{x}_j^{n+1}|\boldsymbol{z})= \begin{cases} \boldsymbol{x}_j^{n+1}+d_j\beta\sum_{l\in N_j} w_{lj} & \text{if } \boldsymbol{x}_j^{n+1}\in S_1^+ \\ \vdots & \\ \boldsymbol{x}_i & \text{if } \boldsymbol{x}_j^{n+1}\in S_i^- \\ \boldsymbol{x}_j^{n+1}-d_j\beta\sum_{l=1}^{i} w_{lj} & \\ \quad +d_j\beta\sum_{l=i+1}^{n_l} w_{lj} & \text{if } \boldsymbol{x}_j^{n+1}\in S_i^+ \\ \vdots & \\ \boldsymbol{x}_j^{n+1}-d_j\beta\sum_{l\in N_j} w_{lj} & \text{if } \boldsymbol{x}_j^{n+1}\in S_{n_l}^+ \end{cases}, \quad (12)$$

where n_l is the number of elements in the set N_j, $2\le i\le n_l-1$, and

$$S_1^+=\left\{u \middle| u\le \boldsymbol{x}_1-d_j\beta\sum_{l\in N_j} w_{lj}\right\}$$

$$\vdots \quad (13)$$

$$S_i^- =\left\{u \middle| \begin{array}{l}\boldsymbol{x}_i+d_j\beta\left(\sum_{l=1}^{i-1} w_{lj}-\sum_{l=i+1}^{n_l} w_{lj}-w_{li}\right)<u \\ \le \boldsymbol{x}_i+d_j\beta\left(\sum_{l=1}^{i-1} w_{lj}-\sum_{l=i+1}^{n_l} w_{lj}+w_{li}\right)\end{array}\right\}$$

$$S_i^+=\left\{u \middle| \begin{array}{l}\boldsymbol{x}_i+d_j\beta\left(\sum_{l=1}^{i} w_{lj}-\sum_{l=i+1}^{n_l} w_{lj}\right)< \\ u\le \boldsymbol{x}_{i+1}+d_j\beta\left(\sum_{l=1}^{i} w_{lj}-\sum_{l=i+1}^{n_l} w_{lj}\right)\end{array}\right\}$$

$$\vdots$$

$$S_{n_l}^+=\left\{u \middle| u>\boldsymbol{x}_{n_l}+d_j\beta\sum_{l\in N_j} w_{lj}\right\}.$$

Note that $S_1^+\cup\cdots\cup S_i^-\cup S_i^+\cup\cdots\cup S_{n_l}^+=\mathbb{R}$ and each $S_\cdot^\cdot$ is disjoint sets. The example of proximal mapping is presented in Figure 1, and it is similar to the soft thresholding operator (Beck and Teboulle 2009b). Therefore, image reconstruction with the proposed l_1 Bowsher prior is conducted by applying the EM update (8) followed by the modified proximal operator update (12). Both the original and proposed Bowsher prior reconstruction algorithms can be accelerated by replacing the EM update with the ordered subset (OS) algorithm.

2.4 Iterative Reweighting

Proposed l_1 Bowsher prior (6) is similar to TV-l_1 regularization which is one of the sparsity-inducing methods (Chambolle *et al.* 2010, Esser 2009). Thus, we can apply the iterative reweighting method to further enforce the sparsity of the proposed l_1 Bowhser prior (Candes *et al.* 2008). The modified prior is given by:

$$R_{l_1}^{IR}(\boldsymbol{x}|\boldsymbol{z})=\sum_j\sum_{l\in N_j} w_{lj}^{IR} w_{lj}|\boldsymbol{x}_l-\boldsymbol{x}_j|\,, \quad (14)$$

$$w_{lj}^{IR}=\frac{1}{w_{lj}|\boldsymbol{x}_l-\boldsymbol{x}_j|+\epsilon}, \quad (15)$$

where $\epsilon>0$ is the design parameter that controls the algorithm's stability, which yields relatively consistent results for its variation (Candes *et al.* 2008). If the $w_{lj}\neq 0$, the prior becomes l_0-norm, the number of non-zero elements, leading to the sparsity of the boundary voxels. In our experiments, $\epsilon=0.1$ was used for both simulation and clinical datasets. For the optimization, the weights w_{lj} of proximal operator (12) at each iteration n is modified to as in Candes *et al.*:

$$w_{lj}^{IR,n}=\frac{1}{w_{lj}|\boldsymbol{x}_l^n-\boldsymbol{x}_j^n|+\epsilon}, \quad (16)$$

Algorithm 1. Modified proximal gradient with ordered subsets

input $\boldsymbol{y}$ and $\boldsymbol{z}$
initialize x
for $n_1 = 1 \ldots n_{\text{out}}$ **do**
  for $n_2 = 1 \ldots n_{\text{subsets}}$ **do**
    $\boldsymbol{x}^{\text{EM}} = \boldsymbol{x} - D(\boldsymbol{x})\nabla L_{n_2}(\boldsymbol{x})$ (ordinary EM using subsets)
      for $j = 1 \ldots n_j$ **do**
        $\boldsymbol{x}_{j,\text{prox}}^{\text{EM}} = \text{prox}_{R_{l_1}}^{D(\boldsymbol{x}_j)^{-1}}\left(\boldsymbol{x}_j^{\text{EM}} | \boldsymbol{z}\right)$ (proximal operator)
      set $\boldsymbol{x} = \boldsymbol{x}_{\text{prox}}^{\text{EM}}$
    end
  end
return

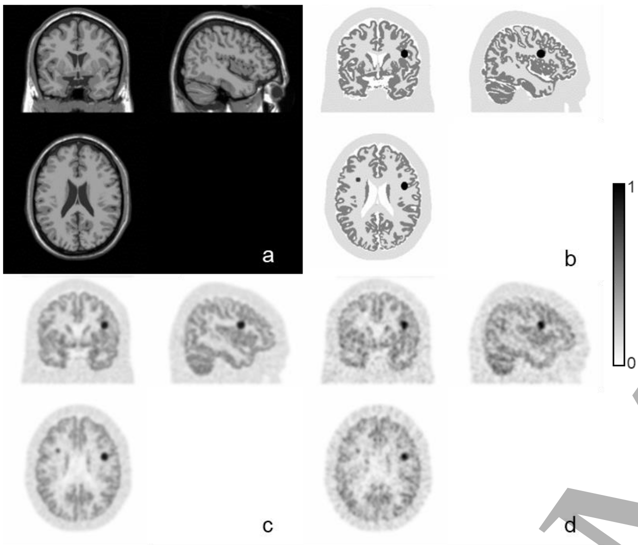

Figure 2. Simulated brain phantoms. (a) MRI and (b) PET (c) OSEM reconstruction with low-level noise and 5 mm Gaussian filter (total 7.0×10^7 prompt counts) (d) OSEM reconstruction with high-level noise and 5 mm Gaussian filter (total 1.4×10^7 prompt counts).

which means that the weight of the current iteration is calculated using the previous images. Accordingly, the optimization problem is not convex anymore because of the l_0-norm, so the convergence to a global solution is not guaranteed. Therefore, the proper initial condition is important and we start this iteratively reweighting scheme after one iteration of OS algorithm. This approach did not cause convergence problems, at least in our experiments, which means that the found solution was not far from the global solution.

2.5 Computer Simulations

We generated the ground truth PET image based on the MR image and its segmentations obtained from BrainWeb (Cocosco *et al.* 1997). The idea was divided into four regions: gray matter (GM), white matter and others (WM and so on), small tumor, and large tumor. We assigned image intensities of 0.5, 0.125, 0.75, and 1 to gray matter, white matter and others, small tumor and large tumor, respectively (Figure 2). Attenuation map was also generated from the ground truth image and a scatter map was acquired by filtering the projections with 50 mm Gaussian FHWM. Scatter fraction was 20%. Two different levels of Poisson noise were added to the projections assuming two different situations: 5 min acquisition (total 7.0×10^7 prompt counts) and 1 min acquisition (total 1.4×10^7 prompt counts) using Siemens Biograph mMR system (Siemens Healthcare, Knoxville, TN), where the number of views in the sinogram was 168. To analyze the results statistically, 15 independent noise realizations are produced. We compared three different image reconstruction strategies: original l_2 Bowsher prior with a relative difference, proposed l_1 Bowsher prior and l_1 Bowsher prior with iterative reweighting. The initial conditions for all the compared algorithms were the output of the first iteration of OSEM. The OS algorithm had 21 subsets and the number of outer iterations was 6. The Bowsher prior was calculated in the nearest 80 voxels. Although the previous report showed the optimal number of selected voxels in the patch (b) was about 10 (Vunckx and Nuyts 2010, Vunckx *et al.* 2012), we also examined larger patch size (20). The regularization parameters for the original Bowsher prior were from 0.1×2^0 to 0.1×2^7 with logarithmic scale 2, and those for the proposed l_1 Bowsher prior were 0.1×2 to 0.1×2^8 divided into the same logarithmic scale. Attenuation and scatter were corrected during image reconstruction, but spatial resolution modeling was not applied.

2.6 Human Data

The proposed method was applied to two different sets of human data acquired using the Siemens Biograph mMR system. One of them was obtained from the PET/MRI scan of a healthy volunteer (59 years old male) acquired 110 min after the injection of 192 MBq [^{18}F]FDG. The PET scan duration was 10 min. A T1-weighted structural MRI was also acquired

Algorithm 2. OS-Modified proximal gradient with iteratively reweighting

1: input $\boldsymbol{y}$ and $\boldsymbol{z}$
2: **initialize** x
3: **for** $n_1 = 1 \ldots n_{\mathrm{out}}$ **do**
4: **for** $n_2 = 1 \ldots n_{\mathrm{subsets}}$ **do**
5: $\boldsymbol{x}^{\mathrm{EM}} = \boldsymbol{x} - D(\boldsymbol{x})\nabla L_{n_2}(\boldsymbol{x})$
6: **for** $j = 1 \ldots n_j$ **do**
7: **if** $n_1 = 1$
8: $\boldsymbol{x}_{j,\mathrm{prox}}^{\mathrm{EM}} = \mathrm{prox}_{R_{l_1}}^{D(x_j)^{-1}}\left(\boldsymbol{x}_j^{\mathrm{EM}}|\boldsymbol{z}\right)$ (eq. (12))
9: **set** $\boldsymbol{x} = \boldsymbol{x}_{\mathrm{prox}}^{\mathrm{EM}}$
10: **else**
11: $\boldsymbol{x}_{j,\mathrm{prox}}^{\mathrm{EM}} = \mathrm{prox}_{R_{l_1}^{IR}}^{D(x_j)^{-1}}\left(\boldsymbol{x}_j^{\mathrm{EM}}|\boldsymbol{z}\right)$ (eq. (12) with weight (16))
12: **set** $\boldsymbol{x} = \boldsymbol{x}_{\mathrm{prox}}^{\mathrm{EM}}$
13: **end**
14: **end**
15: **end**
16: **end**
17: **return**

using the ultrafast gradient-echo sequence and reconstructed into a 208 × 256 × 256 matrix with voxel sizes of 1.0 × 0.98 × 0.98 mm (An *et al.* 2016).

The other set was the PET/MRI data of a patient with head and neck cancer 76 years old female. Both T1- and T2-weighted MRIs were acquired using a turbo spin-echo sequence, whereas the [18F]FDG PET scan was obtained after 110 min injection of 256 MBq of the radiotracer. The dimension of T1-weighted image was 290 × 320 × 42 with voxel sizes of 0.69 × 0.69 × 4.92 mm and that of T2-weighted image was 640 × 640 × 42 with voxel sizes of 0.34 × 0.34 × 4.92 mm. Of the MRIs, only T1-weighted MR images were used for the regularized PET reconstruction. Retrospective use of all human data was approved by the Institutional Review Board of our institute.

Deep learning-based super-resolution along the *z*-axis was performed because the slice thickness of the acquired MR image was thicker than that of the PET scan (Kang *et al.* 2021). The SPM12 (SPM12; University of College London, UK) program was used to re-slice the MR images to have the same voxel size and dimension as that of the PET scan. The Fourier rebinning (FORE) algorithm was applied to the pre-corrected PET sinogram data, and 2D projection and the backprojection algorithm were used (Defrise *et al.* 1997). The same regularization parameters or post-filters as those used in the computer simulation were applied. The number of voxels selected within the patch was fixed at 20, which was showed quantitatively better performances in the simulation study. The regularization parameters for the original Bowsher prior were from 0.1×2^{-1} to 0.1×2^{7} with logarithmic scale 2, and those for the proposed l_1 Bowsher prior were 0.1×2^{0} to 0.1×2^{8}. For the clinical data, consistent with standard OSEM-reconstruction methods for brain imaging on our clinical scanner, we did not apply spatial resolution modeling for OSEM nor for the 3 prior models.

2.7 Image Analysis

Standard deviation (STD) and bias in the PET image intensity were calculated for each region in the simulation study:

$$\mathrm{Bias}^{\mathrm{region}} = \frac{1}{n}\sum_i \boldsymbol{x}_{\mathrm{GT},i}^{\mathrm{region}} - \boldsymbol{x}_{\mathrm{recon},i}^{\mathrm{region}}, \tag{17}$$

$$\mathrm{STD} = \sqrt{\frac{\sum_i\left(\boldsymbol{x}_{\mathrm{GT},i}^{\mathrm{region}} - \overline{\boldsymbol{x}}_{\mathrm{recon}}^{\mathrm{region}}\right)^2}{n^{\mathrm{region}} - 1}}, \tag{18}$$

where $\boldsymbol{x}_{\mathrm{recon}}$ is a reconstructed image of the given region (GM, WM, and tumors), $\overline{\boldsymbol{x}}_{\mathrm{recon}}^{\mathrm{region}}$ is the mean value over the given region, n^{region} is the number of voxels, and $\boldsymbol{x}_{\mathrm{GT}}^{\mathrm{region}}$ is the ground truth value of each region. The number n is the number of instances in the ensemble (n=15).

From the [18F]FDG brain PET of a healthy volunteer, we calculated the mean uptake level (kBq/ml) in the frontal lobe, cingulate cortex, superior parietal gyrus, and lateral temporal gyrus using regions of interest (ROI) drawn only on the gray matter pixels shown in the MRI. The SPM12 program was used to extract gray matter and the above ROIs were defined in the AAL template (Tzourio-Mazoyer *et al.* 2002, Ashburner and Friston 2005). The standard deviation of the white matter

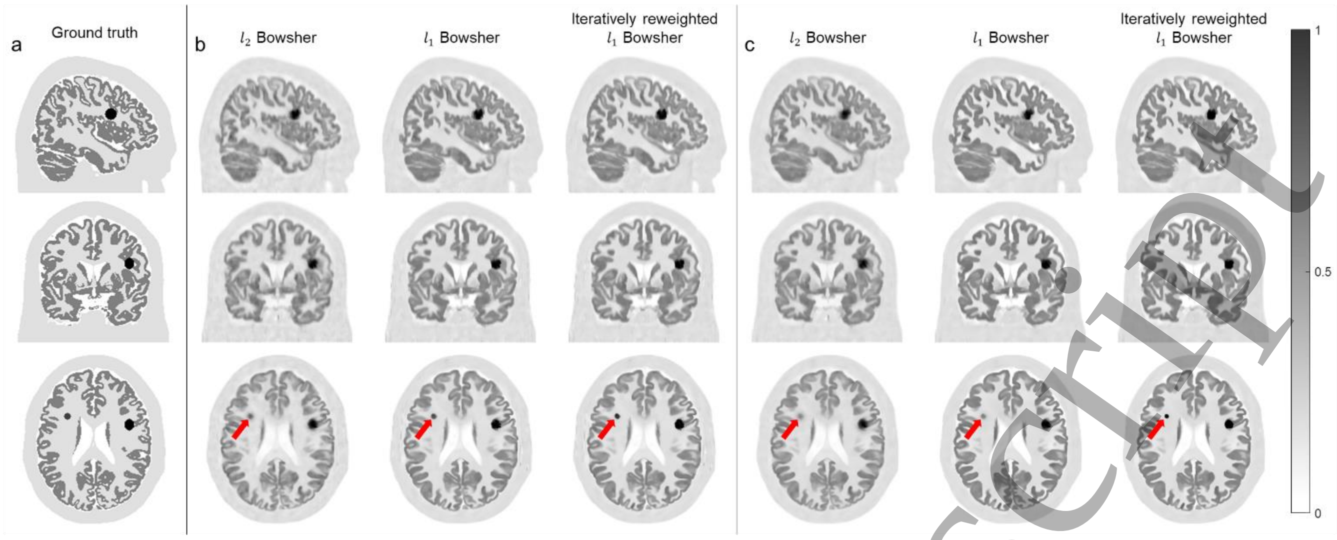

Figure 3. The representative reconstructed images using original $\boldsymbol{l_2}$ Bowsher prior, proposed $\boldsymbol{l_1}$ Bowsher prior and its iterative reweighting variation under low-level noise circumstances (total 7.0×10^7 prompt counts). Red arrow indicates the position of small lesions. a. Ground truth, b. 10 voxels selection in the given patch (nearest 80 voxels) and c. 20 voxels selections in the given patch.

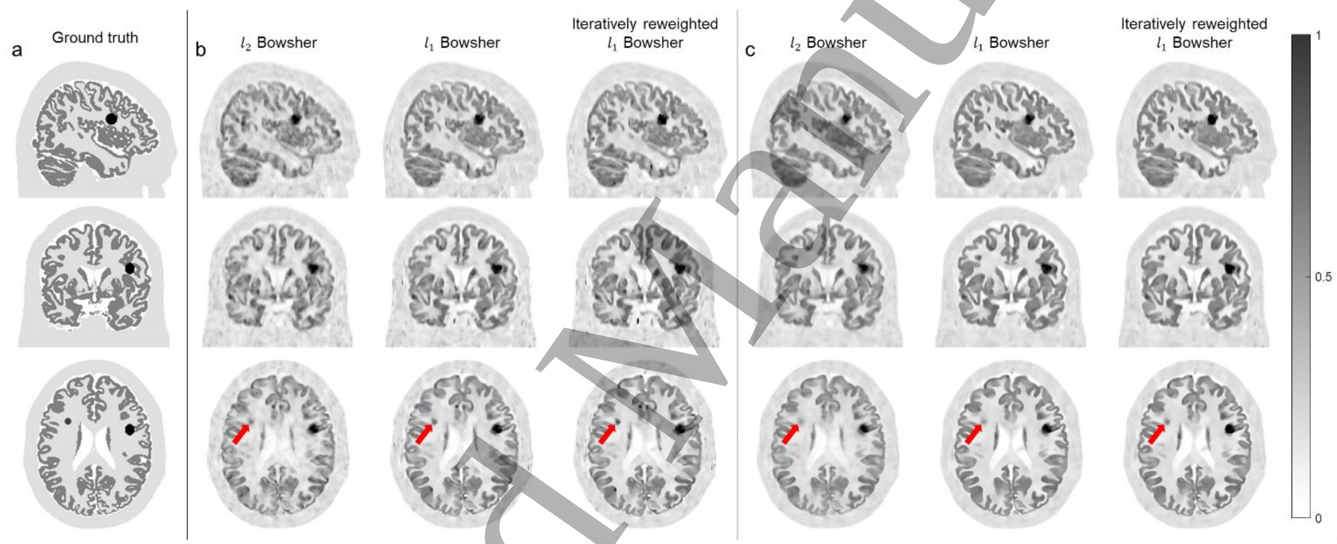

Figure 4. The representative reconstructed images using original l_2 Bowsher prior, proposed l_1 Bowsher prior and its iterative reweighting variation under high-level noise circumstances (total 1.4×10^7 prompt counts). Red arrow indicates the position of small lesions. a. Ground truth, b. 10 voxels selection in the given patch (nearest 80 voxels) and c. 20 voxels selections in the given patch.

pixel value was obtained because the white matter exhibits uniform FDG uptake. We focused on two lesions (large and small) with high uptake in the patient with head and neck cancer.

3. Results

3.1 Simulation with Brain Phantom

The proposed l_1 Bowsher prior recovered the detailed structure of the GM and tumors well even under high-level noise circumstances. Figures 3 and 4 show the representative reconstruction results for different noise levels (low and high) and patch sizes (10 and 20 voxels). Fifth regularization parameters (0.1×2^4 for l_2 Bowsher prior and 0.1×2^5 for others) were chosen for the visualization. The PET intensity in the large lesion was also less smeared with the proposed methods. Although the original Bowsher prior over-smoothed the small tumor, the proposed l_1 Bowsher prior methods preserved the shape and intensity of the small lesion. Figures 5 and 6 show the bias map for different noise levels (low and high). The proposed methods yielded lower bias under both low and high-level noise circumstances. Moreover, the bias of artificial lesions, especially for small lesion, were lower in the iteratively reweighted l_1 Bowsher prior than all the other reconstruction methods. This phenomenon also can be observed in the bias-STD plot (Figure 7) for each simulated region (Gray matter, white matter, large lesion and small

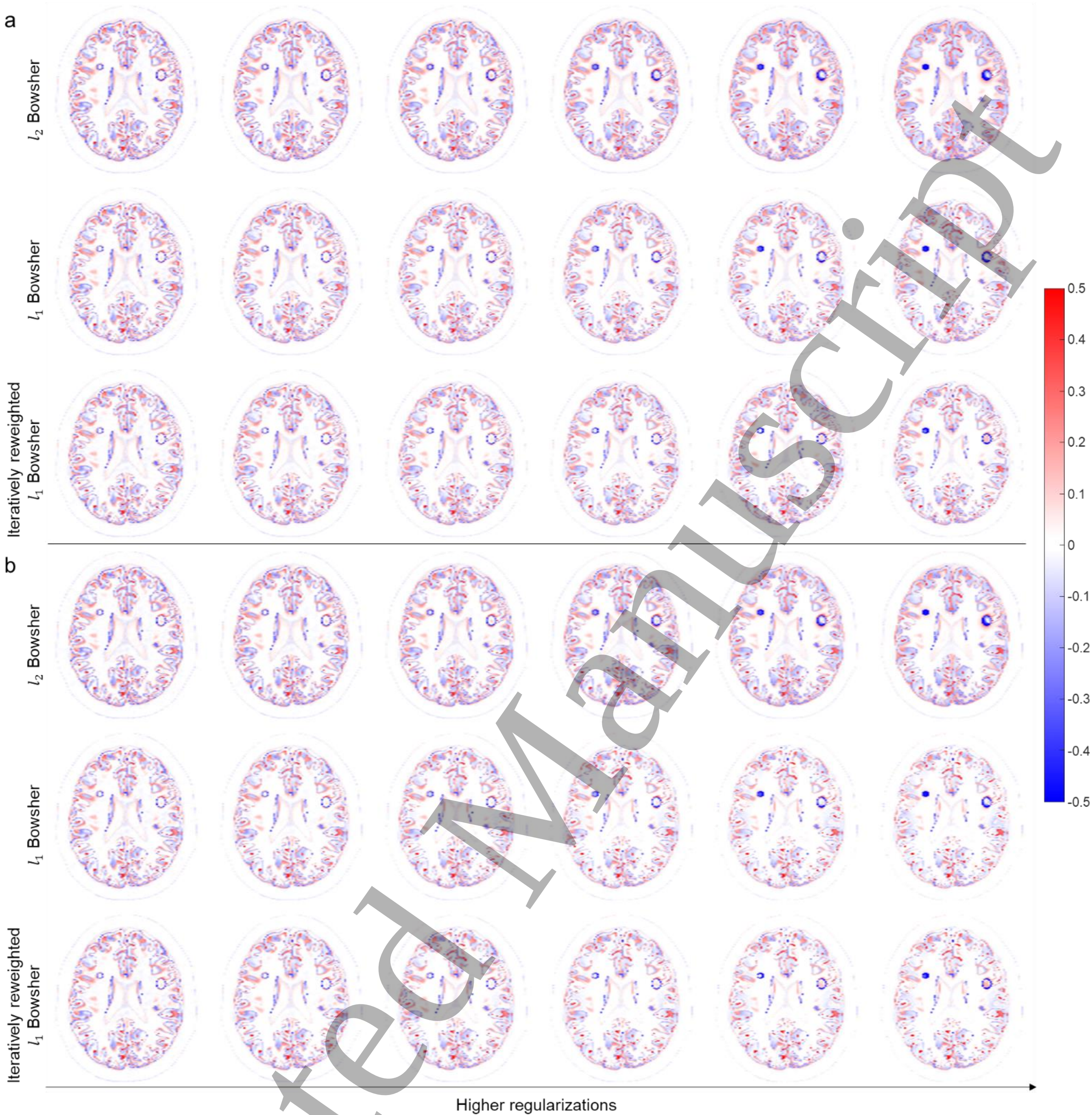

Figure 5. Bias map of computer simulation results using original l_2 Bowsher prior, proposed l_1 Bowsher prior and its iterative reweighting variation under the low-level noise circumstances (total 7.0×10^7 prompt counts). a. 10 voxels selection in the given patch (nearest 80 voxels) and b. 20 voxels selection in the given patch.

lesion). Both the bias and STD were suppressed by the proposed l_1 Bowsher prior methods as the regularization parameter increases, however, the bias became greater with l_2 Bowsher prior. The bias for artificial lesions with iteratively reweighted l_1 Bowsher prior yielded the lowest value.

3.2 Human Data

As described in the Methods section, 20 voxels were selected in the patch for all reconstruction, which showed the better performance. With respect to the human data, the proposed l_1 Bowsher prior methods outperformed the original l_2 Bowsher prior in preserving the detailed structures while suppressing the noise. As depicted in the [^{18}F]FDG PET image of the healthy volunteer (Figure 8), the original l_2 Bowsher prior with high regularization parameters yielded a blurred

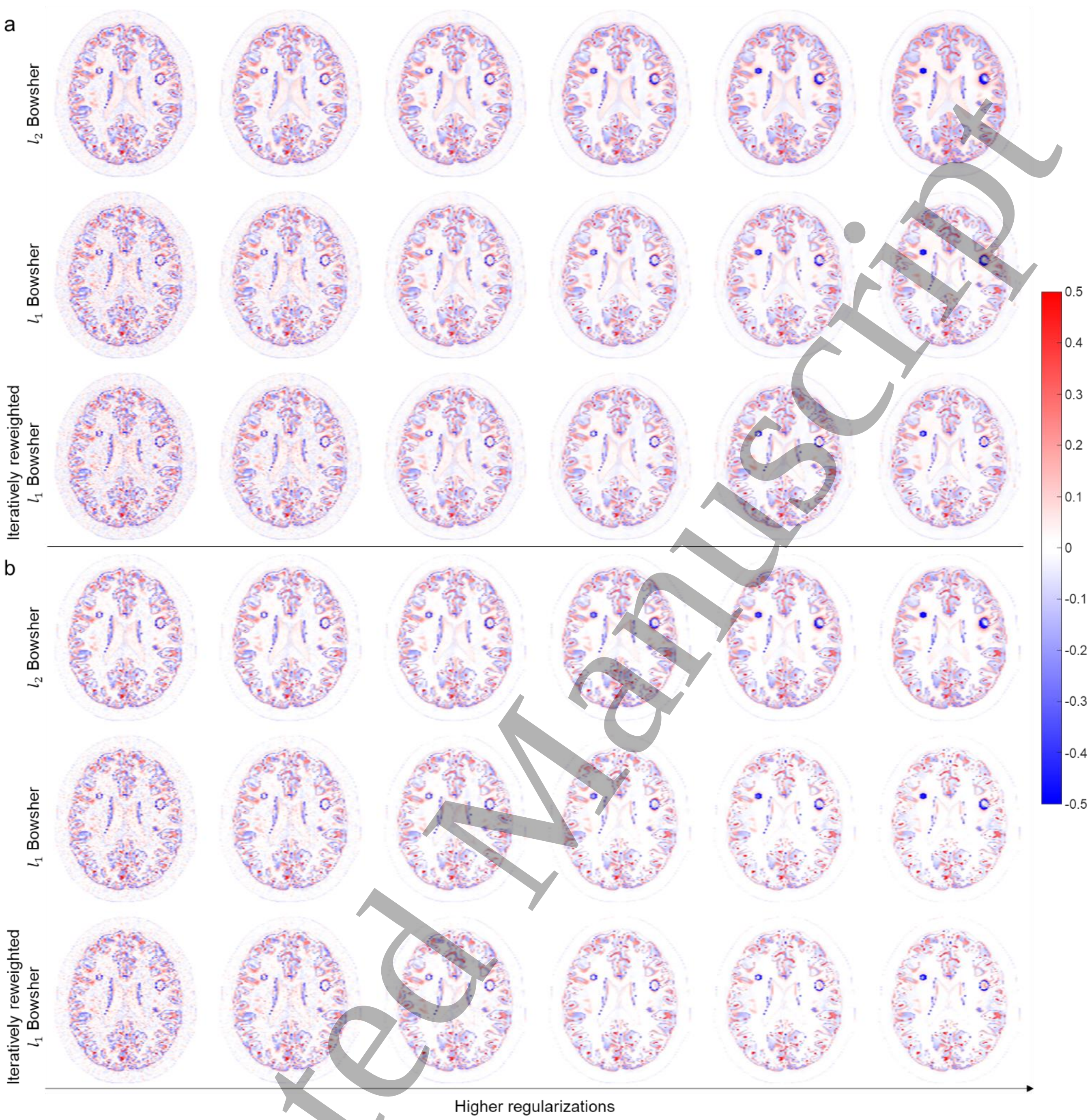

Figure 6. Bias map of computer simulation results using original l_2 Bowsher prior, proposed l_1 Bowsher prior and its iterative reweighting variation under the high-level noise circumstances (total 1.4×10^7 prompt counts). a. 10 voxels selection in the given patch (nearest 80 voxels) and b. 20 voxels selection in the given patch.

shape and decreased the uptake in some gyri, as highlighted with red boxes, and most subcortical regions, such as in the striatum. However, the proposed l_1 Bowsher prior did not indicate such an adverse impact of the high regularization parameter.

These findings were confirmed in the quantitative analysis summarized in Figure 9 that shows the STD of uptake in white matter versus the mean uptake in four different gray matter regions (frontal lobe, cingulate cortex, superior parietal gyrus, and lateral temporal gyrus). Iteratively reweighted l_1 Bowsher prior showed higher uptake than other methods with similar STD. The uptake in the gray matter decreased as the regularization parameter increased when the original Bowsher prior was used. However, the uptake level was more constant with the l_1 Bowsher prior.

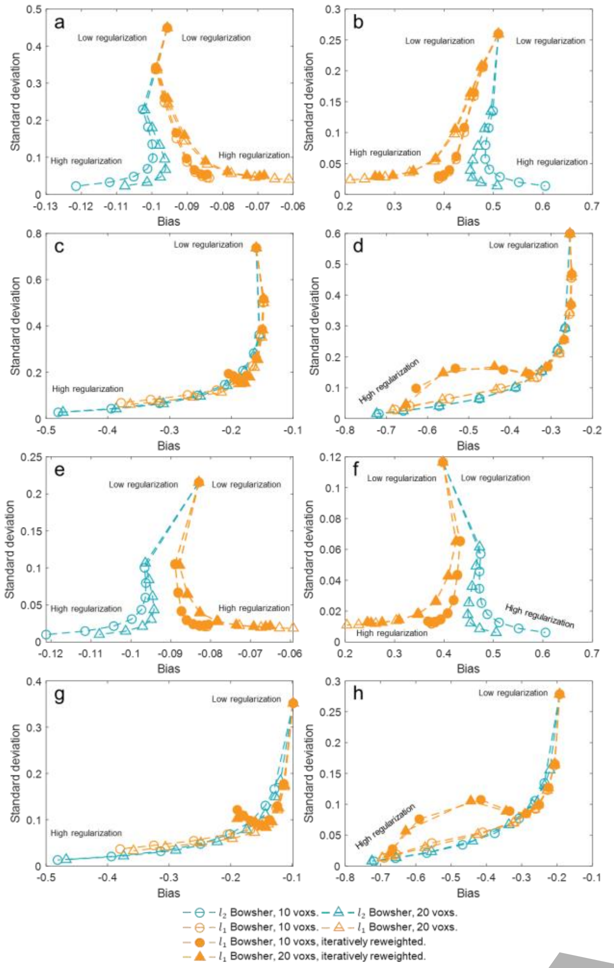

Figure 7. Bias-STD plot of computer simulation results using original l_2 Bowsher prior, proposed l_1 Bowsher prior and its iterative reweighting variation for four regions and two noise levels. a. Gray matter (GM) for low-level noise, b. white matter (WM) for low-level noise, c. large lesion for low-level noise, d. small lesion for low-level noise, e. GM for high-level noise, f. WM for high-level noise, g. large lesion for high-level noise and h. small lesion for high-level noise. Plots start from zero regularization.

In addition, the l_1 Bowsher prior methods better preserved the increased PET uptake in the small lesion that showed low contrast in the structural T1 MRI used for the guiding anatomy as compared to the original Bowsher prior (Figure 10). It should be noted that the T2 MRI images presented in Figure 10 as supporting evidence of the malignancy of the tumor were not used in the anatomy-guided reconstruction.

4. Discussions

In this study, we propose an MRI-guided regularized PET reconstruction based on a new l_1 Bowsher prior and its application with the iterative reweighting scheme. In these methods, (12) plays a pivotal role in incorporating side information into the reconstruction process. The proposed proximal operator described in this equation is similar to the soft-thresholding operator used in the Lasso regression (Tibshirani 1996). Both operators commonly cause the sparsity of their solution, leading to better detectability of small lesions.

As demonstrated by the simulation and real data, better contrast between the background and small lesions with abnormal PET uptake was obtained by applying the proposed l_1 Bowsher prior and its iterative reweighting variation as compared to the original l_2 Bowsher prior. The performance of the proposed method was particularly superior when such lesions are not shown in the MRI used for the regularized PET reconstruction (Figures 3, 4, and 10). The original l_2 Bowsher prior leads to smeared PET intensity in small lesions when there is low contrast between the tumor and surrounding tissue in the anatomical prior. This is because, in (2) and (6), tumor voxels are not distinguishable based on the difference in voxels in the anatomical image. However, the proposed l_1 Bowsher prior enables to preserve the edges between the tumor and the surrounding tissue in PET because of the intrinsic edge-preserving property of the prior based on the l_1-norm. Moreover, enhanced sparseness by iterative reweight enlarged this effect. The proposed method also showed sharper boundaries than the original method when the boundaries of the MRI structures were blurred (striatum in Figure 8; Note that it is unknown whether the boundary should be sharp or the sharpness is an artifact of the l_1-norm). In addition, the proposed method demonstrated smaller bias and less hyper-parameter dependency in PET intensity estimation in the regions (GM and WM) with matched anatomical boundaries in PET and MRI (Figures 5 and 6). The proposed l_1 Bowsher prior methods well preserve the mean uptake level of ROI even with the high regularization parameter although there is a trade-off between the standard deviation and the mean uptake level of ROI in the original l_2 Bowsher prior.

Introducing iterative reweight scheme in the reconstruction with l_1 Bowsher prior allowed better visualization of small hot regions compared to the vanilla l_1 Bowsher prior as shown in Figures 3, 4, and 10. It would be because, as mentioned above, the iterative reweighting enhances the sparseness of the prior. It originally aimed to approximate the optimization process from the l_1 relaxation to the l_0 minimization (Candes *et al.* 2008). The sparsity of the intensity difference defined in (2) is important when the matched anatomical information is not provided because the uptake of these regions will be smoothed by the prior. However, if the optimization algorithm can preserve the sparseness, hot uptake surrounded by warm background can be preserved as shown in Figures 3, 4 and 10. However, promoting sparseness of the prior sometimes leads to a side effect. Under the high-level noise circumstance, it is not an easy task to distinguish noise and true signal in the image, resulting in worse denoising performance compared to the vanilla l_1 Bowsher prior (Figure 10, third row). The ϵ is another control parameter and finding optimal settings

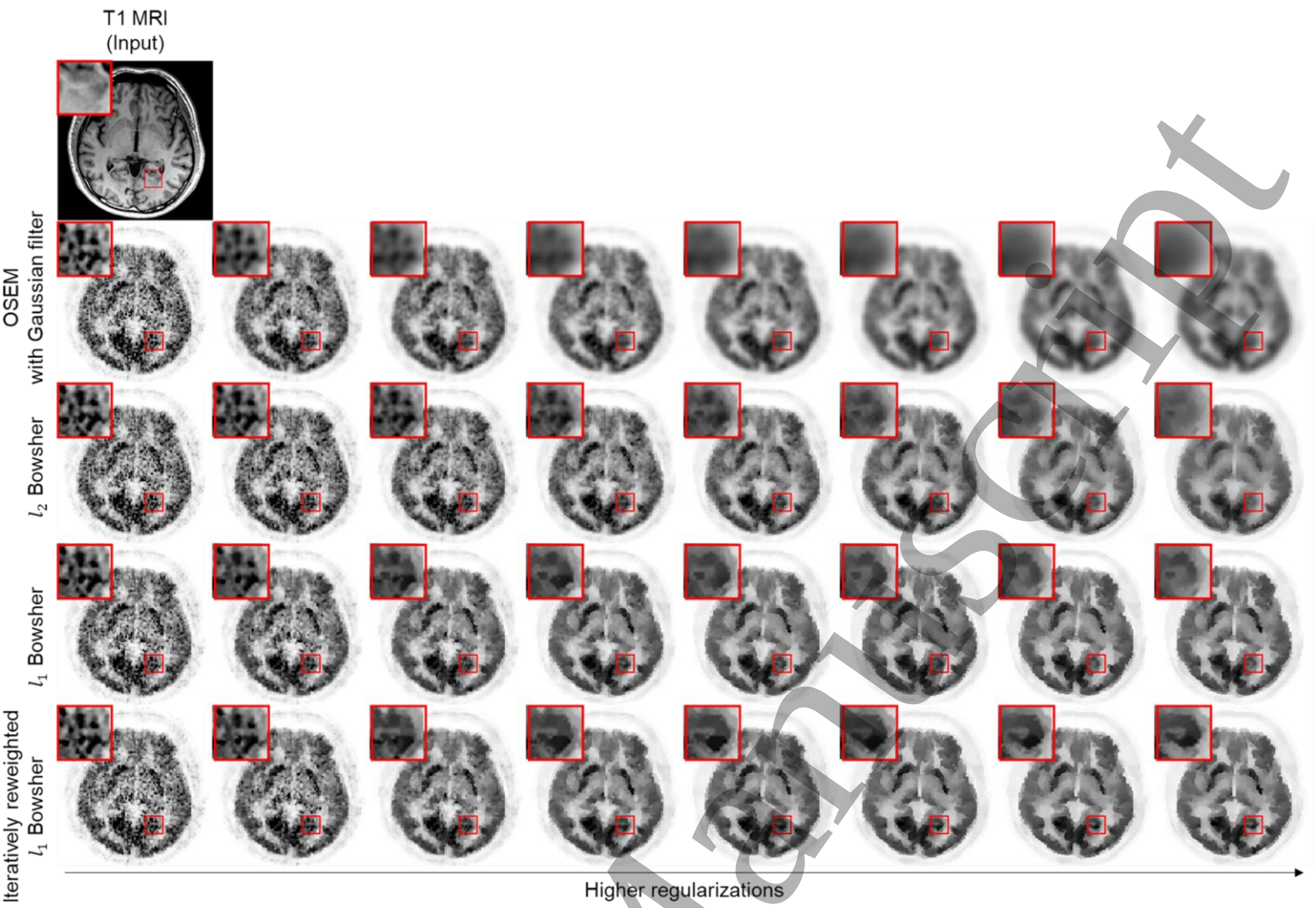

Figure 8. Reconstructed brain [^{18}F]FDG images of a healthy volunteer. The magnified region is highlighted by the red box. The FWHMs of Gaussian filter were from 1 mm to 8 mm, and the results of the first regularization parameter among 9 implementations ($0.1{\times}2^{-1}$ for l_2 Bowsher and $0.1{\times}2^{0}$ for others) were not shown.

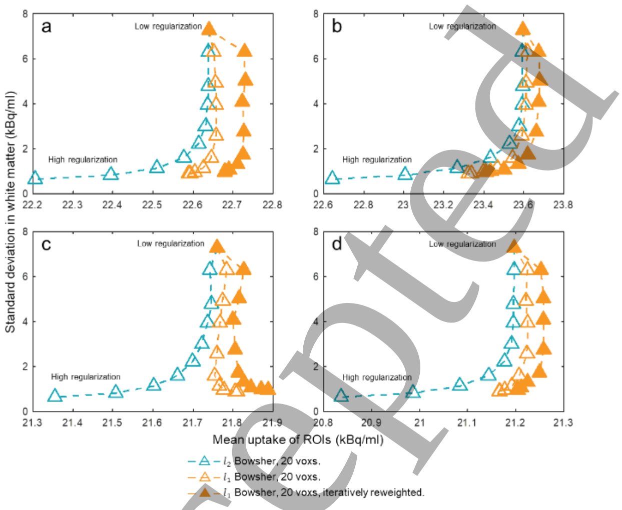

Figure 9. Quantitative analysis on four different regions in brain [^{18}F]FDG PET images. a. frontal lobe, b. cingulate cortex, c. superior parietal gyrus and d. lateral temporal gyrus. Plots start from zero regularization.

including the number of patches and the number of selected voxels is a future direction of the research.

Although this is the first study to apply l_1-norm to the Bowsher prior as far as we know, the l_1-norm has been investigated extensively in the more general context of Bayesian (or penalized likelihood) image reconstruction. Various total variation (TV) minimization approaches have been proposed to improve the image quality of CT and emission tomography (Rudin *et al.* 1992, Sawatzky *et al.* 2008, Guo *et al.* 2009, Ahn *et al.* 2012, Wang *et al.* 2014, Ehrhardt *et al.* 2019, Knoll *et al.* 2017, Niu *et al.* 2014, Gu *et al.* 2018, Burger *et al.* 2014, Son *et al.* 2014). In the most TV approaches, the l_1-norm of the discretized image gradient is used to regularize the fidelity optimization while preserving the edge information. In general, the TV-l_1 model suppresses noise in the uniform region more effectively than the l_2-norm regularization. However, TV-l_1 regularization often causes so-called "staircase" artifacts, yielding multiple flat regions separated by sharp boundaries. For PET images with high-level noise and low spatial resolution, the edges produced by TV prior might be inaccurate.

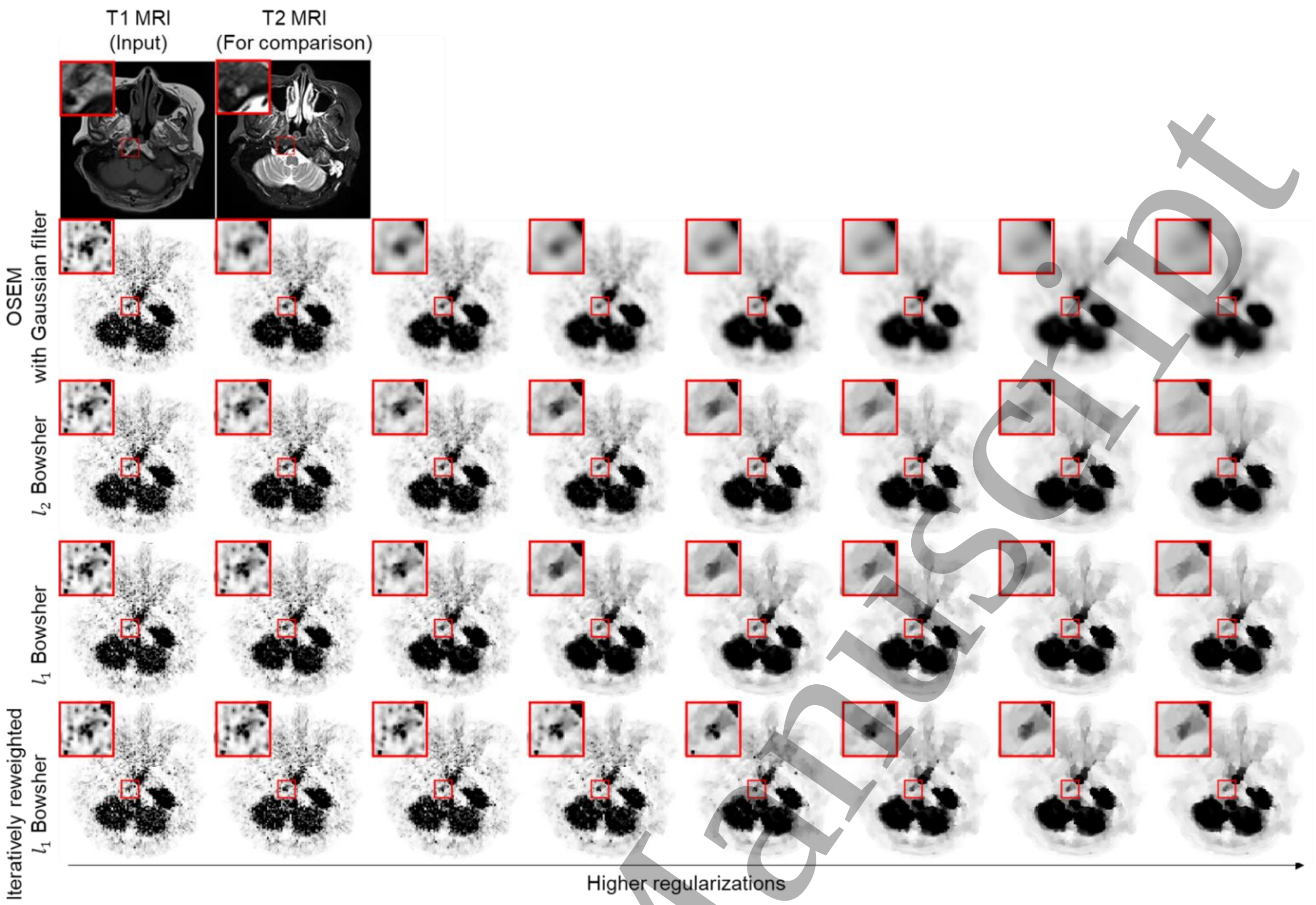

Figure 10. Reconstructed [^{18}F]FDG images of a head and neck cancer patient. Only T1 MR image was used for the anatomy-guided reconstruction. The FWHMs of Gaussian filter were from 1 mm to 8 mm, and the results of the first regularization parameter among 9 implementations (0.1×2^{-1} for l_2 Bowsher and 0.1×2^{0} for others) were not shown.

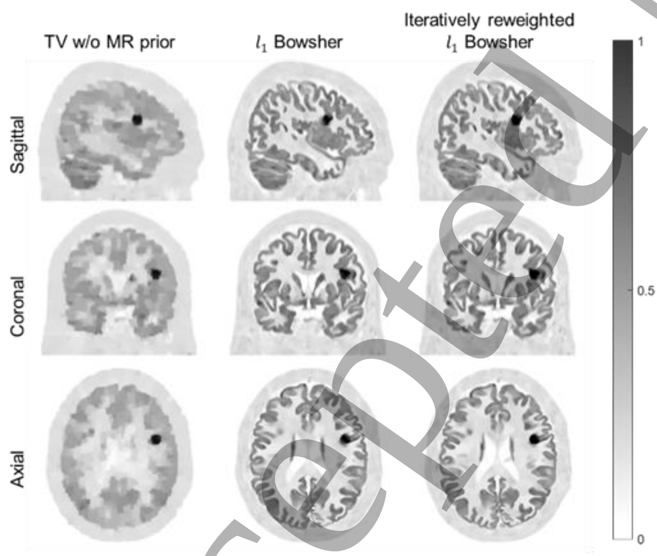

Figure 11. Comparison between TV prior using EM-TV algorithm (Sawatzky *et al.*) and l_1 Bowsher methods. Low count simulation data was used and the regularization parameter for TV prior was 4 and the others were the same with Figure 4 (0.1×2^5).

The shortcomings of the l_1-norm regularization could be alleviated by the anatomical prior because the edge-preserving property of the l_1-norm regularization is guided by the anatomical prior (Figure 11). However, the "staircase" artifacts still appear when the regularization parameter is high (Figures 8 and 10), so further investigations to mitigate the artifacts are needed. Another significant difference in this study from others is that the l_1-norm was applied to the Gibbs prior calculated using the distance between local neighboring pixels. Although many previous studies have used l_1-norm with TV prior (Beck and Teboulle 2009a, Goldstein and Osher 2009, Sawatzky *et al.* 2008), there are relatively few studies on l_1-norm regularization with other prior than TV for solving the inverse problems (Wang *et al.* 2012, Liu *et al.* 2019). Wang *et al*. applied l_1-norm directly to the solution vector and used the barrier function as well as the projection method to find the update equation. Liu *et al.* used both TV and l_1-norm of the image vector in the cost function, which is minimized by a fast iterative shrinkage-thresholding algorithm (FISTA) (Beck and Teboulle 2009b). Both studies examined l_1-norm of the image vector, however, we modified potential function of the Gibbs prior from l_2-norm to l_1-norm. To minimize the proposed l_1 Bowsher prior, modified proximal gradient was calculated and combined with the ordinary EM update.

A limited number of segmentation-free anatomy-guided reconstruction methods have been proposed so far. One of them is the kernel method that assumes that the PET image is a linear function of the transformed anatomical features from the MRI. The kernel-based method that encodes prior information into the PET projection model is another (Hutchcroft *et al.* 2016). In this method, patch-based MR image features are employed to form the kernel matrix. Because this kernel method incorporates anatomical information in the maxim likelihood formulation rather than in the penalized likelihood framework, it is amenable to ordered subsets. However, this approach also suffers from the over-smoothing of PET intensity in the regions where the PET uptake pattern differs from the anatomical side information. A parallel level set (PLS) prior between the anatomical and reconstructed PET image (Ehrhardt *et al.* 2016) is more robust to the discrepancy between the PET uptake pattern and anatomical side information. Nevertheless, using a differentiable prior requires well-defined parameter settings during the optimization process. Moreover, there is a report that the asymmetrical Bowsher prior shows better performance than the PLS method (Schramm *et al.* 2018). Our proposed method is also based on the Bowsher prior, but we have incorporated it into the edge-preserving property of the l_1-norm. The optimization of the cost function is easy to implement using the proximal gradient algorithm and the closed-form solution of the proximal operator. Similar to the original Bowsher prior, the proposed method can be applied to multiple MRI pulse sequences. As presented in Figure 10, various MR images with various pulse sequences were acquired during routine PET/MRI studies. The weight used in (6) can be modified by combining information from the multiple MRI pulse sequences.

In this study, we applied the FORE algorithm to pre-corrected sinogram for scatter, random and attenuation to reconstruct real patient PET images using proposed prior models (Defrise *et al.* 1997). This would cause problems in terms of performance such as degraded sensitivity and resolution. However, the same optimization schemes can be used by replacing only the projection and backprojection parts to the 3D methods.

Another limitation of this study is that spatial resolution of the PET scanner was not modeled during image reconstruction. To reconstruct accurately the activities of the various regions, anatomical information can be helpful. However, accurate modeling of spatial resolution is generally also important. In addition, the levels of blur and accuracy of modeling might substantially affect the absolute and relative performance of l_1- and l_2-norm priors. The spatial resolution would be accurately modeled for the computer-simulation studies because there isn't much blur in the simulation projection data and the reconstruction uses that same model of only very minor blur. This accurate modeling of spatial resolution is one reason why all priors were able to drive the bias to low values in the computer-simulation study.

An approach to anatomy-guided functional image enhancement using deep neural networks is emerging, as deep learning is outperforming conventional approaches based on numerical and statistical signal processing in several different areas (Vincent *et al.* 2010, Xie *et al.* 2012, Krizhevsky *et al.* 2012, Agostinelli *et al.* 2013, Simonyan and Zisserman 2014, He *et al.* 2015, Dey *et al.* 2018, Mansour 2018). Beyond simple noise reduction by recovering high-statistics PET images from the pair of anatomical image and low-statistics PET scan, more sophisticated concepts such as super-resolution and partial volume correction of PET are now being handled using deep learning (Rigie *et al.* 2018, Song *et al.* 2019, Xu *et al.* 2017). Generation of anatomical images or the standard template from PET data using deep neural networks proposed for PET spatial normalization and attenuation correction (Choi and Lee 2018, Kang *et al.* 2018, Hwang *et al.* 2019, Hwang *et al.* 2018) can be potentially utilized for reducing PET noise and enhancing its spatial resolution and image contrast. These methods have the potential for providing anatomical side information to be used for anatomy-guided PET image reconstruction. Including pre-trained deep neural networks that utilize anatomical side information for enhancing PET into PET-iterative reconstruction would also be an interesting future research topic (Adler and Öktem 2018, Gupta *et al.* 2018, Kim *et al.* 2018, Gong *et al.* 2018).

5. Conclusions

In this study, we proposed an l_1-norm-based Bowsher prior. The proximal gradient algorithm was exploited to solve the penalized likelihood function, and a modified proximal operator for EM-based reconstruction was provided. The iterative reweighting scheme that enforces sparseness of the prior improved both qualitative and quantitative results. The results from the computer simulation support the fact that our proposed methods yield a better quantification of tumors as well as the GM and WM than the previous approaches. Besides, clinical data suggest that the proposed prior method might better visualize small regions than l_2 Bowsher prior. Therefore, these methods will be useful for improving the PET image quality based on the anatomical information provided by other anatomical imaging systems. Nevertheless, further evaluation of the proposed method with more clinical data and spatial resolution modeling will be necessary.

Acknowledgments

This work was supported by grants from the National Research Foundation of Korea (NRF) funded by the Korean Ministry of Science, ICT and Future Planning (grant no. NRF-2016R1A2B3014645).

References

Adler J and Öktem O 2018 Learned primal-dual reconstruction *IEEE Trans. Med. Imaging* **37** 1322-32

Agostinelli F, Anderson M R and Lee H 2013 Adaptive multi-column deep neural networks with application to robust image denoising *Adv. Neural Inf. Process. Syst.* pp 1493-501

Ahn S, Kim S M, Son J, Lee D S and Sung Lee J 2012 Gap compensation during PET image reconstruction by constrained, total variation minimization *Med. Phys.* **39** 589-602

An H J, Seo S, Kang H, Choi H, Cheon G J, Kim H-J, Lee D S, Song I C, Kim Y K and Lee J S 2016 MRI-based attenuation correction for PET/MRI using multiphase level-set method *J. Nucl. Med.* **57** 587-93

Artzy E, Elfving T and Herman G T 1979 Quadratic optimization for image reconstruction, II *Computer Graphics and Image Processing* **11** 242-61

Ashburner J and Friston K J 2005 Unified segmentation *Neuroimage* **26** 839-51

Baete K, Nuyts J, Van Laere K, Van Paesschen W, Ceyssens S, De Ceuninck L, Gheysens O, Kelles A, Van den Eynden J and Suetens P 2004a Evaluation of anatomy based reconstruction for partial volume correction in brain FDG-PET *Neuroimage* **23** 305-17

Baete K, Nuyts J, Van Paesschen W, Suetens P and Dupont P 2004b Anatomical-based FDG-PET reconstruction for the detection of hypo-metabolic regions in epilepsy *IEEE Trans. Med. Imaging* **23** 510-9

Bai B, Li Q and Leahy R M 2013 Magnetic resonance-guided positron emission tomography image reconstruction *Semin. Nucl. Med.* **43** 30-44

Beck A and Teboulle M 2009a Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems *IEEE Transactions on Image Processing* **18** 2419-34

Beck A and Teboulle M 2009b A fast iterative shrinkage-thresholding algorithm for linear inverse problems *SIAM journal on imaging sciences* **2** 183-202

Bowsher J E, Johnson V E, Turkington T G, Jaszczak R J, Floyd C E and Coleman R E 1996 Bayesian reconstruction and use of anatomical a priori information for emission tomography *IEEE Trans. Med. Imaging* **15** 673-86

Bowsher J E, Yuan H, Hedlund L W, Turkington T G, Akabani G, Badea A, Kurylo W C, Wheeler C T, Cofer G P and Dewhirst M W 2004 Utilizing MRI information to estimate F18-FDG distributions in rat flank tumors *IEEE Symposium Conference Record Nuclear Science 2004.* pp 2488-92

Burger M, Müller J, Papoutsellis E and Schönlieb C B 2014 Total variation regularization in measurement and image space for PET reconstruction *Inverse Problems* **30** 105003

Candes E J, Wakin M B and Boyd S P 2008 Enhancing sparsity by reweighted ℓ 1 minimization *Journal of Fourier analysis and applications* **14** 877-905

Chambolle A, Caselles V, Cremers D, Novaga M and Pock T 2010 An introduction to total variation for image analysis *Theoretical foundations and numerical methods for sparse recovery* **9** 227

Choi H and Lee D S 2018 Generation of structural MR images from amyloid PET: Application to MR-less quantification *J. Nucl. Med.* **59** 1111-7

Cocosco C A, Kollokian V, Kwan R K-S, Pike G B and Evans A C 1997 Brainweb: Online interface to a 3D MRI simulated brain database *Neuroimage*

Defrise M, Kinahan P E, Townsend D W, Michel C, Sibomana M and Newport D F 1997 Exact and approximate rebinning algorithms for 3-D PET data *IEEE Trans. Med. Imaging* **16** 145-58

Delso G, Fürst S, Jakoby B, Ladebeck R, Ganter C, Nekolla S G, Schwaiger M and Ziegler S I 2011 Performance measurements of the Siemens mMR integrated whole-body PET/MR scanner *J. Nucl. Med.* **52** 1914-22

Dey D, Chaudhuri S and Munshi S 2018 Obstructive sleep apnoea detection using convolutional neural network based deep learning framework *Biomedical Engineering Letters* **8** 95-100

Ehrhardt M J, Markiewicz P, Liljeroth M, Barnes A, Kolehmainen V, Duncan J S, Pizarro L, Atkinson D, Hutton B F, Ourselin S, Thielemans K and Arridge S R 2016 PET Reconstruction With an Anatomical MRI Prior Using Parallel Level Sets *IEEE Trans. Med. Imaging* **35** 2189-99

Ehrhardt M J, Markiewicz P and Schönlieb C-B 2019 Faster PET reconstruction with non-smooth priors by randomization and preconditioning *Phys. Med. Biol.* **64** 225019

Esser E 2009 Applications of Lagrangian-based alternating direction methods and connections to split Bregman *CAM report* **9** 31

Fessler J A and Hero A O 1995 Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms *IEEE Transactions on Image Processing* **4** 1417-29

Goffin K, Van Paesschen W, Dupont P, Baete K, Palmini A, Nuyts J and Van Laere K 2010 Anatomy-based reconstruction of FDG-PET images with implicit partial volume correction improves detection of hypometabolic regions in patients with epilepsy due to focal cortical dysplasia diagnosed on MRI *Eur. J. Nucl. Med. Mol. Imaging* **37** 1148-55

Goldstein T and Osher S 2009 The Split Bregman Method for L1-Regularized Problems *SIAM Journal on Imaging Sciences* **2** 323-43

Gong K, Guan J, Kim K, Zhang X, Yang J, Seo Y, El Fakhri G, Qi J and Li Q 2018 Iterative PET image reconstruction using convolutional neural network representation *IEEE Trans. Med. Imaging* **38** 675-85

Gourion D and Noll D 2002 The inverse problem of emission tomography *Inverse Problems* **18** 1435

Gu C, Zeng D, Lin J, Li S, He J, Zhang H, Bian Z, Niu S, Zhang Z, Huang J, Chen B, Zhao D, Chen W and Ma J 2018 Promote quantitative ischemia imaging via myocardial perfusion CT iterative reconstruction with tensor total generalized variation regularization *Phys. Med. Biol.* **63** 125009

Guo H, Renaut R A, Chen K and Reiman E 2009 FDG–PET parametric imaging by total variation minimization *Comput. Med. Imaging Graph.* **33** 295-303

Gupta H, Jin K H, Nguyen H Q, McCann M T and Unser M 2018 CNN-based projected gradient descent for consistent CT image reconstruction *IEEE Trans. Med. Imaging* **37** 1440-53

He K, Zhang X, Ren S and Sun J 2015 Deep Residual Learning for Image Recognition *arXiv:1512.03385*

Hebert T and Leahy R 1989 A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors *IEEE Trans. Med. Imaging* **8** 194-202

Hutchcroft W, Wang G, Chen K T, Catana C and Qi J 2016 Anatomically-aided PET reconstruction using the kernel method *Phys. Med. Biol.* **61** 6668

Hwang D, Kang S K, Kim K Y, Seo S, Paeng J C, Lee D S and Lee J S 2019 Generation of PET attenuation map for whole-body time-of-flight 18F-FDG PET/MRI using a deep neural network trained with simultaneously reconstructed activity and attenuation maps *J. Nucl. Med.* jnumed. 118.219493

Hwang D, Kim K Y, Kang S K, Seo S, Paeng J C, Lee D S and Lee J S 2018 Improving the Accuracy of Simultaneously Reconstructed Activity and Attenuation Maps Using Deep Learning *J. Nucl. Med.* **59** 1624-9

Judenhofer M S, Wehrl H F, Newport D F, Catana C, Siegel S B, Becker M, Thielscher A, Kneilling M, Lichy M P and Eichner M 2008 Simultaneous PET-MRI: a new approach for functional and morphological imaging *Nat. Med.* **14** 459

Kang S K, Seo S, Shin S A, Byun M S, Lee D Y, Kim Y K, Lee D S and Lee J S 2018 Adaptive template generation for amyloid PET using a deep learning approach *Hum. Brain Mapp.* **39** 3769-78

Kang S K, Shin S A, Seo S, Byun M S, Lee D Y, Kim Y K, Lee D S and Lee J S 2021 Deep Learning-Based 3D Inpainting of Brain MR Images *Sci. Rep.* **11** 1673

Kaufman L 1993 Maximum likelihood, least squares, and penalized least squares for PET *IEEE Trans. Med. Imaging* **12** 200-14

Kim K, Wu D, Gong K, Dutta J, Kim J H, Son Y D, Kim H K, El Fakhri G and Li Q 2018 Penalized PET reconstruction using deep learning prior and local linear fitting *IEEE Trans. Med. Imaging* **37** 1478-87

Knoll F, Holler M, Koesters T, Otazo R, Bredies K and Sodickson D K 2017 Joint MR-PET Reconstruction Using a Multi-Channel Image Regularizer *IEEE Trans. Med. Imaging* **36** 1-16

Ko G B, Yoon H S, Kim K Y, Lee M S, Yang B Y, Jeong J M, Lee D S, Song I C, Kim S-k and Kim D 2016 Simultaneous multiparametric PET/MRI with silicon photomultiplier PET and ultra-high-field MRI for small-animal imaging *J. Nucl. Med.* **57** 1309-15

Krizhevsky A, Sutskever I and Hinton G E 2012 Imagenet classification with deep convolutional neural networks *Adv. Neural Inf. Process. Syst.* pp 1097-105

Lange K and Carson R 1984 EM reconstruction algorithms for emission and transmission tomography *J. Comput. Assist. Tomogr.* **8** 306-16

Levin C S, Maramraju S H, Khalighi M M, Deller T W, Delso G and Jansen F 2016 Design features and mutual compatibility studies of the time-of-flight PET capable GE SIGNA PET/MR system *IEEE Trans. Med. Imaging* **35** 1907-14

Liu T, Rong J, Gao P, Pu H, Zhang W, Zhang X, Liang Z and Lu H 2019 Regularized reconstruction based on joint L1 and total variation for sparse-view cone-beam X-ray luminescence computed tomography *Biomed. Opt. Express* **10** 1-17

Louis A K and Natterer F 1983 Mathematical problems of computerized tomography *Proceedings of the IEEE* **71** 379-89

Mansour R F 2018 Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy *Biomedical Engineering Letters* **8** 41-57

Mehranian A, Belzunce M A, Niccolini F, Politis M, Prieto C, Turkheimer F, Hammers A and Reader A J 2017 PET image reconstruction using multi-parametric anato-functional priors *Phys. Med. Biol.* **62** 5975-6007

Niu S, Gao Y, Bian Z, Huang J, Chen W, Yu G, Liang Z and Ma J 2014 Sparse-view x-ray CT reconstruction via total generalized variation regularization *Phys. Med. Biol.* **59** 2997-3017

Novosad P and Reader A J 2016 MR-guided dynamic PET reconstruction with the kernel method and spectral temporal basis functions *Phys. Med. Biol.* **61** 4624-44

Nuyts J, Baete K, Bequé D and Dupont P 2005 Comparison between MAP and postprocessed ML for image reconstruction in emission tomography when anatomical knowledge is available *IEEE Trans. Med. Imaging* **24** 667-75

Nuyts J, Beque D, Dupont P and Mortelmans L 2002 A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography *IEEE Trans. Nucl. Sci.* **49** 56-60

Parikh N and Boyd S 2014 Proximal algorithms *Foundations and Trends® in Optimization* **1** 127-239

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

Qi J and Leahy R M 2006 Iterative reconstruction techniques in emission computed tomography *Phys. Med. Biol.* **51** R541-78

Rigie D, Schramm G, Vahle T, Shepherd T, Nuyts J and Boada F 2018 Approximating MRI-Based

Anatomically Guided PET Reconstruction with a Convolutional Neural Network *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)* pp 1-3

Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica D: Nonlinear Phenomena* **60** 259-68

Sangtae A and Fessler J A 2003 Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms *IEEE Trans. Med. Imaging* **22** 613-26

Sawatzky A, Brune C, Wubbeling F, Kosters T, Schafers K and Burger M 2008 Accurate EM-TV algorithm in PET with low SNR *2008 IEEE Nuclear Science Symposium Conference Record* pp 5133-7

Schramm G, Holler M, Rezaei A, Vunckx K, Knoll F, Bredies K, Boada F and Nuyts J 2018 Evaluation of Parallel Level Sets and Bowsher's Method as Segmentation-Free Anatomical Priors for Time-of-Flight PET Reconstruction *IEEE Trans. Med. Imaging* **37** 590-603

Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography *IEEE Trans. Med. Imaging* **1** 113-22

Simonyan K and Zisserman A 2014 Very deep convolutional networks for large-scale image recognition *arXiv preprint arXiv:1409.1556*

Son J, Kim S M and Lee J S 2014 A strategy to reduce blocky pattern and contrast loss in emission tomography reconstruction with reduced angular sampling and total variation minimization *Biomedical Engineering Letters* **4** 362-9

Song T-A, Chowdhury S R, Yang F and Dutta J 2019 Super-resolution PET imaging using convolutional neural networks *arXiv preprint arXiv:1906.03645*

Tang J and Rahmim A 2015 Anatomy assisted PET image reconstruction incorporating multi-resolution joint entropy *Phys. Med. Biol.* **60** 31-48

Tibshirani R 1996 Regression shrinkage and selection via the lasso *Journal of the Royal Statistical Society: Series B (Methodological)* **58** 267-88

Tikhonov A N 1963 On the solution of ill-posed problems and the method of regularization *Dokl. Akad. Nauk* pp 501-4

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B and Joliot M 2002 Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain *Neuroimage* **15** 273-89

Vincent P, Larochelle H, Lajoie I, Bengio Y and Manzagol P-A 2010 Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion *Journal of machine learning research* **11** 3371-408

Vunckx K, Atre A, Baete K, Reilhac A, Deroose C M, Laere K V and Nuyts J 2012 Evaluation of Three MRI-Based Anatomical Priors for Quantitative PET Brain Imaging *IEEE Trans. Med. Imaging* **31** 599-612

Vunckx K and Nuyts J 2010 Heuristic modification of an anatomical Markov prior improves its performance *IEEE Nuclear Science Symposuim & Medical Imaging Conference* pp 3262-6

Wang C, Hu Z, Shi P and Liu H 2014 Low dose PET reconstruction with total variation regularization *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* pp 1917-20

Wang G and Qi J 2012 Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization *IEEE Trans. Med. Imaging* **31** 2194-204

Wang Q, Wang H, Zhang R, Wang J, Zheng Y, Cui Z and Yang C 2012 Image reconstruction based on L1 regularization and projection methods for electrical impedance tomography *Rev. Sci. Instrum.* **83** 104707

Xie J, Xu L and Chen E 2012 Image denoising and inpainting with deep neural networks *Adv. Neural Inf. Process. Syst.* pp 341-9

Xu J, Gong E, Pauly J and Zaharchuk G 2017 200x low-dose PET reconstruction using deep learning *arXiv preprint arXiv:1712.04119*

Yoon H S, Ko G B, Kwon S I, Lee C M, Ito M, Song I C, Lee D S, Hong S J and Lee J S 2012 Initial results of simultaneous PET/MRI experiments with an MRI-compatible silicon photomultiplier PET scanner *J. Nucl. Med.* **53** 608-14