Incorporating Anatomical Side Information into PET Reconstruction Using Nonlocal Regularization

Van-Giang Nguyen and Soo-Jin Lee, *Member, IEEE*

***Abstract*— With the introduction of combined positron emission tomography (PET)/computed tomography (CT) or PET/magnetic resonance imaging (MRI) scanners, there is an increasing emphasis on reconstructing PET images with the aid of the anatomical side information obtained from X-ray CT or MRI scanners. In this paper, we propose a new approach to incorporating prior anatomical information into PET reconstruction using the nonlocal regularization method. The nonlocal regularizer developed for this application is designed to selectively consider the anatomical information only when it is reliable. As our proposed nonlocal regularization method does not directly use anatomical edges or boundaries which are often used in conventional methods, it is not only free from additional processes to extract anatomical boundaries or segmented regions, but also more robust to the signal mismatch problem that is caused by the indirect relationship between the PET image and the anatomical image. We perform simulations with digital phantoms. According to our experimental results, compared to the conventional method based on the traditional local regularization method, our nonlocal regularization method performs well even with the imperfect prior anatomical information or in the presence of signal mismatch between the PET image and the anatomical image.**

***Index Terms*— Positron emission tomography (PET), statistical image reconstruction, penalized likelihood, anatomical priors, nonlocal regularization.**

I. Introduction

POSITRON emission tomography (PET) has played an important role as one of the nuclear medicine imaging techniques that provide functional information about physiological processes in the human body. In PET a positron-emitting radioactive tracer isotope is introduced into the patient and localizes within one or more organs based on its biochemical properties. After an appropriate uptake period, the radioisotope undergoes positron emission decay and emits a positron, which travels a short distance before colliding and annihilating with an electron of a nearby atom. This annihilation produces pairs of gamma rays moving in approximately opposite directions [1], [2]. These gamma ray photons easily escape from the human body and are recorded by the detectors placed around the body and grouped into projection data. The two-dimensional (2-D) or 3-D distribution of radioisotope concentrations can be visualized by reconstructing from the projection data acquired over an adequate range of angles about the body.

Over the last two decades, statistical methods have been a topic for PET reconstruction, since they can not only conveniently incorporate a system model needed to account for physics of image formation, but also allow the incorporation of *a priori* information on the underlying source distribution. (An extensive overview of statistical reconstruction methods and references can be found in [3]). For maximum *a posteriori* (MAP) approaches, priors reflect the local spatial characteristics of the underlying source and regularize otherwise unstable maximum likelihood (ML) solutions. However, since most of the prior models are designed by the generic assumptions on the underlying source, they tend to work well only under specific conditions.

Recently, attempts have also been made to incorporate the prior anatomical information, which is obtained from high-resolution magnetic resonance (MR) or X-ray computed tomography (CT) images, into the statistical image reconstruction methods. (The prior models in these cases are non-generic; the prior information differs for different objects.). One popular approach is to use Markov random field (MRF) models in the context of a Bayesian MAP framework in which *a priori* anatomical information is modeled as a Gibbs distribution [4], [5]. Several different approaches, which include the Bayesian joint mixture model [6], the method of using the mutual information and joint entropy between the PET image and the anatomical image [7], the level set method [8], the method of using tissue composition model [9], and the minimum cross-entropy method [10], have also been proposed.

An important issue in these methods is that, there always remain the problems of position mismatch due to the position alignment error and signal mismatch due to the indirect relationship between the anatomical and PET images [4], [5], [8], [9]; the intensity changes in PET images are not consistent with those in anatomical images, which can cause artifacts in the reconstructed images. Thanks to the recent development of multimodal imaging systems which provide coregistered functional and anatomical images, such as combined PET/CT and PET/MRI systems, the position mismatch problem has been

Manuscript received September 13, 2012; revised January 16, 2013 and March 28, 2013; accepted May 13, 2013. Date of publication June 3, 2013; date of current version August 30, 2013. This work was supported by Basic Science Research Program of the National Research Foundation of Korea funded by the Ministry of Education under Grant 2011-0014325. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jeffrey A. Fessler.

V.-G. Nguyen was with the Department of Electronic Engineering, Paichai University, Daejeon 302-735, Korea. He is now with the Department of Information Systems, Le Quy Don Technical University, Hanoi 10000, Vietnam (e-mail: giangnv@mta.edu.vn).

S.-J. Lee is with the Department of Electronic Engineering, Paichai University, Daejeon 302-735, Korea (e-mail: sjlee@pcu.ac.kr).

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

Digital Object Identifier 10.1109/TIP.2013.2265881

alleviated [11]. (An extensive overview of the multimodality systems can be found in [12].) In this work we assume that the PET and anatomical images are perfectly aligned and the registration errors between the two images are negligible.

For signal mismatch, it is usually classified into two types. If there is an intensity change in the pixels in the PET image while there is no such change in the (labeling of) anatomical image, it is called missing anatomical information (label). In the opposite case, where there is no intensity change in the pixels in the PET image while there is in the anatomical image, it is called false anatomical information (label) [5].

In order to reduce the signal mismatch error, several different approaches have been proposed. In [4], both pixel values and edges were jointly estimated. The edges in this case were modeled as the binary line processes proposed by Geman and Geman [13]. The line processes signal the presence of an edge between two adjacent pixels. During the reconstruction process, the quadratic penalty encourages smoothness except where edges are detected by the line processes turned on. By using prior anatomical information, the line processes are turned on more easily in those regions of the functional image corresponding to anatomical boundaries. Since this method involves binary variables for the line processes, its overall objective function becomes non-convex, which requires a complex annealing process for optimization. In [5], the anatomical labels derived from an anatomical image were first blurred and then used to indicate the location of edges in the functional PET image via a Bayesian method. While this method can reduce the effects of signal mismatch, it results in high variance in the regions containing false anatomical labels [5]. In [10], an edge-preserving filter derived from the anatomical image was applied to PET image reconstruction via cross entropy minimization. Though this method can preserve the functional edges matched with anatomical edges, it has a limitation in preserving the functional edges when their corresponding anatomical edges are missing. More recently, Cheng-Liao and Qi [8] proposed to use a level set prior which can model not only the smoothness of functional images, but also the similarity between functional boundaries and their corresponding anatomical boundaries. In general, since these existing methods usually involve segmented anatomical regions or boundaries, they have limitations in incorporating complex anatomical structures which are far from piecewise flat.

Recently, efforts have been made to develop methods to incorporate anatomical information into reconstruction without segmenting anatomical images or obtaining anatomical boundary maps. In [11] and [14], joint entropy between the emission distribution and the anatomical image was introduced where the intensity of the anatomical image was directly used. In [7], spatial information was introduced into the joint entropy priors by adding features that capture the underlying anatomical morphology. The features were derived from the scale-space theory, which is a framework for analyzing image structures at different scales. The scale-space features in this case were defined to put a high value on strong boundaries in the anatomical and functional images, and a less value on internal detail and noise.

Another method that directly uses anatomical image was proposed in [15] and was recently re-validated in [11]. In this case, a position dependent neighborhood for the MAP approach was defined by selecting the neighbors that are most similar in the anatomical image. While this method encourages similar emission values in those pixels that are alike in the anatomical image, it is vulnerable in the case of inconsistencies between PET and anatomical images.

In this work, an adaptive weighting method, which can selectively incorporate prior anatomical information without need for explicit boundary or edge information from the anatomical image, is used. The main key to this work is to use the "self-similarity" property of an image, which underlies the fact that every small patch in an image has many similar patches in the same image. It was first exploited in image restoration and resulted in the well-known nonlocal means algorithm [16]. The essence of the nonlocal means algorithm is to perform a weighted averaging process using the natural redundancy of information in an image; the weight of a pixel reflects the similarity between the patch centered at the pixel and the corresponding patch centered at the reference pixel. This principle has been successfully applied to many other applications such as super-resolution imaging [17], [18], compressed sensing recovery [18], and image inpainting [19]. In medical imaging, it has been applied to MRI denoising [20], MRI reconstruction [21], MRI upsampling [22], [23], and CT image reconstruction [24]. Recently, it has also been applied to emission tomography reconstruction [25]–[27].

Our motivation of using the nonlocal method for incorporating anatomical information into PET image reconstruction is that, the nonlocal method can capture a variety of useful image features, such as edges, points and texture, in images [16], [28]. Furthermore, it is less sensitive to noise than other single intensity methods [16]. In this work we show that the nonlocal method can be successfully applied to incorporating anatomical side information into PET reconstruction and has great potential not only to improve the quality of reconstruction but also to resolve the signal mismatch issue better than other single intensity methods.

Our overall reconstruction algorithm is derived from the penalized-likelihood approach in which the penalization is performed by a patch-based nonlocal regularization (NLR) as described in Section II. Section III describes our simulation procedure. In Section IV our experimental results that show the performance of our proposed methods are presented. In Section V our final conclusion on our idea of using the NLR method to incorporate anatomical side information into PET reconstruction is drawn.

II. Methods

A. Nonlocal Regularization for PET Reconstruction Using Prior Anatomical Information

In PET reconstruction, the penalized-likelihood (PL) approach is to seek the underlying source image, whose estimate is denoted as $\mathbf{f}$, by minimizing the following objective function:

$$E(\mathbf{f}) = L(\mathbf{g}|\mathbf{f}) + \gamma\, R(\mathbf{f}), \tag{1}$$

where $L(\mathbf{g}|\mathbf{f})$ is the Poisson negative log likelihood of the projection data $\mathbf{g}$ given $\mathbf{f}$ of the underlying source image, $R(\mathbf{f})$ is the regularization term for $\mathbf{f}$, and γ is a positive parameter. In (1), $L(\mathbf{g}|\mathbf{f}) = \sum_i (\bar{g}_i - g_i \log \bar{g}_i)$, where g_i is the number of photon counts in the i-th detector pair, and $\bar{g}_i = \sum_j \mathcal{H}_{ij} f_j + r_i$ where $\mathcal{H}_{ij} \geq 0$ denotes the element of the forward projection (or the system) matrix, r_i is the mean number of background events such as scattered and random events.

The regularization term is usually defined as

$$R(\mathbf{f}) = \sum_j \sum_{k \in N_j} \omega_{jk} \phi(f_j - f_k), \tag{2}$$

where ω_{jk} is the positive parameter, $\phi(\cdot)$ is the penalty function, and N_j is a local neighborhood of the pixel located at j. In this work only convex penalties are considered.

In the conventional PL approach with local regularization terms, the penalty function is usually specified by a set of pixels that belong to N_j. The weight ω_{jk} is shift invariant and inversely proportional to the distance between the pixel located at j and its neighbor located at k. For the conventional local regularization-based method using anatomical boundary information (referred to as LR-AB), the weight ω_{jk} is set to zero if the intensity difference between the two anatomical pixels located at j and k is greater than a threshold. Otherwise, ω_{jk} is set to one.

In this work, the weight ω_{jk} is defined so that it is proportional to the similarity between the two patches centered at j and k in the PET image. In order to incorporate the prior anatomical information, ω_{jk} is redefined so that it can also reflect the similarity between the two patches centered at j and k in the anatomical image. Our proposed nonlocal regularization-based method using the anatomical prior (NLR-AP) can be represented by

$$\hat{\mathbf{f}} = \arg\max_{\mathbf{f}} \{-L(\mathbf{g}|\mathbf{f}) - \gamma R(\mathbf{f})\}, \tag{3}$$

where the new nonlocal regularization term that includes the anatomical side information is defined as follows:

$$R(\mathbf{f}) = \sum_j \sum_{k \in \Omega_j} \frac{\omega_{jk}}{W_j} \phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right). \tag{4}$$

In (4), $\mathbf{f}(N_j) = \left(f_{j'}, j' \in N_j\right)$ denotes the patch centered at j, and

$$\phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right) = \sum_{p=1}^{P} G_\sigma(p) \phi(f_{j(p)} - f_{k(p)}), \tag{5}$$

where $j(p)$ and $k(p)$ denote the p-th pixels in the patches centered at j and k, respectively, P denotes the number of pixels in a patch, Ω_j is the search window for the pixel located at j, $W_j = \sum_{k \in \Omega_j} \omega_{jk}$, and G_σ is a Gaussian kernel of standard deviation σ. The weight ω_{jk} that reflects self-similarities in both PET and anatomical images is given by

$$\omega_{jk} = \exp\left(-\frac{\phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right)}{h^2}\right) A_{jk}, \tag{6}$$

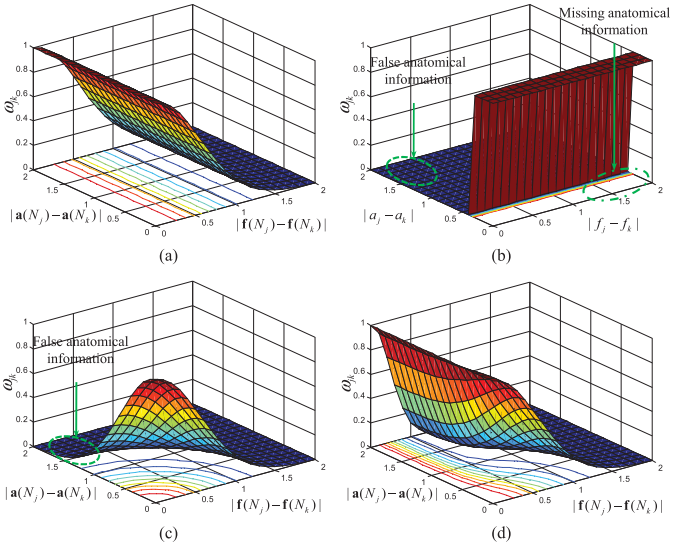

Fig. 1. Sensitivity of ω_{jk} with respect to functional edges and anatomical edges: (a) The standard NLR method that does not use anatomical information; (b) The LR-AB method that erroneously decreases the weight at the location with false anatomical boundaries but increases the weight at the location where anatomical boundaries are missing; (c) The nonlocal regularization method using the independent anatomical prior (NLR-IAP), whose weight factors are specified by Eq. (11), can account for missing anatomical information but has no ability to account for false anatomical information; (d) Our proposed NLR-AP method which adaptively signals the existence of roughness in the anatomical image only when there is similar roughness in the current estimate of the PET image.

where h is a positive parameter and A_{jk} is given by

$$A_{jk} = \exp\left(-\frac{\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)}{h_A^2}\right) + \left(1 - \exp\left(-\frac{\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)}{h_A^2}\right)\right) \exp\left(-\frac{\phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right)}{h_F^2}\right). \tag{7}$$

In (7), $\mathbf{a}$ denotes the anatomical image, the parameters h_A and h_F control the degree of smoothness in the anatomical image and estimated PET image, respectively, both of which affect the current estimate of the PET image.

Note that A_{jk} described in (7) takes into account the degree of roughness in the anatomical image using the term $\exp\left(-\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)/h_A^2\right)$. If the difference between the two anatomical patches $\mathbf{a}(N_j)$ and $\mathbf{a}(N_k)$ is negligible, then $\exp\left(-\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)/h_A^2\right) \approx 1$ and $A_{jk} \approx 1$. In this case, the weight ω_{jk} is determined solely by the patch-based difference in the PET image $\mathbf{f}$ so that the signal mismatch issue due to the missing anatomical edges can be handled. (See Fig. 1(d)). In contrast, if the difference between the two anatomical patches $\mathbf{a}(N_j)$ and $\mathbf{a}(N_k)$ is considerably large, then $A_{jk} \approx \exp\left(-\phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right)/h_F^2\right)$. In this case, A_{jk} takes the role of verifying whether there exists comparably large difference between the patches centered at j and k in the previously estimated PET image $\mathbf{f}$. If so, the weight ω_{jk} is reduced by the multiplication factor A_{jk}, thereby convincing the existence of edges in the PET image. However, if the

difference in the patches centered at j and k in the PET image is negligible, which corresponds to a signal mismatch case due to the false anatomical information, then $A_{jk} \approx 1$ and the weight ω_{jk} is again determined solely by the patch-based difference in the PET image (See Fig. 1(d)). Note that, in general, $\exp\left(-\phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right)/h_F^2\right) \leq A_{jk} \leq 1$. Unlike the conventional regularization function in (2), the NLR function in (4) regularizes two patches whenever the similarity between them, reflected in the weight ω_{jk}, is high.

In our method, as $\mathbf{f}$ is updated iteratively, the space-variant weight $\boldsymbol{\omega}$ is also updated iteratively. This unfavorable update of $\boldsymbol{\omega}$ changes the objective function in (3) at each iteration. However, as shown below, under certain conditions, the objective function can be frozen so that it does not change during iterations.

Inspired by the nonlocal regularization for inverse problems, which includes compressive sensing recovery, super-resolution imaging [18], and exemplar-based image inpainting [19], the weight $\boldsymbol{\omega}$ can be regarded as a set of parameters to be estimated during the reconstruction process. In this work, to estimate $\boldsymbol{\omega}$, we use the maximum entropy approach and consider the following joint optimization:

$$\begin{aligned}(\hat{\mathbf{f}}, \hat{\boldsymbol{\omega}}) &= \arg\max_{\mathbf{f},\boldsymbol{\omega}} \Phi(\mathbf{f}, \boldsymbol{\omega}) \\ &= \arg\max_{\mathbf{f},\boldsymbol{\omega}} \left\{-L(\mathbf{g}|\mathbf{f}) - \gamma\left[R(\mathbf{f}, \boldsymbol{\omega}) - h^2 H(\boldsymbol{\omega})\right]\right\},\end{aligned} \tag{8}$$

where

$$\begin{aligned}R(\mathbf{f}, \boldsymbol{\omega}) = \sum_j \sum_{k\in\Omega_j} &\omega_{jk} \\ &\times\left[\phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right) + \tau\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)\right],\end{aligned} \tag{9}$$

$H(\boldsymbol{\omega})$ is the summation of local entropy associated with the weight $\boldsymbol{\omega}$

$$H(\boldsymbol{\omega}) = -\sum\nolimits_j \sum\nolimits_{k\in\Omega_j} \omega_{jk} \log(\boldsymbol{\omega}_{jk}), \tag{10}$$

and the weight $\boldsymbol{\omega}$ satisfies the constraint $\sum_k \omega_{jk} = 1\ \forall j$. In this case, h is the positive parameter which weights the summation of local entropy of the weight $\boldsymbol{\omega}$ relative to the regularization term $R(\mathbf{f}, \boldsymbol{\omega})$ (h can also be considered as a filter parameter; the larger its value, the smoother the reconstructed image), τ is the positive parameter that balances the contribution of the patch differences between the PET image and the anatomical image to the nonlocal regularizer. The summation of local entropy $H(\boldsymbol{\omega})$ in (8) is to constrain the weight $\boldsymbol{\omega}$ to have large entropy [18], [19].

Although the function $\Phi(\mathbf{f}, \boldsymbol{\omega})$ is not jointly concave in $\mathbf{f}$ and $\boldsymbol{\omega}$, it is concave with respect to $\mathbf{f}$ and also with respect to $\boldsymbol{\omega}$. A maximizer $(\hat{\mathbf{f}}, \hat{\boldsymbol{\omega}})$ to (8) can be found by using alternating optimization; an iterative procedure is performed for maximizing $\Phi(\mathbf{f}, \boldsymbol{\omega})$ jointly over both $\mathbf{f}$ and $\boldsymbol{\omega}$ by alternating restricted maximization over each variable. It can be shown that, maximizing $\Phi(\mathbf{f}, \boldsymbol{\omega})$, while fixing $\mathbf{f}$, results in the following update for ω_{jk}:

$$\begin{aligned}\omega_{jk} = \frac{1}{W_j} &\exp\left(-\frac{\phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right)}{h^2}\right) \\ &\times \exp\left(-\tau\frac{\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)}{h^2}\right),\end{aligned} \tag{11}$$

where

$$\begin{aligned}W_j = \sum\nolimits_k &\exp\left(-\frac{\phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right)}{h^2}\right) \\ &\times \exp\left(-\tau\frac{\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)}{h^2}\right)\end{aligned}$$

is the normalizing constant. When $\boldsymbol{\omega}$ is fixed, (8) becomes the standard nonlocal regularization method. Therefore, the joint optimization in (8) is a special case of our proposed NLR-AP method. The special case occurs when the anatomical information is independently used without referring to the estimated PET image $\mathbf{f}$. In this case, the false anatomical information can affect the reconstruction (See Fig. 1(c)). Mathematically, the equivalence is achieved when (7) is reduced to $A_{jk} = \exp\left(-\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)/h_A^2\right)$ with $h_A = h/\sqrt{\tau}$. This NLR method using the independent anatomical prior is referred to as NLR-IAP.

In summary, if $A_{jk} = \exp\left(-\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)/h_A^2\right)$ and $h_A = h/\sqrt{\tau}$, the NLR-AP method becomes the NLR-IAP method, which is a joint minimization problem of the fixed objective function given in (8). Otherwise, the objective function for the NLR-AP method is not fixed.

B. Other Related Methods

Among the recent works using nonlocal regularization for image restoration and image reconstruction ([17], [20], [21], [23], [24], [27], [29]), [23] is closely related to our work in that a nonlocal approach was applied to incorporating the side information for image super-resolution in MRI.

In [23], the problem of image super-resolution from a single low-resolution image (also known as image up-scaling) was considered. In particular, it aims to generate a high-resolution brain image from a single low-resolution T2-weighted image with the aid of a high-resolution T1-weighted image obtained from another modality. In this case, the T1-weighted image was regarded as a high-resolution anatomical image. The authors proposed to solve the problem by using a regularization method where the regularizer is given by

$$R(\mathbf{f}, \mathbf{a}) = \sum_j \left\| f_j - d_{NLM}(f_j, \mathbf{a}) \right\|^2, \tag{12}$$

where $d_{NLM}(f_j, \mathbf{a})$ is the denoised version of f_j with the support of anatomical image $\mathbf{a}$, and is given by

$$d_{NLM}(f_j, \mathbf{a}) = \alpha_j \sum_{k\in\Omega_j} \omega_{jk}(\mathbf{a}) f_k + (1-\alpha_j) \sum_{k\in\Omega_j} \omega_{jk}(\mathbf{f}) f_k. \tag{13}$$

In (13) $\omega(\mathbf{a})$ and $\omega(\mathbf{f})$ are the weighted graphs estimated on the anatomical image $\mathbf{a}$ and reconstructed image $\mathbf{f}$, respectively,

which are defined by

$$\omega_{jk}(\mathbf{a}) = \exp\left(-\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)/h_A^2\right) \Big/ \sum_{k\in\Omega_j} \exp\left(-\phi\left(\mathbf{a}(N_j) - \mathbf{a}(N_k)\right)/h_A^2\right), \quad (14)$$

$$\omega_{jk}(\mathbf{f}) = \exp\left(-\phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right)/h^2\right) \Big/ \sum_{k\in\Omega_j} \exp\left(-\phi\left(\mathbf{f}(N_j) - \mathbf{f}(N_k)\right)/h^2\right), \quad (15)$$

and α_j is the local correlation between $\omega(\mathbf{a})$ and $\omega(\mathbf{f})$ at the pixel located at j. Note that the weights in the first term on the right side of (13) are computed from the anatomical image and those in the second term are computed from the current estimate of the reconstructed image. The anatomical image is used in the denoising process in (13), which eventually affects the reconstruction process via (12) only when the two weighted graphs centered at the j-th pixel are correlated.

Our proposed method is different from this method in that, while this method first performs a nonlocal denoising process of the anatomical image and regularizes by minimizing the difference between the current estimate and the denoised version of the estimate, our method directly performs nonlocal regularization by minimizing the quadratic penalty with the weights computed from both the anatomical image and the PET image.

C. Optimization Method

Since the solution for maximizing the NLR-AP objective function described in (3) cannot be found in a closed form, an iterative method needs to be used. A variety of fast and globally convergent optimization algorithms for convex objectives have been proposed [30]–[32]. To achieve fast convergence, most of the algorithms use block-iterative methods based on the ordered-subsets (OS) principle [33]. In this work, to efficiently find $\hat{\mathbf{f}}$ in (3), we use the complete-data ordered subsets expectation maximization (COSEM) algorithm [30], which is free from a seriously inconvenient user-specific relaxation schedule required in conventional relaxed OS methods.

In the COSEM algorithm, a 'complete data objective' is jointly minimized with respect to both the complete data $\boldsymbol{C}$ and the underlying PET image $\mathbf{f}$. The element C_{ij} of the complete data $\boldsymbol{C}$, which is real and positive, obeys the constraint $\sum_j C_{ij} = g_i$. In the standard COSEM algorithm, the update equations are obtained by directly differentiating the complete data objective with respect to C_{ij} and f_j, and setting the result to zero while satisfying constraints on $\boldsymbol{C}$. For the objective function given in (3), the update equation for C_{ij} remains the same as the conventional COSEM algorithm. However, the update equation for f_j cannot be derived by directly differentiating the complete data objective with respect to C_{ij} and f_j since the objective function involves the regularization term that introduces coupling between the patches, which results in complex interaction between source pixels. To decouple the regularization term, the separable surrogate approach can be used as done in [34] and [35]. With n as the index for the outer iteration and l as the subset iteration for the COSEM-PL algorithm, the surrogate $R_s(\mathbf{f}; \mathbf{f}^{(n,l-1)})$ for $R(\mathbf{f})$ in (4) is given by

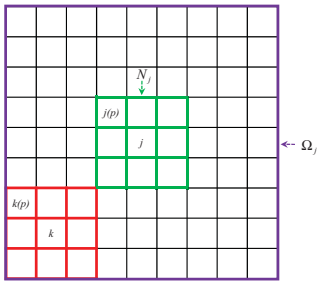

Fig. 2. A neighborhood system in the nonlocal regularization method.

$$R_s(\mathbf{f}; \mathbf{f}^{(n,l-1)}) = \sum_j \sum_{k\in\Omega_j} \frac{\omega_{jk}}{2W_j} \left[\sum_{p=1}^{P} G_\sigma(p) \left(\phi\left(2f_{j(p)} - f_{j(p)}^{(n,l-1)} - f_{k(p)}^{(n,l-1)}\right) + \phi\left(2f_{k(p)} - f_{j(p)}^{(n,l-1)} - f_{k(p)}^{(n,l-1)}\right)\right)\right], \quad (16)$$

where the convexity of the penalty function $\phi(\cdot)$ has been used to obtain the surrogate [34], [35] $R_s(\mathbf{f}; \mathbf{f}^{(n,l-1)}) \geq R(f)$.

The surrogate regularization term $R_s(\mathbf{f}; \mathbf{f}^{(n,l-1)})$ is now separable with respect to f_j so that,

$$R_s(f_j; \mathbf{f}^{(n,l-1)}) = \sum_{j(p)\in N_j} G_\sigma(p) \sum_{k(p)\in\Omega_{j(p)}} \frac{\omega_{j(p)k(p)}}{W_{j(p)}} \phi\left(2f_j - f_j^{(n,l-1)} - f_k^{(n,l-1)}\right) \quad (17)$$

where $k = k(p) + j - j(p)$ (See Fig. 2).

Therefore, we have the following surrogate objective function which is separable with respect to f_j

$$E_s(f_j; \mathbf{f}^{(n,l-1)}, \mathbf{C}^{(n,l)}) = -\sum_i C_{ij}^{(n,l)} \log f_j + \sum_i \mathcal{H}_{ij} f_j + \gamma R_s(f_j; \mathbf{f}^{(n,l-1)}) \quad (18)$$

When the penalty function $\phi(\cdot)$ has the quadratic form $\phi(\xi) = \xi^2$, the minimizer of (18) can be directly found by differentiating $E_s(f_j; \mathbf{f}^{(n,l-1)}, \boldsymbol{C}^{(n,l)})$ with respect to f_j and set the result to zero. The solution in this case is positive, unique and of closed-form. Note that, if the penalty function is convex nonquadratic, the closed-form update for $f_j^{(n,l)}$ will be replaced by 1-D optimization as done in [36]. Before describing the COSEM-PL algorithm in detail we first define the following notations as done in [37]: $B_j^{(n,l)} \stackrel{\text{def}}{=} \sum_i C_{ij}^{(n,l)}, \forall j$, $A_j^{(n,l)} \stackrel{\text{def}}{=} \sum_{i\in S_l} C_{ij}^{(n,l)}, l = 1, ..., L \ \forall j$, and $D_j \stackrel{\text{def}}{=} \sum_i \mathcal{H}_{ij}, \forall j$ where L is the number of subsets. The COSEM-PL algorithm first subdivides the projection data into an ordered sequence of L disjoint subsets $S_l : l = 1, ..., L$. An iteration of the COSEM-PL algorithm is defined as a single pass through all of L subsets. Given $\mathbf{f}^{(0)}$ as an initial image, the outline of the COSEM-PL algorithm is described in Table I.

In summary, to find a solution for our NLR method, the following procedure is used:

- Initialize $\mathbf{f}^{(0)}$
- **For** $m = 1, ...,$

TABLE I
OUTLINE FOR THE COSEM-PL ALGORITHM

Initialize $\bar{g}_i = \sum_j \mathcal{H}_{ij} f_j^{(0)}, \forall i$,

$A_j^{(0,l)} = f_j^{(0)} \sum_{i \in S_l} \frac{g_i}{\bar{g}_i} \mathcal{H}_{ij}, \forall j, \ l = 1, ..., L \quad B_j^{(0,L)} = f_j^{(0)} \sum_i \frac{g_i}{\bar{g}_i} \mathcal{H}_{ij}, \forall j$

for each iteration $n = 1, ...$

$\mathbf{f}^{(n,0)} = \mathbf{f}^{(n-1)}$,

$B_j^{(n,0)} = B_j^{(n-1,L)}, \forall j$

for each subset $l = 1, ..., L$

$\bar{g}_i = \sum_j \mathcal{H}_{ij} f_j^{(n,l-1)}$ for $\forall i \in S_l$,

$A_j^{(n,l)} = f_j^{(n,l-1)} \sum_{i \in S_l} \frac{g_i}{\bar{g}_i} \mathcal{H}_{ij}$, for $\forall j$

$B_j^{(n,l)} = A_j^{(n,l)} - A_j^{(n-1,l)} + B_j^{(n,l-1)}$, for $\forall j$

$f_j^{(n,l)} = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$, for $\forall j$ where

$a = 8\gamma \sum_{j(p) \in N_j} G_\sigma(p) \sum_{k(p) \in \Omega_{j(p)}} \frac{\omega_{j(p)k(p)}}{W_{j(p)}}$,

$b = D_j - 4\gamma \sum_{j(p) \in N_j} G_\sigma(p) \sum_{k(p) \in \Omega_{j(p)}} \frac{\omega_{j(p)k(p)}}{W_{j(p)}} (f_j^{(n,l-1)} + f_k^{(n,l-1)})$,

$c = -B_j^{(n,l)}$

end

$\mathbf{f}^{(n)} = \mathbf{f}^{(n,L)}$,

end

- Set $\mathbf{f} = \mathbf{f}^{(m-1)}$,
- Update the weights $\omega_{jk} \ \forall j, k$ using (6),
- Update image $\mathbf{f}^{(m)}$ using the COSEM-PL algorithm described in Table I.

In our application, $\mathbf{f}$ and ω are updated at each iteration of the COSEM-PL algorithm.

III. SIMULATIONS

In PET reconstruction, due to limited width of detector bins and the low photon counts, images are usually reconstructed at a relatively low-resolution. In reconstructing PET images using prior anatomical information, since the anatomical images obtained from MR or CT scans have relatively higher resolution than the PET images, they are usually down-sampled to match the resolution of the PET image [5]. Here we note that, instead of simply down-sampling the high-resolution anatomical images, we can also extend our simulations to increase the pixel resolution of PET images up to the level of the high-resolution anatomical image by modeling the underlying PET image on a finer-grid (with the smaller pixel size) and performing our nonlocal regularization method using the side information obtained from the high-resolution anatomical image. In this section, the results from the extended simulation are presented along with the results from the simulation for the original low-resolution PET reconstruction.

To evaluate the performance of our proposed method, we performed 2-D simulations using the 128 × 128 software phantoms A and B shown in Figs. 3(a) and 4(a), respectively.

TABLE II
VALUES OF THE SMOOTHING PARAMETER γ USED IN SIMULATIONS

Phantom (Resolution)	LR-QD	LR-AB	NLR	NLR -AP	NLR -IAP
A(L)	0.0024	0.0024	0.0012	0.0012	0.0012
A(H)	0.04	0.04	0.005	0.005	0.005
B(L)	0.00055	0.00055	0.0002	0.0002	0.0002
B(H)	0.0028	0.0028	0.001	0.001	0.001

Phantom A consists of piecewise flat regions with activities of 4:1:0. Phantom B was derived from a digitized rhesus monkey autoradiograph. The intent of using such an autoradiograph for phantom B is that it contains more realistic edge structure than a simple piecewise constant phantom, such as phantom A. The phantoms in Fig. 3(a) and Fig. 4(a) were derived from their original 256 × 256 phantoms, Fig. 3(b) and Fig. 4(b), respectively, by summing up four adjacent pixels in each high-resolution phantom to generate a corresponding pixel in the associated low-resolution phantom. The low- and high-resolution software anatomical phantoms are shown in Figs. 3(c) and 3(d), respectively, for phantom A, and Figs. 4(c) and 4(d), respectively, for phantom B. (The low-resolution anatomical phantoms were obtained by down sampling the high-resolution phantoms by using the 8-nearest neighbor interpolation method.) The intensity ratio of anatomical image for phantom A was 4:2:1:0. The phantoms were designed to contain unmatched region boundaries between the functional PET image and the anatomical image. For the set of phantom A, the region denoted R1 in the PET image has no corresponding region in the anatomical image, the boundary of the region denoted R2 perfectly matches with its corresponding anatomical boundary, and the boundary of the region denoted R3 partially matches with its corresponding anatomical boundary. For the set of phantom B, two tumors with high activity were inserted into the PET image. Both tumors have no corresponding regions in the anatomical image. However, the region denoted R2 partially matches with the anatomical boundary.

To generate projection data with independent Poisson noise, the low-resolution phantoms were used. The total projection counts in this case were approximately 400,000 for phantom A and 1,000,000 for phantom B. The width of the detector bins was set to the length equal to the pixel size in the conventional (referred to as low-resolution) reconstruction, which was twice as long as the pixel size in the high-resolution reconstruction.

In this work we tested the following two PL reconstruction methods: (i) our proposed nonlocal regularization-based method using anatomical prior (NLR-AP), (ii) the conventional local regularization-based method using anatomical boundary information (LR-AB) [5]. In both methods, the penalty function was set to $\phi(\xi) = \xi^2$. Therefore, if the anatomical information is ignored, the NLR-AP method reduces to the PL method using simple nonlocal regularization (referred to as NLR) and the LR-AB method reduces to the conventional PL method with a quadratic prior (referred to as LR-QD). We also tested the NLR-IAP method that incorporates the anatomical

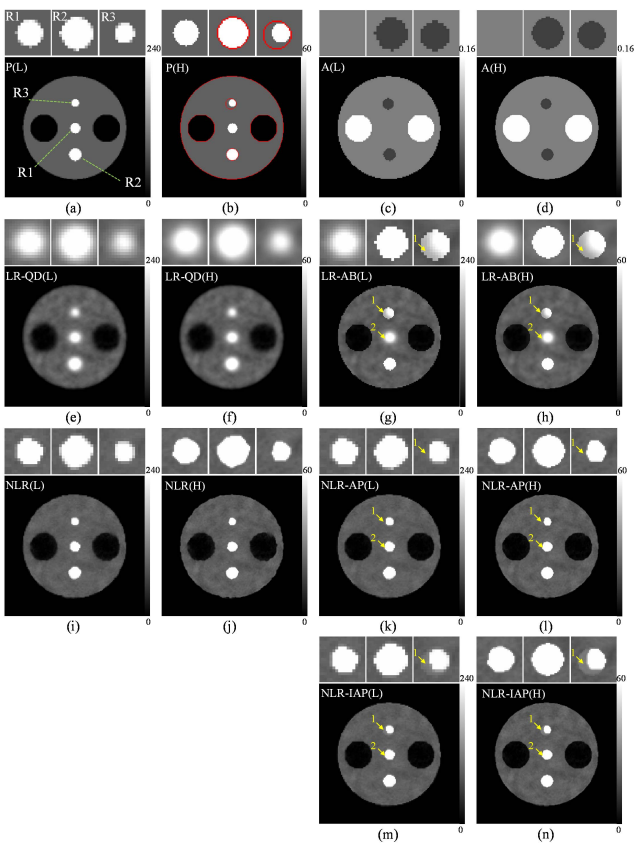

Fig. 3. Reconstructions for phantom A: (a) Low-resolution phantom (used to generate simulated projection data for all reconstructions); (b) High-resolution phantom with anatomical edges superposed; (c) Low-resolution anatomical image; (d) High-resolution anatomical image; (e)(f) LR-QD; (g)(h) LR-AB; (i)(j) NLR; (k)(l) NLR-AP; (m)(n) NLR-IAP. P(L) and P(H) denote low- and high-resolution phantoms, respectively, and A(L) and A(H) denote low- and high-resolution anatomical images, respectively. Arrows designate mismatched regions; #1 for false anatomical information and #2 for missing anatomical information. The three small images on top of each figure show zoomed-in regions.

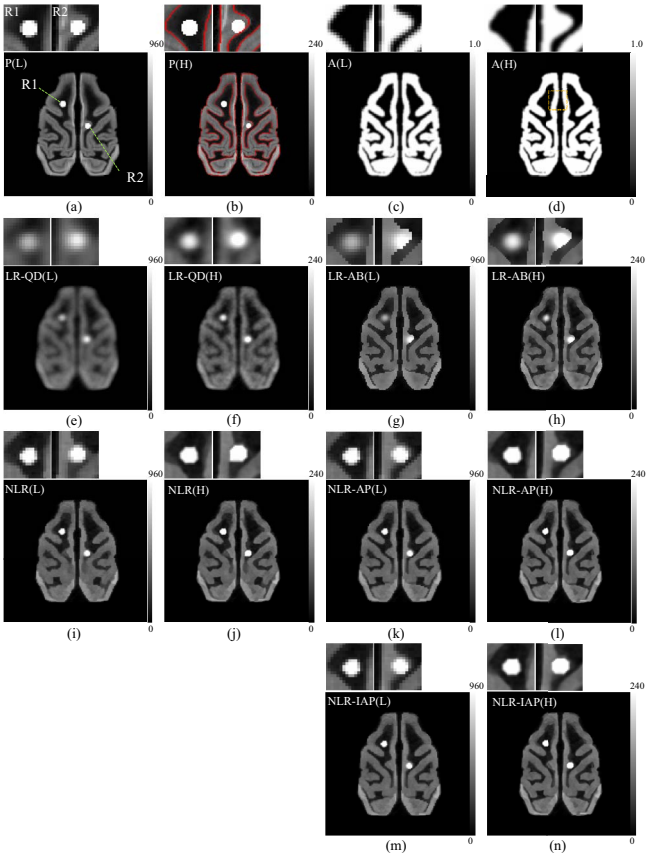

Fig. 4. Reconstructions for phantom B: (a) Low-resolution phantom (used to generate simulated projection data for all reconstructions); (b) High-resolution phantom with anatomical edges superposed; (c) Low-resolution anatomical image; (d) High-resolution anatomical image; (e)(f) LR-QD; (g)(h) LR-AB; (i)(j) NLR; (k)(l) NLR-AP; (m)(n) NLR-IAP. P(L) and P(H) denote low- and high-resolution phantoms, respectively, and A(L) and A(H) denote low- and high-resolution anatomical images, respectively. The two small images on top of each figure show zoomed-in regions.

prior independently without referring to the PET image. The number of iterations was set to 100 for all methods. In implementing the above algorithms, the system matrix $\mathcal{H}$ could be modeled to include physical factors in PET imaging, such as the positron range blurring, geometrical sensitivity, attenuation, detector blurring, and detector sensitivity [2]. However, to avoid any side effects, which can be caused by the physical factors, in investigating the efficacy of our NLR-AP method, we ignored the physical factors by using a simple geometric projection matrix in modeling the system matrix. To accurately account for finite-width detectors in modeling $\mathcal{H}$, however, we employed the strip area-based method [38], which calculates the area of intersection between a pixel and the ray strip formed by detector pair. (For three-dimensional reconstruction, the distance-driven method [39] or the separable footprint method [40] can be used to model the system matrix.) In the strip area-based method, an element of the system matrix, $\mathcal{H}_{ij}$, is modeled as the weight calculated by the intersecting area s_{ij} between the pixel located at j and the strip formed by the i-th finite-width detector pair. For PET, each element of the system matrix can be normalized so that [10] $\mathcal{H}_{ij} = s_{ij} / \sum_i s_{ij}$.

TABLE III
WINDOW SIZES AND VALUES OF FILTER PARAMETERS USED IN SIMULATIONS

Phantom (Resolution)	$\lvert\Omega_j\rvert$	$\lvert N_j\rvert$	h	h_F
A(L)	7×7	3×3	48	18
A(H)	11×11	3×3	12	4.5
B(L)	7×7	3×3	84	36
B(H)	11×11	5×5	25	10

The anatomical label maps for the LR-AB method were generated by manually segmenting the anatomical phantom images. The initial estimate for $\hat{\mathbf{f}}$ in all reconstruction methods was initialized to the image reconstructed by 15 iterations of the maximum likelihood expectation maximization algorithm [41]. For the smoothing parameter γ, we empirically chose different values for different methods so that the resulting reconstructions had approximately the same background noise level. The resulting values of γ are summarized in Table II.

The search window Ω_j and the neighborhood window N_j (or the similarity window in the context of a nonlocal means

approach) were set to the values described in Table III. The parameter h_A was set to [20]

$$h_A = \sqrt{2P\beta\hat{\sigma}_A^2}. \tag{19}$$

In (19), $\hat{\sigma}_A$ is the estimated standard deviation of noise in the anatomical image. For Gaussian noise, $\hat{\sigma}_A$ is given by $\hat{\sigma}_A = \sqrt{\sum_j \tilde{\varepsilon}_j^2 / N_A}$, where N_A is the number of pixels in the anatomical image, and $\tilde{\varepsilon}_j$ is the pseudo-residual [42] given by $\tilde{\varepsilon}_j = \sqrt{(4/5)}\left((1/4)\sum_{k\in M_j} a_k - a_j\right)$ for 2-D images, where M_j is the nearest 4-neighbors of the pixel located at j. For data-dependent noise, $\hat{\sigma}_A(j)$ is dependent on the intensity of the j-th pixel in the anatomical image and is estimated directly from the anatomical image using the estimation method proposed in [43]. The parameter β is a constant determined by the noise model in the anatomical image. For Gaussian noise, β is set to 1. In fact (19) has proven useful for Rician noise in MRI [20], [23].

The parameters h and h_F are also related to the variance of noise in the PET image. However, since the PET image contains signal-dependent noise [44] which is different from additive Gaussian noise, the method of pseudo-residuals cannot be directly used to estimate variance of the noise in the PET image. In this work, while the value of h was chosen so that the NLR method results in the same background noise level as the LR-QD and LR-AB methods, the value of h_F was chosen so that the NLR-AP method results in the same background noise level as the NLR method. The values of h and h_F used in the simulations are listed in Table III. Note that, since the last term involving h_F in (7) is used to confirm the large difference between the two patches in the PET image in conjunction with that in the anatomical image, h_F is usually set to a value less than h as shown in Table III.

To evaluate the quantitative performance of reconstructed images, we computed point-wise bias and standard deviation (STD) images. A bias image b_j is defined as

$$b_j = (1/N)\sum_{n=1}^{N}(\hat{f}_j^n - f_j), \tag{20}$$

where $\hat{f}_j^n$ is the pixel located at j in the n-th reconstruction and the summation is over $N = 100$ independent noise trials. (To display the bipolar bias image, an intermediate gray scale of 128 out of 256 levels was used as zero bias.) A standard deviation image, s_j, is defined as

$$s_j = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(\hat{f}_j^n - \bar{f}_j)^2}, \tag{21}$$

where $\bar{f}_j$ is the mean of $\hat{f}_j$ over the noise trials defined as $\bar{f}_j = (1/N)\sum_{n=1}^{N}\hat{f}_j^n$. The smoothing parameter γ in this case was chosen so that different methods resulted in approximately the same level of STD in the background region.

IV. Results

Figs. 3 and 4 show anecdotal reconstructions for phantoms A and B, respectively. The regions designated by the arrows in Figs. 3(g)(h)(k)(l)(m)(n) are mismatched regions between the PET image and the anatomical image; the arrow #1 is for the false anatomical information and the arrow #2 is for the missing anatomical information. Several qualitative observations may be noted here. Clearly, the PL method performs better when the anatomical information is incorporated into the reconstruction process. For the LR-AB method, since it uses the boundary information derived from the anatomical image, unless any false information is involved, it performs well in both low-resolution and high-resolution reconstructions as shown in Figs. 3(g) and (h). However, it leads to a significant error when the boundaries in the anatomical image and those in the PET image are not matched as shown in Figs. 3(g) and 3(h) (see the region designated by arrow #1) On the other hand, our proposed NLR-AP method, which does not require additional boundary information, performs better than the LR-AB method when the image resolution is increased. Comparisons of Figs. 3(g) with 3(k) and Figs. 3(h) with 3(l) show the better performance of the NLR-AP method than that of the LR-AB method when the resolution is increased. Comparisons of Figs. 3(k) with (m) and Figs. 3(l) with (n) show that, unlike the NLR-AP method, the NLR-IAP method results in errors in the false anatomical region (designated by the arrow #1). This is due to the fact that, while our NLR-AP method confirms the existence of edges whenever there exist edges in the anatomical image by referring to the PET image, the NLR-IAP method independently incorporates the anatomical information without referring to the PET image. Therefore, if there exist false anatomical regions, the NLR-IAP reconstruction produces false edges in the corresponding regions. Comparisons of Figs. 3(l) with 3(k) and 3(l) with 3(h) show that the high-resolution reconstruction using our proposed method improves the reconstruction accuracy. (See the region designated by the arrow #1.)

TABLE IV
Mean Absolute Percentage Errors (%) Calculated from 100 Noise Trials Using Phantom A

Method\ Resolution	Low: Full Image (ROIs R1/R2/R3)	High: Full Image (ROIs R1/R2/R3)
LR-QD	24.04 (30.21/25.05/37.68)	23.85 (29.34/24.44/36.91)
LR-AB	18.92 (30.34/12.48/29.51)	17.14 (29.42/6.21/29.67)
NLR	12.95 (14.63/13.85/19.26)	12.51 (13.68/11.94/16.66)
NLR-AP	11.43 (14.70/9.53/14.70)	9.73 (13.71/5.07/14.04)
NLR-IAP	10.77 (14.71/8.85/14.84)	9.13 (13.72/3.76/14.53)

Table IV summarizes the mean absolute percentage errors (MAPEs) of the reconstructions and the regions of interests (ROIs), R1, R2, and R3 calculated from 100 noise trials. The sizes of ROIs were 72, 107, and 42 pixels for R1, R2, and R3, respectively. To calculate MAPEs of the high-resolution reconstructions under the same condition as the low-resolution reconstructions, the high-resolution reconstructions were first down-sampled to the low-resolution images and their MAPEs were calculated with respect to the low-resolution phantom A. As shown in Table IV, regardless of the level of

the resolution, the NLR-based methods clearly outperform the conventional LR-based methods. For the comparison of NLR-IAP with NLR-AP, the overall MAPEs of the NLR-IAP reconstructions appear smaller than those of the NLR-AP reconstructions. This is presumably due to the fact that, as the NLR-IAP method is strongly affected by anatomical boundaries, it usually performs well for piecewise flat reconstructions when the signal mismatch problem is not severe. This is also true in R2 for a perfectly matched anatomical region. On the other hand, the MAPEs of R1, which contains no anatomical edges, are about the same for the two methods. Finally, the NLR-AP method, as expected, outperforms the NLR-IAP method in R3 which contains false anatomical information. Note that the MAPEs for the high-resolution reconstructions are smaller than those for the low-resolution reconstructions. The MAPEs measured from the ROIs also support this superior performance of the high-resolution reconstructions.

For the phantom B whose anatomical boundaries are not as sharp as those of the phantom A, the incorporation of the anatomical information improves accuracy in both LR-AB and NLR-AP reconstructions (see Figs. 4(g)(h) and (k)(l).) While the visual improvements from LR-AB to NLR-AP are not stunning, close inspection reveals that the NLR-AP reconstruction, indeed captures subtle aspects of the autoradiograph phantom B that are missed by the LR-AB method; while the LR-AB method attempts to flatten the 'ramp-like' gradually changing regions (see Figs. 4(g) and 4(h)), the NLR-AP method avoids the over-smoothing behavior of the LR-AB method as seen in Fig. 4(l). For this phantom, due to the absence of false anatomical boundaries, the results from NLR-AP and NLR-IAP are almost identical. The MAPEs of the reconstructions and the two ROIs shown in Fig. 4(a) are summarized in Table V, which supports the results from our visual inspection. The sizes of ROIs were 31 and 22 pixels for R1 and R2, respectively.

We also calculated point-wise bias and STD images. To compare the high-resolution reconstructions in the same condition as the low-resolution reconstructions, the reconstructed high-resolution images were down-sampled to the low-resolution images. Note again that the smoothing parameter γ in this case was chosen so that different methods resulted in approximately the same STD level in the background region. Bias images are bipolar, with a value of zero displayed as an intermediate gray, and with darker/lighter regions corresponding to negative/positive bias. In STD images, intensity represents the positive STD value.

Fig. 5 shows our bias and STD images for phantom A. The results indicate that the incorporation of anatomical information immediately reduces the bias error in both the LR-AB and the NLR-AP methods. Comparisons of Fig. 5(d) with Fig. 5(h) and Fig. 5(c) with Fig. 5(g) reveal that our proposed NLR-AP method outperforms the LR-AB method by significantly reducing the bias error. The areas designated by arrows in Figs. 5(c)(d) and 5(g)(h) show that the NLR-AP method outperforms the LR-AB method in bias where anatomical regions are missing (arrow #2) or false (arrow #1). Comparisons of Figs. 5(g) with 5(h) show that increasing the pixel resolution in NLR-AP reconstruction and then down-sampling to conventional resolution reduces bias. Finally, comparisons of Figs. 5(g) with 5(i) and 5(h) with 5(j) show that the NLR-IAP method has all the advantages of the proposed NLR-AP method except that it has no ability to resolve the false anatomical information. (See the arrow #1's in Figs. 5(i) and (j) where the false anatomical boundary resulted in high bias error.)

TABLE V
MEAN ABSOLUTE PERCENTAGE ERRORS (%) CALCULATED FROM 100 NOISE TRIALS USING PHANTOM B

Method\ Resolution	Low: Full Image (ROIs R1/R2)	High: Full Image (ROIs R1/R2)
LR-QD	29.48 (40.37/42.99)	23.24 (28.93/30.87)
LR-AB	22.72 (40.71/35.52)	19.02 (28.98/26.91)
NLR	18.24 (17.52/16.85)	17.13 (15.92/14.97)
NLR-AP	14.95 (16.59/17.19)	14.87 (15.86/14.86)
NLR-IAP	14.54 (16.55/17.47)	14.75 (15.86/14.86)

Fig. 6 shows our bias and STD images for phantom B. The results once again indicate that the incorporation of anatomical information improves reconstruction accuracy in bias at the same level of STD. For this realistic phantom, the errors in LR-AB reconstruction are caused mainly by the inconsistency between the segmented anatomical image and the anatomical phantom image; the anatomical phantom image contains more complex edge structure than the segmented anatomical image. The NLR-AP reconstructions (Figs. 6(g) and 6(h)) reveal better accuracy than the LR-AB reconstructions (Figs. 6(c) and 6(d)) in bias at the same level of STD. Visual comparisons of Figs. 6(g) with 6(i) and 6(h) with 6(j) show that, since the false anatomical information does not exist in this particular phantom, the NLR-IAP method performs slightly better than the NLR-AP method in most areas. Note that, in point-wise STD, the results from both LR-QD and LR-AB (Figs. 6(k)(l) and 6(m)(n), respectively) have less fluctuations than those from NLR and NLR-AP (Figs. 6(o)(p) and 6(q)(r), respectively). However, it can be easily noted that the high STD values in the NLR and NLR-AP/IAP reconstructions appear only around the high-contrast edges.

Finally, to evaluate the quantitative performance of our proposed method focused in the tumor regions in both phantoms A and B, we performed additional ROI studies using the contrast recovery coefficient (CRC) [45]. Figs. 7(a) and 8(a) show the preselected ROIs for the phantoms A and B, respectively, which are identical to those shown in Figs. 3(a) and 4(a), respectively. In both Figs. 7(a) and 8(a), the circles denoted B indicate the background regions.

The contrast recovery coefficient (CRC) in each tumor region is defined as

$$CRC^n_{\mathrm{ROI}} = CR^n_{\mathrm{ROI}} / CR^0_{\mathrm{ROI}}, \tag{22}$$

where $CR^n_{\mathrm{ROI}} = \left| Z^n_{\mathrm{ROI}} - Z^n_{\mathrm{B}} \right| / Z^n_{\mathrm{B}}$ with $Z^n_{\mathrm{ROI}} =$

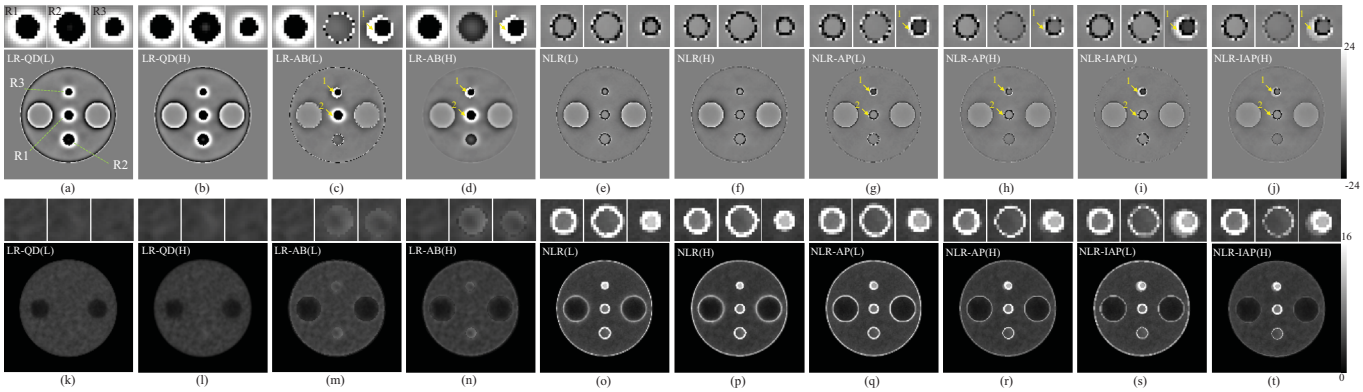

Fig. 5. Point-wise bias/STD images for phantom A: (a)(b) bias for LR-QD; (c)(d) bias for LR-AB; (e)(f) bias for NLR; (g)(h) bias for NLR-AP; (i)(j) bias for NLR-IAP; (k)(l) STD for LR-QD; (m)(n) STD for LR-AB; (o)(p) STD for NLR; (q)(r) STD for NLR-AP; (s)(t) STD for NLR-IAP. (L) and (H) denote low- and high-resolution, respectively. The three small images on top of each figure show zoomed-in regions.

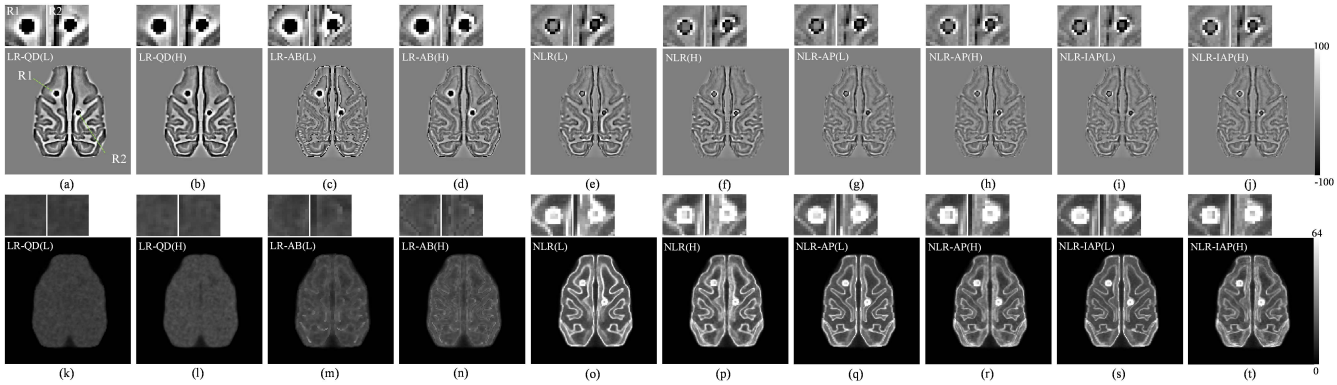

Fig. 6. Point-wise bias/STD images for phantom B: (a)(b) bias for LR-QD; (c)(d) bias for LR-AB; (e)(f) bias for NLR; (g)(h) bias for NLR-AP; (i)(j) bias for NLR-IAP; (k)(l) STD for LR-QD; (m)(n) STD for LR-AB; (o)(p) STD for NLR; (q)(r) STD for NLR-AP; (s)(t) STD for NLR-IAP. (L) and (H) denote low- and high-resolution, respectively. The two small images on top of each figure show zoomed-in regions.

$(1/m)\sum_{j\in \mathrm{ROI}} \hat{f}_j^n$ denoting the mean activity in each region at the n-th noise trial, m the number of pixels in the ROI, Z_{B}^n the mean activity in the background region, and CR_{ROI}^0 the true contrast in the phantom. The ensemble mean of CRC is defined as

$$\overline{CRC}_{\mathrm{ROI}} \stackrel{\mathrm{def}}{=} (1/N)\sum_{n=1}^{N} CRC_{\mathrm{ROI}}^n, \tag{23}$$

and the regional standard deviation is defined as

$$STD_{\mathrm{ROI}}^{CRC} = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(CRC_{\mathrm{ROI}}^n - \overline{CRC}_{\mathrm{ROI}})^2}, \tag{24}$$

where N is the total number of noise trials.

Fig. 7 shows regional CRC-STD curves for phantom A in which each point represents the mean and STD of the CRC evaluated from $N=100$ independent noise trials. To observe the effect of smoothing over a range of γ, the nine different γ values were used; γ was set to 0, 0.0000375, 0.000075, 0.00015, 0.0003, 0.0006, 0.0012, 0.0024, 0.0048 for the original low-resolution reconstructions and 0, 0.0003125, 0.000625, 0.00125, 0.0025, 0.005, 0.01, 0.02, 0.04 for the high-resolution reconstructions. Therefore, each curve contains nine CRC-STD points for each reconstruction algorithm. For each curve, the starting point with the highest STD corresponds to $\gamma = 0$. The reconstructions with the same resolution share the same starting point.

According to the results from the ROI studies, unless the degree of signal mismatch was severe, the methods that used the anatomical information always yielded better results in CRC than the methods without using the anatomical information. In particular, for the cases of signal mismatch in R1 and R3, our proposed NLR-AP/IAP methods provided better results with higher CRC than the LR-AB method over a wide range of γ. For R2, which was the case of the region with the matched boundary, both the NLR-AP and NLR-IAP methods revealed the higher CRC and smaller STD than the LR-AB method. For R1, since the anatomical information is completely missing, there is no difference between the NLR-AP method and the NLR-IAP method, which is verified by the overlapped CRC mean-STD curves for both methods. For R2 which contained a perfectly matched anatomical boundary, the NLR-AP method still outperformed the NLR-IAP method in the low-resolution setting. However, the performance difference between the two methods in the high-resolution setting was minor. For R3 which contained false anatomical information, the NLR-AP method performed better than the NLR-IAP method as shown Fig. 7(d). By increasing the pixel resolution of the PET image as described in Sec. III, a better performance was achieved in all ROIs. Note that the results shown in Fig. 7 were measured in the low-resolution setting where the reconstructed high-resolution images were down-sampled

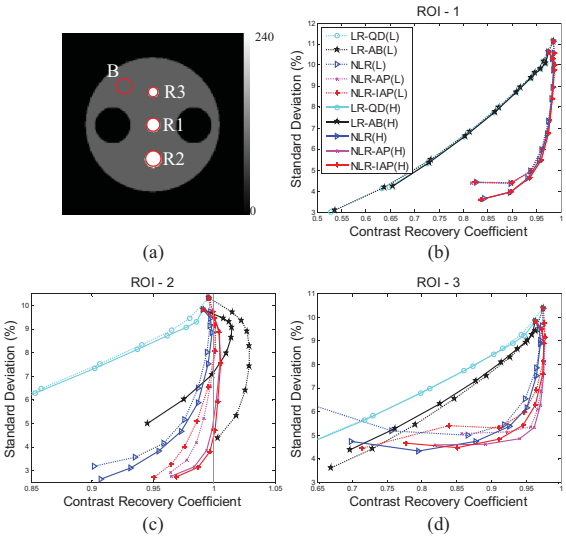

Fig. 7. Regional CRC-STD plots for ROI studies with phantom A: (a) Phantom with regional boundaries superposed where R1 has no corresponding anatomical boundary, R2 has a perfectly matched anatomical boundary, and R3 has a partially matched anatomical boundary; (b)-(d) Mean versus STD of CRC curves for the preselected 3 ROIs.

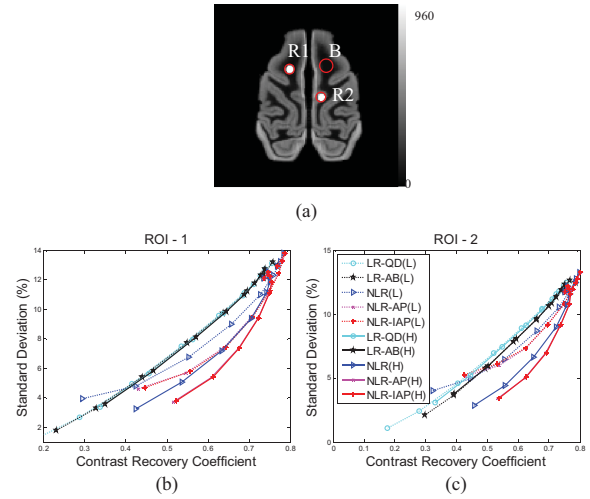

Fig. 8. Regional CRC-STD plots for ROI studies with phantom B: (a) Phantom with regional boundaries superposed where R1 has no corresponding anatomical boundary and R2 has a partially matched anatomical boundary; (b)-(c) Mean versus STD of CRC curves for the preselected 2 ROIs.

to the low-resolution images before computing the regional mean and STD. Comparison between the NLR method and the proposed methods (NLR-AP and NLR-IAP) shows that the proposed methods clearly outperform the NLR method in R2 where the anatomical boundary and the functional boundary are perfectly matched. When the anatomical boundary is completely missing as shown in R1, the NLR method and the proposed methods achieve the same performance. Finally, when the anatomical boundary and the functional boundary are not perfectly matched as shown in R3, with the properly chosen smoothing parameter, the proposed methods still perform better than the NLR method. These observations indicate that our proposed methods not only utilize the reliable anatomical information selectively, but also compromise the partially false anatomical information.

Fig. 8 shows regional CRC-STD curves for phantom B in which each point represents the mean and STD of the CRC evaluated from $N = 100$ independent noise trials. To observe the effect of smoothing over a range of γ, the nine different values were used; γ was set to 0, 0.0000125, 0.000025, 0.00005, 0.0001, 0.0002, 0.0004, 0.0008, 0.0016 for the low-resolution reconstructions and 0, 0.0000625, 0.000125, 0.00025, 0.0005, 0.001, 0.002, 0.004, 0.008 for the high-resolution reconstructions. In both ROIs, our proposed NLR-AP/IAP methods revealed better results with higher CRC than the LR-AB method and the NLR method without anatomical information. Similarly to the case of the phantom A, by increasing the pixel resolution of the PET image, a better performance was also achieved in all ROIs in this phantom.

V. Conclusions

We have developed a nonlocal regularization method that incorporates anatomical side information into PET reconstruction. In our method, to model the self-similarity property in both the underlying (PET) image and the anatomical image, a patch-based regularization term, which is based on nonlocal comparison of patches, was used. The experimental results show that our proposed method significantly improves both the quantitative and qualitative performance even with imperfect anatomical information or in the presence of signal mismatch between the PET image and the anatomical image. According to the quantitative results from our extended simulation study to increase the pixel resolution of the PET image up to the level of the high-resolution anatomical image, the high-resolution NLR-AP method further improves its performance when they are evaluated in the original low-resolution setting by down-sampling the reconstructed images.

Our NLR-AP method differs from the conventional methods in the following aspects: (i) It uses patch-based nonlocal regularization rather than conventional pixel-based local smoothing; (ii) The weight ω_{jk} in our nonlocal regularizer is modeled to adaptively signal the existence of roughness in the anatomical image only when there is similar roughness in the current estimate of the PET image. (Our method incorporates the anatomical information into the reconstruction process only when it is likely to be correlated to the PET data [4], thereby reducing the signal mismatch error.); (iii) It directly uses anatomical images without additional processes to extract anatomical boundaries or segmented regions [4], [5].

Our method is computationally expensive since it is time-consuming to update the values of the image pixels and the projection bins as shown in Table I. Fortunately, the COSEM-PL algorithm is parallelizable. All operations can be performed independently and simultaneously, either for all pixels or all projection bins. In our implementation, though it was not reported in this paper, we parallelized the COSEM-PL algorithm using the graphics processing unit (GPU) and achieved a significant improvement in computational speed. To maximize the computational efficiency of the GPU, we used

the single subset for the COSEM-PL algorithm. The details on the acceleration of the COSEM algorithm using GPU can be found in [46].

In conclusion, our proposed NLR-AP method has a great potential to enhance the accuracy of PET image reconstruction.

REFERENCES

[1] J. M. Ollinger and J. A. Fessler, "Positron-emission tomography," *IEEE Sign. Process. Mag.*, vol. 14, no. 1, pp. 43–55, Jan. 1997.

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

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

[4] G. Gindi, M. Lee, A. Rangarajan, and I. G. Zubal, "Bayesian reconstruction of functional images using anatomical information as priors," *IEEE Trans. Med. Imag.*, vol. 12, no. 4, pp. 670–680, Dec. 1993.

[5] C. Comtat, P. E. Kinahan, J. A. Fessler, T. Beyer, D. W. Townsend, M. Defrise, and C. Michel, "Clinically feasible reconstruction of 3D whole-body PET/CT scan using blurred anatomical labels," *Phys. Med. Biol.*, vol. 47, no. 1, pp. 1–20, 2002.

[6] A. Rangarajan, I.-T. Hsiao, and G. Gindi, "A Bayesian joint mixture framework for the integration of anatomical information in functional image reconstruction," *J. Math. Imag. Vis.*, vol. 12, no. 3, pp. 199–217, 2000.

[7] S. Somayajula, C. Panagiotou, A. Rangarajan, Q. Li, S. R. Arridge, and R. M. Leahy, "PET image reconstruction using information theoretic anatomical priors," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 537–549, Mar. 2011.

[8] J. Cheng-Liao and J. Qi, "PET image reconstruction with anatomical edge guided level set prior," *Phys. Med. Biol.*, vol. 56, no. 21, pp. 6899–6918, 2011.

[9] S. Sastry and R. E. Carson, "Multimodality Bayesian algorithm for image reconstruction in positron emission tomography: A tissue composition model," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 750–761, Dec. 1997.

[10] B. A. Ardekani, M. Braun, B. F. Hutton, I. Kanno, and H. Iida, "Minimum cross-entropy reconstruction of PET images using prior anatomical information," *Phys. Med. Biol.*, vol. 41, no. 11, pp. 2497–2517, 1996.

[11] K. Vunckx, A. Atre, K. Baete, A. Reilhac, C. M. Deroose, K. Van Laere, and J. Nuyts, "Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 599–612, Mar. 2012.

[12] D. W. Townsend, "Multimodality imaging of structure and function," *Phys. Med. Biol.*, vol. 53, no. 4, pp. R1–R39, 2008.

[13] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[14] J. Nuyts, "The use of mutual information and joint entropy for anatomical priors in emission tomography," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, Nov. 2007, pp. 4149–4154.

[15] J. Bowsher, H. Yuan, L. W. Hedlund, and T. G. Turkington, "Utilizing MRI information to estimate F18-FDG distributions in rat flank tumors," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, Oct. 2004, pp. 2488–2492.

[16] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model Simul.*, vol. 4, no. 2, pp. 490–530, 2005.

[17] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 36–51, Jan. 2009.

[18] G. Peyré, S. Bougleux, and L. Cohen, "Non-local regularization of inverse problems," *Inverse Problems Imag.*, vol. 5, no. 2, pp. 511–530, 2011.

[19] P. Arias, G. Facciolo, V. Caselles, and G. Sapiro, "A variational framework for exemplar-based image inpainting," *Int. J. Comput. Vis.*, vol. 93, no. 3, pp. 319–347, 2011.

[20] P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, "An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 27, no. 4, pp. 425–441, Apr. 2008.

[21] G. Adluru, T. Tasdizen, M. C. Schabel, and E. V. R. DiBella, "Reconstruction of 3D dynamic contrast-enhanced magnetic resonance imaging using nonlocal means," *J. Mag. Res. Imag.*, vol. 32, no. 5, pp. 1217–1227, 2010.

[22] J.V. Manjón, P. Coupé, A. Buades, V. Fonov, D. L. Collins, and M. Robles, "Nonlocal MRI upsamling," *Med. Image Anal.*, vol. 14, no. 6, pp. 784–792, 2010.

[23] F. Rousseau, "A non-local approach for image super-resolution using intermodality priors," *Med. Image Anal.*, vol. 14, no. 4, pp. 594–605, 2010.

[24] Z. Tian, X. Jia, B. Dong, Y. Lou, and S. B. Jiang, "Low-dose 4D CT reconstruction via temporal nonlocal means," *Med. Phys.*, vol. 38, no. 3, pp. 1359–1365, 2011.

[25] V.-G. Nguyen and S.-J. Lee, "Nonlocal-means approaches to anatomy-based PET image reconstruction," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, Oct. 2010, pp. 3273–3277.

[26] V.-G. Nguyen and S.-J. Lee, "Anatomy-based PET image reconstruction using nonlocal regularization," *Proc. SPIE*, vol. 8313, p. 83133T, Feb. 2012.

[27] G. Wang and J. Qi, "Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization," *IEEE Trans. Med. Imag.*, vol. 31, no. 12, pp. 2194–2204, Dec. 2012.

[28] A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," in *Proc. Int. Conf. Comput. Vis.*, 1999, pp. 1033–1038.

[29] Y. Lou, X. Zhang, S. Osher, and A. Bertozzi, "Image recovery via nonlocal operators," *J. Sci. Comput.*, vol. 42, no. 2, pp. 185–197, 2010.

[30] I. T. Hsiao, A. Rangarajan, and G. Gindi, "A new convergent MAP reconstruction algorithm for emission tomography using ordered subsets and separable surrogates," in *Proc. IEEE Int. Symp. Biomed. Imag. Conf. Rec.*, Jan. 2002, pp. 409–412.

[31] A. De Pierro and M. E. B. Yamagishi, "Fast EM-like methods for maximum 'a posteriori' estimates in emission tomography," *IEEE Trans. Med. Imag.*, vol. 20, no. 4, pp. 280–288, Apr. 2001.

[32] S. Ahn and J. A. Fessler, "Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms," *IEEE Trans. Med. Imag.*, vol. 22, no. 5, pp. 613–626, May 2003.

[33] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 601–609, Dec. 1994.

[34] A. De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Trans. Med. Imag.*, vol. 14, no. 1, pp. 132–137, Mar. 1995.

[35] A. Rangarajan, P. Khurd, I. T. Hsiao, and G. Gindi, "Convergence proofs for the COSEM-ML and COSEM-MAP algorithms," Med. Image Process. Lab, Dept. Radiol., SUNY Stony Brook, Univ., Stony Brook, NY, USA, Tech. Rep. MIPL-03-1, 2003.

[36] V.-G. Nguyen, S.-J. Lee, and M. N. Lee, "GPU-accelerated 3-D Bayesian image reconstruction from Compton scattered data," *Phys. Med. Biol.*, vol. 56, no. 9, pp. 2817–2836, 2011.

[37] I.-T. Hsiao, A. Rangarajan, P. Khurd, and G. Gindi, "An accelerated convergent ordered subsets algorithm for emission tomography," *Phys. Med. Biol.*, vol. 49, no. 11, pp. 2145–2156, 2004.

[38] S.-C. B. Lo, "Strip and line path integrals with a square pixel matrix: A unified theory for computational CT projections," *IEEE Trans. Med. Imag.*, vol. 7, no. 4, pp. 355–363, Dec. 1988.

[39] B. De Man and S. Basu, "Distance-driven projection and backprojection in three dimensions," *Phys. Med. Biol.*, vol. 49, no. 11, pp. 2463–2475, 2004.

[40] Y. Long, J.A. Fessler, and J.M. Balter, "3-D forward and back-projection for X-ray CT using separable footprints," *IEEE Trans. Med. Imag.*, vol. 29, no. 11, pp. 1839–1850, Nov. 2010.

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

[42] T. Gasser, L. Sroka, and C. J. Steinmetz, "Residual variance and residual pattern in nonlinear regression," *Biometrika*, vol. 73, no. 3, pp. 625–633, 1986.

[43] P. Gravel, G. Beaudoin, and J. De Guise, "A method for modeling noise in medical images," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1221–1232, Oct. 2004.

[44] H. H. Barrett, D. W. Wilson, and B. M. W. Tsui, "Noise properties of the EM algorithm: I. Theory," *Phys. Med. Biol.*, vol. 39, no. 5, pp. 833–846, 1994.

[45] J.-S. Liow and S. C. Strother, "Practical tradeoffs between noise, quantitation, and number of iterations for maximum likelihood-based reconstructions," *IEEE Trans. Med. Imag.*, vol. 10, no. 4, pp. 563–571, Jun. 1991.

[46] V.-G. Nguyen and S.-J. Lee, "Graphics processing unit-accelerated iterative tomographic reconstruction with strip-integral system model," *Opt. Eng.*, vol. 51, no. 9, pp. 093203-1–093203-11, 2012.

Van-Giang Nguyen received the B.S. degree in computer science from Vietnam Military Technical Academy, Vietnam, in 2005, and the M.S. and Ph.D. degrees in electronic engineering from Paichai University, Daejeon, Korea, in 2009 and 2012, respectively. Since 2013, he has been with the Department of Information Systems, Le Quy Don Technical University, Hanoi, Vietnam. His current research interests include image processing, computer vision, and their applications to tomographic reconstruction.

Soo-Jin Lee (M'92) received the B.S. and M.S. degrees in electronic engineering from Sogang University, Seoul, Korea, in 1984 and 1986, respectively, and the Ph.D. degree in electrical engineering from Stony Brook University, Stony Brook, NY, USA, in 1995. From 1986 to 1991, he was a Research Staff Member with LG Electronics, Inc., Anyang, Korea, working on the development of magnetic resonance imaging systems. In 1995, he was a Post-Doctoral Research Associate with the Department of Radiology, Stony Brook University. He was a Senior Research Scientist with the Korea Institute of Science and Technology, Seoul, in 1996. In 1997, he joined the Faculty of Paichai University, Daejeon, Korea, where he is currently a Professor of electronic engineering. From 2004 to 2005, he was a Visiting Associate Professor with the Department of Molecular and Medical Pharmacology, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA, USA. His current research interests include computational imaging and medical image processing with an emphasis on the tomographic reconstruction problem.