IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 20, 2021, accepted February 6, 2021, date of publication February 11, 2021, date of current version February 24, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3058807

Dynamic PET Image Reconstruction Incorporating Multiscale Superpixel Clusters

SHUANGLIANG CAO, YURU HE, HONGYAN ZHANG, WENBING LV, LIJUN LU, AND WUFAN CHEN

School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China
Guangdong Provincial Key Laboratory of Medical Image Processing, Southern Medical University, Guangzhou 510515, China

Corresponding authors: Lijun Lu (ljlubme@gmail.com) and Wufan Chen (chenwf@fimmu.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 81871437, and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2019A1515011104. The work of Lijun Lu was supported by the Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme, in 2018.

ABSTRACT Dynamic positron emission tomography (PET) image reconstruction is challenging due to the low-count statistics of individual frames. This study proposes a novel reconstruction framework aiming to enhance the quantitative accuracy of individual dynamic frames via the introduction of priors based on multiscale superpixel clusters. The clusters are derived from pre-reconstruction composite images using superpixel clustering followed by fuzzy c-means (FCM) clustering. A multiscale aggregation is exploited during the superpixel clustering to generate multiscale superpixel clusters. Then, maximum a posteriori (MAP) PET reconstruction with different-scale clusters is separately applied to individual frame and fused to generate the final result. Using realistic simulated dynamic brain PET data, the quantitative performance of the proposed method is investigated and compared with the maximum-likelihood expectation-maximization (MLEM), Bowsher method, and kernelized expectation-maximization (the kernel method). The proposed method achieves substantial improvements in both visual and quantitative accuracy (in terms of the signal-to-noise ratio and contrast versus noise performances). The method is also tested with a 60 min ^{18}F-FDG rat study performed with an Inveon small animal PET scanner. The proposed method is shown to outperform the other methods via improvements in visual and quantitative accuracy (in terms of noise versus the mean intensity of the region of interest).

INDEX TERMS Image reconstruction, maximum a posteriori, positron emission tomography, superpixel clustering.

I. INTRODUCTION

Positron emission tomography (PET), a functional imaging modality widely used in oncology, cardiology, and neurology, can measure radiotracer distribution *in vivo* [1]–[3]. Dynamic PET scan generally results in multiple time frames ranging from seconds to minutes in a series starting from the moment of tracer injection. With the addition of tracer kinetic modeling, dynamic PET is capable of quantifying physiologically or biochemically important parameters at the level of the voxel or region of interest (ROI) [4], [5]. However, accurate tracer kinetic modeling requires very short time frames, which result in low statistic counts and high noise in each frame [6]–[8].

The associate editor coordinating the review of this manuscript and approving it for publication was Yunjie Yang.

The tomography image reconstruction is inherently an ill-posed problem, which is further accentuated with low-count projection data from short individual time frames. Conventional analytical reconstruction, such as filtered back projection (FBP) [9], [10], often results in noisy images. Statistical image reconstruction, such as the maximum-likelihood expectation-maximization (MLEM) algorithm [11], [12], can exploit the statistical property of the detected data and produce improved reconstructed images compared with FBP. However, direct MLEM estimates of PET images often exhibit high variances at low counts with increasing iterations.

The maximum a posteriori (MAP) framework aims to tackle the ill-posedness inherent in PET image reconstruction through the incorporation of prior models [13]. Conventional prior models focus on local neighborhoods and subsequently

penalize inter-voxel intensity differences through different penalty functions such as the quadratic prior [14]. More sophisticated priors have been proposed for improved PET image reconstruction [15]–[18]. An alternative approach seeks to control the penalization of inter-voxel differences via the incorporation of information from segmented or non-segmented anatomical images. The classic anatomical prior voiding need for segmented anatomical information was proposed by Bowsher *et al*. [19]. This prior (referred to as the Bowsher prior) encourages smoothing over an anatomy-dependent neighborhood, defined by selecting a set of the most similar neighbors in the anatomical image [20]. Ehrhardt *et al*. derived a prior based on the structural similarity between PET and MR images measured by the alignment of PET and MR gradients [21].

Inspired by the kernel method, a model of PET image intensity at each pixel can be expressed as a function of a set of features and incorporated into the forward model of PET projection data. Thus, the coefficients of the features can be estimated by the maximum likelihood (ML), and the corresponding PET image can be derived. This novel PET image reconstruction is named kernelized expectation-maximization (KEM). The features can be derived from the composite images for dynamic PET imaging [22], MR images for static PET imaging [23], PET and MRI information for the list-mode reconstruction of static images [24], and the composite images and sinograms for high temporal-resolution dynamic PET reconstruction [25]. KEM reconstruction has also been applied in direct parametric imaging by combining the spatial kernel with tracer kinetic models such as the spectral model [26] or the Patlak graphical model [27].

In the different context of post-reconstruction dynamic PET image analysis, a number of clustering-based techniques have been proposed to reduce noise in kinetic analysis. Factor analysis seeks to decompose dynamic cardiac PET images into different tissue types based on their unique temporal signatures to improve quantification of physiological function [28]–[30]. In oncologic whole-body imaging, the principal component analysis (PCA) approach has been used to enhance the distinction of tumors in dynamic FDG images compared with conventional static standard uptake value (SUV) images [31]. Janssen *et al*. applied k-means clustering to the slopes of time activity curves (calculated based on the last few time frames of FDG uptake) to differentiate between tumors and healthy tissues [32]. The k-means and Gaussian Mixture Models clustering algorithms were applied prior to voxel-wise kinetic modeling, which improves the classification of time activity curves between active and non-active neurotransmitter states [33].

In contrast to the abovementioned post-reconstruction methods, our group proposed a "3.5D" image reconstruction utilizing clustering to enhance the image reconstruction via the generation of clusters of neighborhood [34]. As an extension of the previous work, the present study utilizes two-step clustering, including superpixel clustering and fuzzy c-means (FCM) clustering. The advantage of two-step clustering is that on the one hand it can exploit the local spatial information by superpixel clustering while on the other hand the global feature information is utilized by FCM clustering. The clustering result is used as a prior for the reconstruction to encourage PET image smoothness within the same cluster. Moreover, multiscale aggregation is performed to further reduce bias and noise. The proposed method can combine the advantages of superpixel clustering, FCM clustering, and multiscale aggregation and thus can make good use of the image prior. The reconstruction performance of the proposed method is compared with the MLEM algorithm, the Bowsher method and the kernel method [22] in simulated and real clinical data reconstruction. The main contributions of this work are as follows.

- This study proposes a novel reconstruction framework aiming to enhance the quantitative accuracy of individual dynamic frames via the introduction of priors based on multiscale superpixel clusters.
- The clusters are derived from pre-reconstruction composite images using superpixel clustering followed by FCM clustering. A multiscale aggregation is exploited during the superpixel clustering to generate multiscale superpixel clusters.
- MAP PET reconstruction with different-scale clusters is separately applied to individual frame and fused to generate the final result.

The rest of this paper is organized as follows. We first introduce the MAP reconstruction framework for PET imaging and describe how the proposed reconstruction method is performed in Section II. Then, we present a computer simulation study in Section III to validate the improvement of the proposed method over existing methods. Section IV presents the results of applying the new method to real clinical data. Finally, a discussion is provided in Section V, and conclusions are drawn in Section VI.

II. THEORY

A. MAP IMAGE RECONSTRUCTION

PET projection data $\boldsymbol{y}$ are well modeled as independent Poisson random variables with the log-likelihood function:

$$L(\boldsymbol{y}|\boldsymbol{x}) = \sum_{i=1}^{M} y_i \log \bar{y}_i - \bar{y}_i - \log y_i!, \tag{1}$$

where $\bar{\boldsymbol{y}}$ is the expectation of the projection data and related to the unknown image $\boldsymbol{x}$ by

$$\bar{\boldsymbol{y}} = \boldsymbol{A}x + \boldsymbol{r}, \tag{2}$$

where $\boldsymbol{A} \in \mathbb{R}^{M\times N}$ is the system matrix with elements a_{ij} denoting the probability of a positron emitted from pixel j resulting in a coincidence at the ith detector pairs, and $\boldsymbol{r}$ represents the expectation of random and scattered events. M and N denote the total number of detector pairs and pixels, respectively. The ML estimate of image $\boldsymbol{x}$ is obtained by

maximizing the log-likelihood function above:

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}\geq 0} L(\boldsymbol{y}|\boldsymbol{x}). \tag{3}$$

However, the ML estimate can be very noisy at convergence. This problem is commonly solved by incorporating an image prior into the MAP framework:

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}\geq 0} L(\boldsymbol{y}|\boldsymbol{x}) - \beta U(\boldsymbol{x}), \tag{4}$$

where β is a regularization parameter that controls the trade-off between resolution and noise, and $U(\boldsymbol{x})$ is the energy function, commonly known as a prior. To derive the MAP estimate, the preconditioned gradient ascent algorithm [18] is used instead of the one-step-late (OSL) approach [16] to avoid the numerical problems that the latter may cause. The resulting iterative update to the MAP estimate is as follows:

$$x_j = x_j^{old} + \left(\frac{\partial L}{\partial x_j} - \beta\frac{\partial U}{\partial x_j}\right) \Bigg/ \left(\frac{\sum_i a_{ij}}{x_j^{old}} + \beta\frac{\partial^2 U}{\partial x_j^2}\right), \tag{5}$$

where all partial derivatives are evaluated in the current reconstruction $\boldsymbol{x}^{old}$, and x_j is the new activity estimate at pixel j. The performance of the MAP reconstruction strongly depends on $U(\boldsymbol{x})$ and regularization parameter β.

B. GENERATION OF THE PRIOR MODEL

1) PROPOSED CLUSTER-BASED PRIOR

To make use of more pixels to further encourage smoothing without causing significant bias, we expand the use of local neighborhoods to those containing all pixels with similar intensity values as clustered together. The prior is defined as:

$$U(\boldsymbol{x}) = \sum_{j=1}^{N}\sum_{k\in c\{j\}} w_{jk}\psi(x_j - x_k), \tag{6}$$

where $c\{j\}$ stands for the cluster in which pixel j is grouped (a detailed description of which can be found in the next section). ψ is the potential function and takes the form of $\psi(s) = s^2$, and w_{jk} is the weight of a given pixel k in the neighborhood $c\{j\}$ of pixel j. As in the previous study, w_{jk} is inversely proportional to the Euclidean distance between pixels j and k. We compute w_{jk} in a particular neighborhood (e.g., 9×9 as in the following simulation experiment) instead of the entire cluster. The weights of more distant pixels within the cluster are negligible, leading to a considerable increase in computational speed.

2) CLUSTERING

The construction of the prior strongly depends on the cluster $c\{j\}$, which determines how accurately the prior describes the nature of the image. In our previous work [34], FCM clustering was performed based on time activity curve of each pixel. However, FCM is sensitive to noise because the local spatial information of pixels is not utilized. Here, we propose a two-step clustering: superpixel clustering followed by FCM clustering. The advantage of the proposed clustering scheme is that superpixel clustering can exploit local information, while FCM can exploit global feature information.

First, superpixel clustering is performed based on the three composite images reconstructed from the summed sinograms of multiple time frames of a dynamic PET. In this study, superpixel clustering is performed using the simple linear iterative clustering (SLIC) method [35] for its computational and memory efficiency and good adherence to the image boundaries. We use the composite images as the input to SLIC, and thus the weighted distance measure $D_{j,l}$ between pixels j and l is defined by:

$$D_{j,l} = \sqrt{d_f^2 + \left(\frac{d_s}{S}\right)^2 m^2}, \tag{7}$$

where d_s is the spatial Euclidean distance between pixels j and l, and $d_f = \sqrt{\sum_{h=1}^{H}(f_j^h - f_l^h)^2}$ is the pixel intensity similarity measure between pixels j and l, where H stands for the number of image channels and f_j^h is the activity value of pixel j in the hth channel. In the present study, H is set to three. $S = \sqrt{(N/K)}$ is the mean superpixel width as a normalization factor, where N is the number of pixels and K is the number of superpixels. Finally, m is the weight between the intensity similarity and spatial proximity. A smaller m results in better boundary adherence with a less regular size and shape.

Then, FCM clustering [36], [37] is carried out to classify superpixels according to the mean image intensity of every superpixel region in the three composite images. The classification number T of FCM can be set to the number of main tissues; for example, T can be three for the human brain data because three main clusters (gray matter, white matter, and background) would be found in the reconstructed image. For a given T value, the FCM clustering minimizes

$$J = \sum_{p=1}^{K}\sum_{q=1}^{T} u_{pq}^{v}\left\|b_p - c_q\right\|^2, \tag{8}$$

where b_p is the vector comprised of the mean intensities of the pth superpixel region in all composite images, c_q is the center of the qth cluster, u_{pq} is the degree of membership of b_p in the qth cluster, and v is the fuzzy partition matrix exponent, used for controlling the degree of fuzzy overlap.

After FCM clustering, a membership matrix is obtained and used to determine each superpixel's classification. However, it is not suitable to directly classify all superpixels into T clusters according to the maximum membership value because some small-sized tissues, e.g., tumors, do not belong to these T clusters. Moreover, it is difficult to make small-sized tumor tissues classified into a separate cluster only by selecting a larger value of T. Given that the superpixels in a tumor region have a relatively low maximum membership, we set a threshold τ to distinguish these superpixels from the T main clusters. The τ value can range from 0 to 1 and should be the lowest value that can ensure distinguishing of the superpixels. Each superpixel with a lower maximum membership than the threshold τ is treated as a cluster separate

from the T clusters. The other superpixels are then divided into T clusters according to the maximum membership.

3) MULTISCALE AGGREGATION

When only single-scale superpixel clustering followed by FCM clustering is performed, the resulting MAP estimate has much bias from the true image. To reduce the bias, multiscale aggregation can be performed as an average of multiple estimates based on multiscale superpixel clustering [38]. For the SLIC method, these multiscale superpixels are generated by varying the number of superpixels that decides the mean superpixel size. Each scaled superpixel clustering is followed by FCM clustering, and the corresponding MAP estimate is derived. Assuming N_S scales are used, the average of all estimates is the final reconstructed result $\boldsymbol{x}^*$:

$$\boldsymbol{x}^* = \frac{1}{N_S}\sum_{n=1}^{N_S}\hat{\boldsymbol{x}}_n, \tag{9}$$

where $\hat{\boldsymbol{x}}_n$ is the image estimate based on the nth single-scale superpixel clustering.

4) IMPLEMENTATION OF THE PROPOSED DYNAMIC PET RECONSTRUCTION

The flowchart of the proposed reconstruction algorithm is shown in Fig. 1. It consists of five steps, elaborated as follows:

Step 1: The dynamic PET sinogram data are rebinned into three composite frames with the same duration. MLEM reconstruction is first used to generate three composite images $\hat{\boldsymbol{x}}_1^{\text{reb}}$, $\hat{\boldsymbol{x}}_2^{\text{reb}}$, and $\hat{\boldsymbol{x}}_3^{\text{reb}}$. Then, each composite image is normalized by dividing its activities by the standard deviation of the activities.

Step 2: Using the composite images, SLIC superpixel clustering is performed according to the number of superpixels K and the compactness factor m.

Step 3: The K superpixels are clustered by FCM (8). Each superpixel is classified according to the maximum membership and the threshold τ.

Step 4: The final clustering result from Step 3 is used as the cluster prior (6) for the MAP reconstruction (5), and a single-scale image estimate $\hat{\boldsymbol{x}}_n$ is obtained.

Step 5: Varying K in the given range, steps 2-4 are repeated. Finally, all scaled image estimates are derived, and their average is the reconstruction result $\boldsymbol{x}^*$ (9) of the proposed method.

III. VALIDATION USING COMPUTER SIMULATION

A. SIMULATION SETUP

The geometry of a GE DST whole-body PET scanner was simulated for dynamic PET scans. A 2D simulated PET phantom (image dimensions 217 × 217) was constructed based on an anatomical model from the BrainWeb database [39]. We aimed to evaluate the reconstruction performance of all methods for the lesions with different sizes and locations. Three different sized tumors were added to the PET phantom, as shown in Fig. 2(a). The largest and smallest tumors were located in the white matter, whereas the medium-sized tumor was positioned across the white matter and gray matter. The scanning schedule was as follows: 4 × 20 s, 4 × 40 s, 4 × 60 s, 4 × 180 s, and 8 × 300 s, which resulted in 24 time frames. The regional time activity curves shown in Fig. 2(c) were assigned to different brain regions. Dynamic activity images were first forward projected to generate noise-free sinograms, and then Poisson noise was introduced to generate the projection data. A 20% uniform background was included to simulate random events. Scatters were not simulated in this study. An attenuation image for 511 keV photons was created using the anatomical model by assigning a value of 0 cm^{-1} to air, 0.146 cm^{-1} to bone, and 0.096 cm^{-1} to other tissues, as shown in Fig. 2(b). Attenuation and random corrections were included into all the reconstruction methods to obtain quantitative images. A total of 3×10^7 events were acquired over 60 min, and ten noisy realizations were simulated.

The 24 frames were rebinned into three high-count composite frames, corresponding to the first 20 min, middle 20 min, and last 20 min. All composite frames were reconstructed using the MLEM algorithm with 100 iterations. The three composite images provided the prior information used in the Bowsher method, kernel method, and proposed method. The neighborhood size was set to 9 × 9 for all methods using the prior images. The feature vector, composed of three intensities at the same pixel positions of the three composite images, was used to search the 20 nearest neighbors for the Bowsher method and the 50 nearest neighbors for the kernel method. For the proposed approach, the number of superpixels K ranged from 1000 to 2600 in steps of 100 for multiscale SLIC, the compactness factor m was fixed to 70 without varying with K value, the cluster number T of FCM was set to 3, and the threshold τ was set to 0.7. Both the Bowsher method and proposed method used a quadratic potential function, and 300 iterations were used to ensure convergence. The MLEM algorithm with post-reconstruction Gaussian filtering was also included for performance comparison. Both the MLEM algorithm and the kernel method are EM-based methods.

B. EVALUATION METRICS

In the quantitative comparison between the different reconstruction methods, the overall image quality was assessed by the signal-to-noise ratio (SNR):

$$\text{SNR} = 10\log_{10}\left(\frac{\left\|x^{\text{true}}\right\|^2}{\left\|x^i - x^{\text{true}}\right\|^2}\right), \tag{10}$$

where x^{true} and x^i denote the ground-truth and reconstructed images, respectively.

The three tumors were chosen as the regions of interest (ROIs), and the contrast recovery coefficient (CRC) is calculated by:

$$\text{CRC} = \frac{1}{N_r}\sum_{i=1}^{N_r}\left(\frac{\bar{R}_i}{\bar{B}_i} - 1\right)\Bigg/\left(\frac{\bar{R}_{\text{true}}}{\bar{B}_{\text{true}}} - 1\right), \tag{11}$$

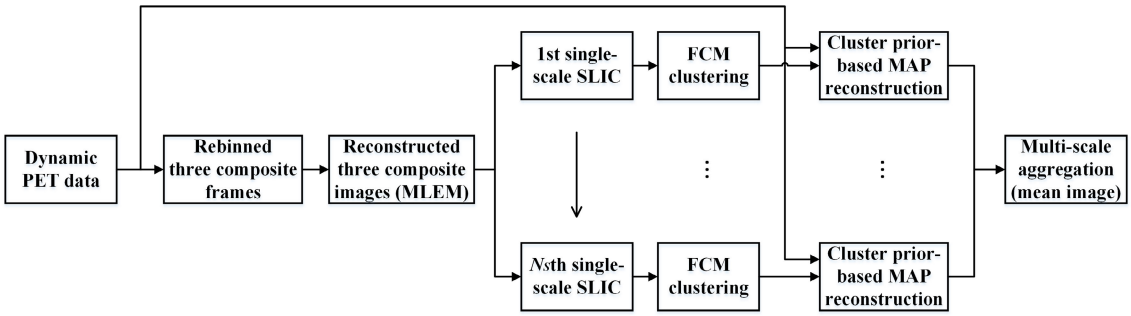

FIGURE 1. The flowchart of the proposed PET reconstruction algorithm.

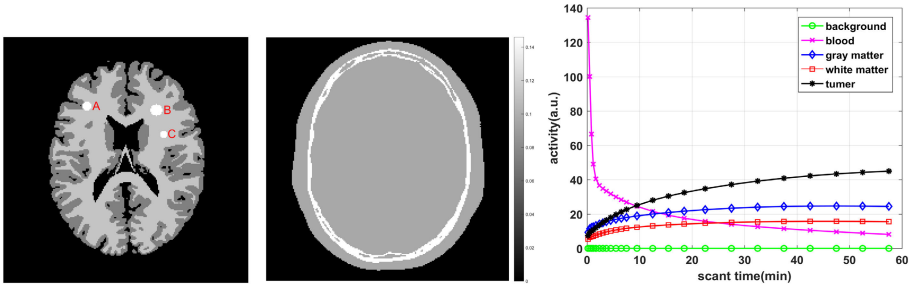

FIGURE 2. Digital brain phantom, attenuation map, and time activity curves. (a) PET phantom composed of the gray matter, white matter and three tumors A, B, C. (b) the corresponding attenuation image. (c) Regional time activity curves.

where N_r is the total number of noisy realizations (N_r= 10 in this simulation), $\bar{R}_i$ is mean intensity of the ROI in the ith realization, $\bar{B}_i$ is the mean intensity of the background, $\bar{R}_{\text{true}}$ is the true ROI mean, and $\bar{B}_{\text{true}}$ is the true background mean. The white matter region was eroded by a 5×5 square structuring element and chosen as the background to calculate the noise standard deviation (SD). The SD was also averaged over all realizations.

The spatially averaged pixel-level normalized bias for each ROI is defined as:

$$\text{Bias}=100\% \times \frac{1}{N_{\text{ROI}}} \sum_{j \in \text{ROI}} \frac{\left| \bar{x}_j - x_j^{\text{true}} \right|}{x_j^{\text{true}}}, \tag{12}$$

where N_{ROI} is the total number of pixels in the ROI and $\bar{x}_j$ is the ensemble mean value of pixel j (i.e., $\bar{x}_j = \frac{1}{N_r} \sum_{i=1}^{N_r} x_j^i$).

The spatially averaged pixel-level coefficient of variation (COV) for each ROI is defined as:

$$\text{COV}=100\% \times \frac{1}{N_{\text{ROI}}} \sum_{j \in \text{ROI}} \frac{\sqrt{\frac{1}{N_r - 1} \sum_{i=1}^{N_r} (x_j^i - \bar{x}_j)^2}}{\bar{x}_j}. \tag{13}$$

C. SIMULATION RESULTS

We mainly compared the reconstruction performance of the different methods for frames 12 and 24, which have 406 k and 2752 k events, respectively.

Fig. 3 shows the true activity image and the images reconstructed by the four methods for the 12th and 24th frames. For each method, a reconstructed image is shown with the highest SNR by varying either the iteration number for the EM-based methods or the regularization parameter β for the regularization-based methods. Frame 12 had a much lower scan count than frame 24, and thus its reconstruction results showed larger differences. The images from the MLEM algorithm had a high level of noise. By incorporating a prior, the kernel method, Bowsher method, and proposed method had substantially reduced noise and better preserved edges and tumors than the MLEM algorithm with post-reconstruction Gaussian filtering for the two studied frames. The proposed method achieved the highest SNR and an optimal visual effect, including the sharpest boundaries between the gray matter and white matter and the least noise in the white matter. The corresponding bias images are shown in Fig. 4. The proposed method had the least bias, especially in the white matter region.

Fig. 5 shows the contrast recovery coefficients (CRCs) of the three tumor regions versus background standard deviation (SD) trade-off achieved with the different methods, obtained by varying either the iteration number or regularization parameter in each method for the 12th and 24th frames. As expected, the methods that utilized a prior had lower background noise than the MLEM algorithm with Gaussian postfiltering at matched CRC level. The proposed approach achieved the highest CRC with the lowest back-

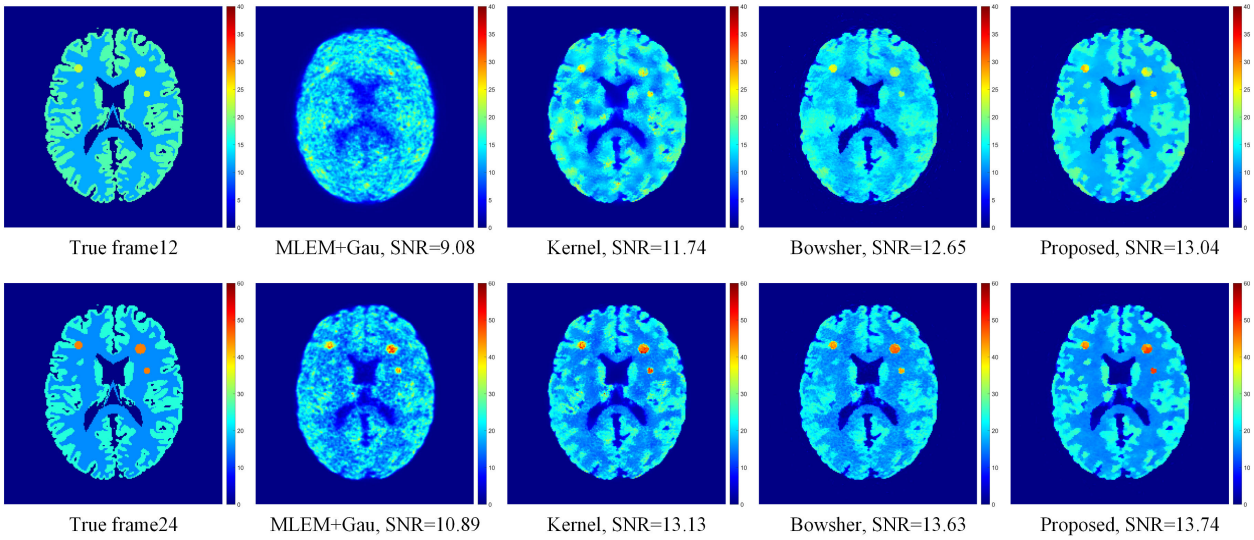

FIGURE 3. True activity image and reconstructed images by different methods for the 12th frame (top row), and 24th frame (bottom row). The reconstruction methods include (from left to right) the MLEM with gaussian postfiltering, kernel method, Bowsher method and proposed approach.

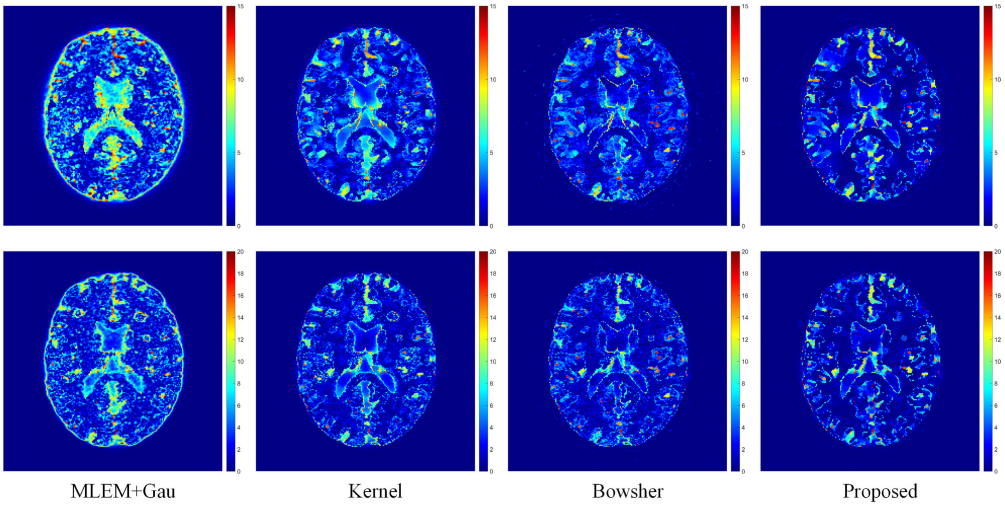

FIGURE 4. Bias images for the 12th frame (top row), and 24th frame (bottom row).

ground noise for all tumors. One of the key reasons why the proposed method can preserve a higher tumor contrast than other methods is that in each single-scale MAP reconstruction process, the smoothness of the tumor region activities is limited in the superpixel regions corresponding to that tumor and does not occur across that tumor and other tissues.

Fig. 6 shows the spatially averaged pixel-level bias versus spatially averaged pixel-level COV trade-off of the three tumor regions for the 12th and 24th frames. The methods incorporating a prior achieved better bias-variance performance than the post-smoothed MLEM algorithm for all tumors. The proposed method had the optimal bias-variance trade-off for the most cases and slightly poorer trade-off than the Bowsher method only for tumor B (the largest tumor) of the 24th frame.

Fig. 7 shows the plots of the image SNR of all time frames for the different methods. The SNR is the mean of the SNRs over the ten noisy realizations, and the error bars indicate the standard deviation of the SNRs over the ten realizations. All methods utilizing a prior outperformed the post-smoothed MLEM algorithm for all frames. The proposed approach achieved higher SNRs than the Bowsher method and kernel method, especially for the low-count frames.

D. ILLUSTRATION ABOUT THE EFFECT OF EACH COMPONENT

In our proposed method, three components, i.e., superpixel clustering, FCM clustering, and multiscale aggregation, were combined for PET image reconstruction. We use the reconstruction of the simulated 12th frame as an example to briefly illustrate the effect of these components. Fig. 8 shows the true

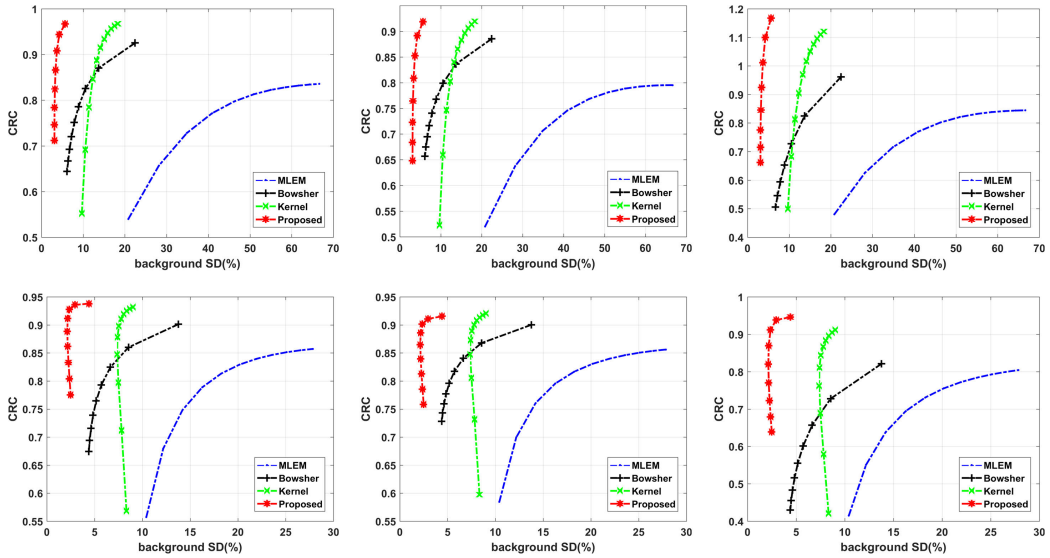

FIGURE 5. CRCs of three tumor regions versus background SD trade-off achieved by different methods for frames 12 (top row), and 24 (bottom row). (From left to right): correspond to the tumors A, B, and C.

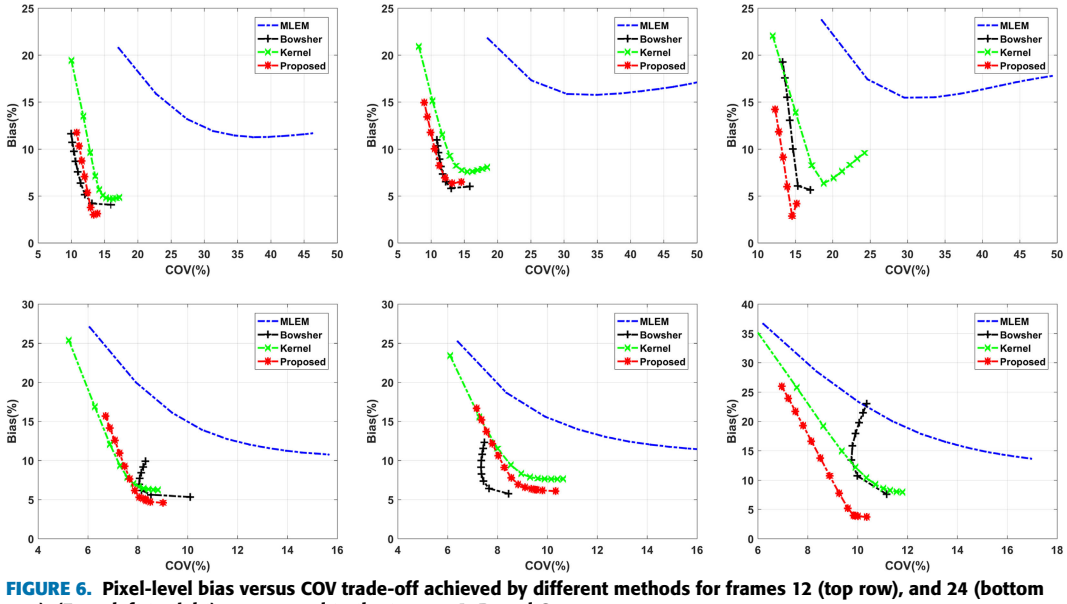

FIGURE 6. Pixel-level bias versus COV trade-off achieved by different methods for frames 12 (top row), and 24 (bottom row). (From left to right): correspond to the tumors A, B, and C.

image and images reconstructed by (a) the MLEM algorithm without any components utilized and the cluster prior-based MAP method with (b) only superpixel clustering ($K = 1500$) utilized, (c) both superpixel clustering ($K = 1500$) and FCM clustering utilized, and (d) all three components utilized (i.e., the proposed method). All MAP methods used the same component parameters, and the regularization parameter β was tuned for each reconstruction. Fig. 8 (b) presents better edges than (a) by using a superpixel clustering-based prior. Fig. 8 (c) illustrates less noise and better tissue contrast than (b) by adding FCM clustering. However, high bias persists in Fig. 8 (c), such as the two regions indicated by red arrows. The bias level was substantially decreased owing to the effect of multiscale aggregation, as shown in Fig. 8 (d). The SNR of Fig. 8 (a)-(d) increases as more components are used.

IV. APPLICATION TO REAL CLINICAL DATA

A. CLINICAL DATA ACQUISITION

A rat with myocardial infarction was scanned by a Siemens Inveon small animal PET scanner. The rat received a bolus injection of 1.15 mCi ^{18}F-FDG. The PET scan started right at the injection and lasted 60 min. The dynamic PET data were divided into 25 frames: 10×3 s, 3×10 s, 4×60 s, 5×300 s, and 3×600 s. A CT scan was acquired to provide the attenuation map for attenuation correction. Attenuation factors were extracted using the vendor software and included in the forward model during reconstruction.

As mentioned in the simulation study, three 20 min composite frames were reconstructed to provide the image prior. The rat data were reconstructed into an image matrix of $217 \times 217 \times 159$ voxels with a voxel size of

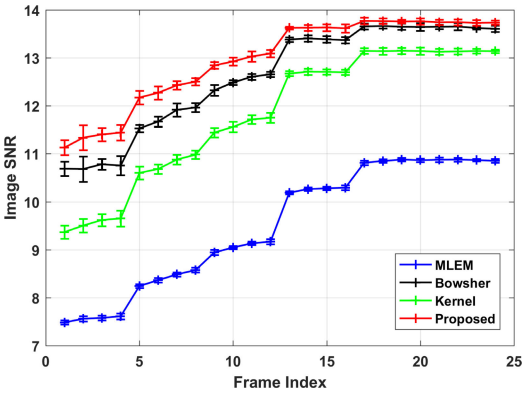

FIGURE 7. Plots of image SNR of all time frames reconstructed by different methods.

$0.46 \times 0.46 \times 0.796$ mm^3 using the MLEM algorithm, kernel method, Bowsher method, and proposed method. For the proposed method, the parameters m and T were set to 50 and 5, respectively. Other parameters were kept the same as in the simulation study.

Since the true activity values were unknown, bias analysis was not performed. To compare the different methods, we plotted the mean activity versus standard deviation (SD) curves of the ROI (the myocardium region indicated by the white arrows in Fig. 9).

B. RESULTS

Fig. 9 shows a transverse slice of reconstructed images of the 18th and 25th frames using the different methods. For the MLEM algorithm and kernel method, 100 iterations were used. For the Bowsher method and proposed method, the regularization parameter β was heuristically selected and set to 0.01 and 0.005 for the 18th and 25th frames, respectively. The reconstructed images from the MLEM algorithm suffered from considerable noise. By using the composite image prior, the kernel method suppressed noise to a certain extent and preserved the tissue contrast. The Bowsher method partly reduced the noise but over smoothed the image structures. In comparison, the proposed method presented better noise suppression and preserved boundaries and tissue contrasts better than other methods.

Fig. 10 shows the mean activity versus standard deviation trade-off of the ROI using the different methods by varying the iteration number or regularization parameter in each method for the 18th and 25th frames. For the myocardium region, the proposed method achieved less noise at the same activity value than other methods.

V. DISCUSSION

A. PARAMETERS SELECTION

The performance of the proposed method is affected by the number of superpixels K and the compactness factor m used for superpixel clustering, the cluster number T and threshold τ used for FCM clustering, and the scale selection for multiscale aggregation. The number of superpixels K should be neither too large nor too small. A large K value results in meaningless, small-sized superpixels, while a small K value leads to large-sized superpixels with excessive loss of detail resulting from some superpixels containing more than one class of tissue. However, for a given image size, an appropriate set of K values could be determined, e.g., an image of size 217×217, it would be appropriate to set K between 1000 and 3000. In other words, the suitable number of superpixels for the SLIC algorithm is proportional to the total number of image pixels. It is easy to select some scales from the range of appropriate K values because the proposed method is not sensitive to multiscale selection. The cluster number T could be equal to or slightly greater than the number of main tissues in the image to be reconstructed; e.g., in our simulation study, T can be 3, 4, 5, or 6 without obvious improvements or degradations of image quality. This finding is consistent with our previous study [34]. Both the compactness factor m and threshold τ are related to the quality of the prior image. The m value and τ value should be small when the composite prior images have less noise, and vice versa. In practice, we prefer to use a smaller m value for better boundary adherence. Thus, for image reconstruction with the proposed method, we can easily select parameters K and T and multiple scales from experience and only need to carefully choose the parameters m and τ.

B. IMPLEMENTATION OF MULTISCALE AGGREGATION

As investigated in [38], multiscale superpixels can be generated by varying the number of superpixels K and the compactness factor m in the SLIC algorithm. However, varying the K value is much more efficient than varying the m value for eliminating the bias according to our experiments. This may be because it is easier to capture the different-scale structural information by varying the K value [40] and because the reconstructed image quality is susceptible to the m value. Thus, we implemented multiscale over-segmentations only by varying the K value. In the simulation study, no fewer than five scales ranging from 1000 to 3000 should be selected for varying the K values. For each single-scale MAP estimate, we used the same compactness factor m, cluster number T, threshold τ, and regularization parameter β to avoid adding new tunable parameters. Furthermore, the computational time of the proposed method is proportional to the number of chosen scales and can be reduced by parallel computing.

C. DISADVANTAGE OF THE PROPOSED METHOD

According to the analysis of the bias images in Fig. 4, the proposed method cannot reconstruct slender structures well. This is because the SLIC algorithm cannot segment the slender structure well from the surroundings.

D. DEEP IMAGE PRIOR

With the development of deep learning, some deep prior models [41], [42] have been proposed for PET reconstruction. These methods incorporate the deep image prior [43] into static PET image reconstruction [41] and dynamic PET image reconstruction using non-negative matrix factorization

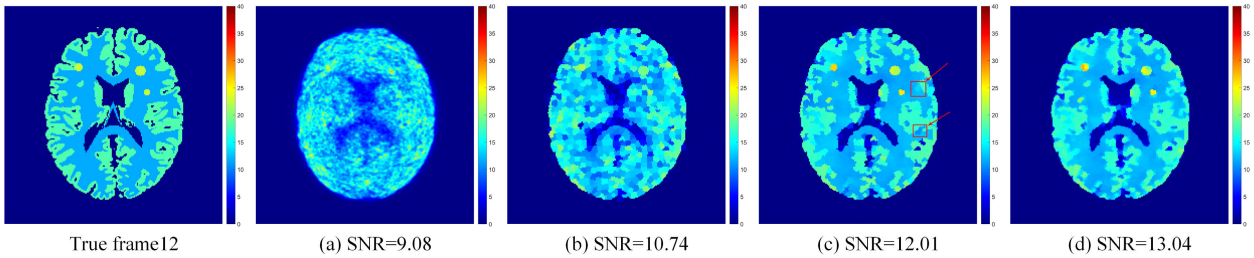

FIGURE 8. True image of frame 12 and reconstructed images by different methods based on different combinations of three components.

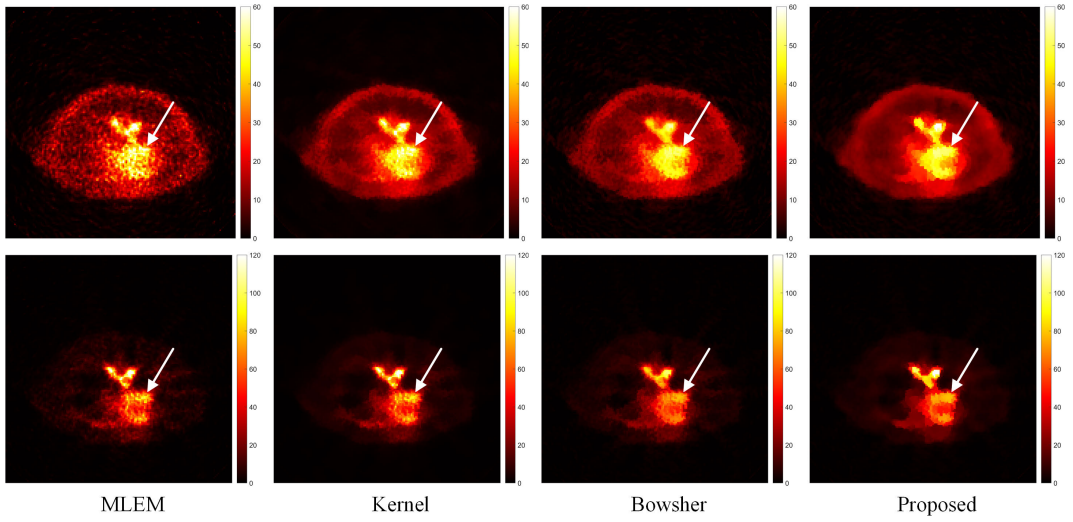

FIGURE 9. Reconstructed images of the 18th (top row) and 25th (bottom row) frames using different methods. The myocardium region was indicated by the white arrows.

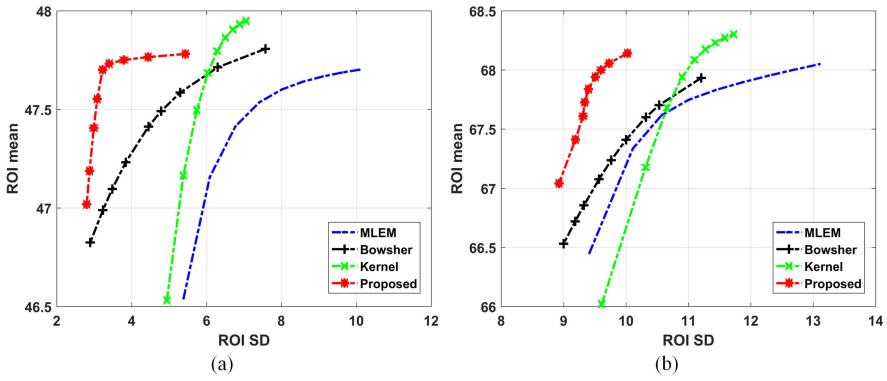

FIGURE 10. Mean activity versus SD trade-off of myocardium ROI achieved by different methods for the (a) 18th and (b) 25th frames. The curves were plotted by varying the iteration number from 20 to 100 for the MLEM algorithm and kernel method and varying the regularization parameter for the Bowsher method and proposed method.

[42] and show better performance visually and quantitatively than some traditional reconstruction algorithms. These deep prior-based methods will be studied and compared with our proposed method in the future.

VI. CONCLUSION

In this paper, we developed a novel cluster prior-based MAP reconstruction method. The clustering consists of superpixel clustering followed by FCM clustering, and multiscale aggregation is used to further improve the image quality. The simulation results show the proposed approach achieved better visual effects and CRC versus SD trade-off than the MLEM algorithm, kernel method, and Bowsher method. The proposed approach can preserve edges and tumors well with the least background noise. The proposed method was applied to reconstruct the dynamic PET images of a rat

with myocardial infarction and achieved visual and quantitative accuracy improvements over the other methods. This reconstruction method can be easily extended to utilize a multimodal anatomical prior or a combination of anatomical and functional priors by changing the input channel of the SLIC algorithm. Applications to human clinical data and the corresponding parameter selection strategy will be studied in the future.

REFERENCES

[1] G. J. Kelloff, J. M. Hoffman, B. Johnson, H. I. Scher, B. A. Siegel, E. Y. Cheng, B. D. Cheson, J. O'Shaughnessy, K. Z. Guyton, D. A. Mankoff, L. Shankar, S. M. Larson, C. C. Sigman, R. L. Schilsky, and D. C. Sullivan, "Progress and promise of FDG-PET imaging for cancer patient management and oncologic drug development," *Clin. Cancer Res.*, vol. 11, no. 8, pp. 2785–2808, Apr. 2005.

[2] T. S. Khan, A. Sundin, C. Juhlin, B. Långström, M. Bergström, and B. Eriksson, "11 C-metomidate PET imaging of adrenocortical cancer," *Eur. J. Nucl. Med. Mol.*, vol. 30, no. 3, pp. 403–410, 2003.

[3] P. Vakil, J. J. Lee, J. J. Mouannes-Srour, C. P. Derdeyn, and T. J. Carroll, "Cerebrovascular occlusive disease: Quantitative cerebral blood flow using dynamic susceptibility contrast MR imaging correlates with quantitative H2[^{15}O] PET," *Radiology*, vol. 266, no. 3, pp. 879–886, Mar. 2013.

[4] H. Watabe, Y. Ikoma, Y. Kimura, M. Naganawa, and M. Shidahara, "PET kinetic analysis—Compartmental model," *Ann. Nucl. Med.*, vol. 20, no. 9, pp. 583–588, Nov. 2006.

[5] M. Bentourkia and H. Zaidi, "Tracer kinetic modeling in PET," *PET Clinics*, vol. 2, no. 2, pp. 267–277, Apr. 2007.

[6] A. Dimitrakopoulou-Strauss, L. Pan, and L. G. Strauss, "Parametric imaging: A promising approach for the evaluation of dynamic PET-18 F-FDG studies-the DKFZ experience," *Hell J. Nucl. Med.*, vol. 13, no. 1, pp. 18–22, 2010.

[7] N. A. Karakatsanis, M. A. Lodge, A. K. Tahari, Y. Zhou, R. L. Wahl, and A. Rahmim, "Dynamic whole-body PET parametric imaging: I. Concept, acquisition protocol optimization and clinical application," *Phys. Med. Biol.*, vol. 58, no. 20, pp. 7391–7418, Oct. 2013.

[8] N. A. Karakatsanis, M. A. Lodge, Y. Zhou, R. L. Wahl, and A. Rahmim, "Dynamic whole-body PET parametric imaging: II. task-oriented statistical estimation," *Phys. Med. Biol.*, vol. 58, no. 20, pp. 7419–7445, Oct. 2013.

[9] L. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," *IEEE Trans. Nucl. Sci.*, vol. 21, no. 3, pp. 21–43, Jun. 1974.

[10] P. E. Kinahan and J. G. Rogers, "Analytic 3D image reconstruction using all detected events," *IEEE Trans. Nucl. Sci.*, vol. 36, no. 1, pp. 964–968, 1989.

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

[12] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *J. Comput. Assist. Tomogr.*, vol. 8, no. 2, pp. 306–316, Apr. 1984.

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

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

[15] S. Geman and D. E. McClure, "Statistical methods for tomographic image reconstruction," *Bull. Int. Stat. Inst.*, vol. 4, pp. 5–21, 1987.

[16] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imag.*, vol. 9, no. 1, pp. 84–93, Mar. 1990.

[17] E. Ü. Mumcuoglu, R. M. Leahy, and S. R. Cherry, "Bayesian reconstruction of PET images: Methodology and performance analysis," *Phys. Med. Biol.*, vol. 41, no. 9, pp. 1777–1807, Sep. 1996.

[18] J. Nuyts, D. Beque, P. Dupont, and L. Mortelmans, "A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 1, pp. 56–60, Feb. 2002.

[19] J. E. Bowsher, H. Yuan, L. W. Hedlund, T. G. Turkington, G. Akabani, A. Badea, W. C. Kurylo, C. T. Wheeler, G. P. Cofer, M. W. Dewhirst, and G. A. Johnson, "Utilizing MRI information to estimate F18-FDG distributions in rat flank tumors," in *Proc. IEEE Symp. Conf. Rec. Nucl. Sci.*, vol. 4, Oct. 2004, pp. 2488–2492.

[20] 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.

[21] M. J. Ehrhardt, P. Markiewicz, M. Liljeroth, A. Barnes, V. Kolehmainen, J. S. Duncan, L. Pizarro, D. Atkinson, B. F. Hutton, S. Ourselin, K. Thielemans, and S. R. Arridge, "PET reconstruction with an anatomical MRI prior using parallel level sets," *IEEE Trans. Med. Imag.*, vol. 35, no. 9, pp. 2189–2199, Sep. 2016.

[22] G. Wang and J. Qi, "PET image reconstruction using kernel method," *IEEE Trans. Med. Imag.*, vol. 34, no. 1, pp. 61–71, Jan. 2015.

[23] W. Hutchcroft, G. Wang, K. T. Chen, C. Catana, and J. Qi, "Anatomically-aided PET reconstruction using the kernel method," *Phys. Med. Biol.*, vol. 61, no. 18, pp. 6668–6683, Sep. 2016.

[24] D. Deidda, N. A. Karakatsanis, P. M. Robson, Y.-J. Tsai, N. Efthimiou, K. Thielemans, Z. A. Fayad, R. G. Aykroyd, and C. Tsoumpas, "Hybrid PET-MR list-mode kernelized expectation maximization reconstruction," *Inverse Problems*, vol. 35, no. 4, Apr. 2019, Art. no. 044001.

[25] G. Wang, "High temporal-resolution dynamic PET image reconstruction using a new spatiotemporal kernel method," *IEEE Trans. Med. Imag.*, vol. 38, no. 3, pp. 664–674, Mar. 2019.

[26] P. Novosad and A. J. Reader, "MR-guided dynamic PET reconstruction with the kernel method and spectral temporal basis functions," *Phys. Med. Biol.*, vol. 61, no. 12, pp. 4624–4644, Jun. 2016.

[27] K. Gong, J. Cheng-Liao, G. Wang, K. T. Chen, C. Catana, and J. Qi, "Direct patlak reconstruction from dynamic PET data using the kernel method with MRI information based on structural similarity," *IEEE Trans. Med. Imag.*, vol. 37, no. 4, pp. 955–965, Apr. 2018.

[28] G. E. Fakhri, A. Sitek, B. Guérin, M. F. Kijewski, M. F. D. Carli, and S. C. Moore, "Quantitative dynamic cardiac 82Rb PET using generalized factor and compartment analyses," *J. Nucl. Med.*, vol. 46, no. 8, pp. 1264–1271, 2005.

[29] R. Klein, R. S. Beanlands, R. W. Wassenaar, S. L. Thorn, M. Lamoureux, J. N. DaSilva, A. Adler, and R. A. deKemp, "Kinetic model-based factor analysis of dynamic sequences for 82-rubidium cardiac positron emission tomography," *Med. Phys.*, vol. 37, no. 8, pp. 3995–4010, Jul. 2010.

[30] Y. C. Cavalcanti, T. Oberlin, N. Dobigeon, C. Fevotte, S. Stute, M.-J. Ribeiro, and C. Tauber, "Factor analysis of dynamic PET images: Beyond Gaussian noise," *IEEE Trans. Med. Imag.*, vol. 38, no. 9, pp. 2231–2241, Sep. 2019.

[31] Y. Anzai, S. Minoshima, G. T. Wolf, and R. L. Wahl, "Head and neck cancer: Detection of recurrence with three-dimensional principal components analysis at dynamic FDG PET," *Radiology*, vol. 212, no. 1, pp. 285–290, Jul. 1999.

[32] M. H. M. Janssen, H. J. W. L. Aerts, M. C. Öllers, G. Bosmans, J. A. Lee, J. Buijsen, D. De Ruysscher, P. Lambin, G. Lammering, and A. L. A. J. Dekker, "Tumor delineation based on time–activity curve differences assessed with dynamic fluorodeoxyglucose positron emission tomography–computed tomography in rectal cancer patients," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 73, no. 2, pp. 456–465, Feb. 2009.

[33] R. Misiunaite, G. I. Angelis, and S. R. Meikle, "Clustering analysis for neurotransmitter response profiles of dynamic PET data," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. (NSS/MIC)*, Oct. 2017, pp. 1–4.

[34] L. Lu, N. A. Karakatsanis, J. Tang, W. Chen, and A. Rahmim, "3.5D dynamic PET image reconstruction incorporating kinetics-based clusters," *Phys. Med. Biol.*, vol. 57, no. 15, pp. 5035–5055, Aug. 2012.

[35] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[36] J. C. Bezdek, *Pattern Recognition With Fuzzy Objective Function Algorithms*. New York, NY, USA: Plenum Press, 1981.

[37] C. L. Chowdhary, M. Mittal, P. A. Pattanaik, and Z. Marszalek, "An efficient segmentation and classification system in medical images using intuitionist possibilistic fuzzy C-mean clustering and fuzzy SVM algorithm," *Sensors*, vol. 20, no. 14, pp. 3903–3922, Jul. 2020.

[38] J. Jiao, P. Markiewicz, N. Burgos, D. Atkinson, B. Hutton, S. Arridge, and S. Ourselin, "Detail-preserving PET reconstruction with sparse image representation and anatomical priors," in *Proc. Int. Conf. Inf. Process. Med. Imag.* Cham, Switzerland: Springer, 2015, pp. 540–551.

[39] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 463–468, Jun. 1998.
[40] L. Bi, J. Kim, A. Kumar, L. Wen, D. Feng, and M. Fulham, "Automatic detection and classification of regions of FDG uptake in whole-body PET-CT lymphoma studies," *Computerized Med. Imag. Graph.*, vol. 60, pp. 3–10, Sep. 2017.
[41] K. Gong, C. Catana, J. Qi, and Q. Li, "PET image reconstruction using deep image prior," *IEEE Trans. Med. Imag.*, vol. 38, no. 7, pp. 1655–1665, Jul. 2019.
[42] T. Yokota, K. Kawai, M. Sakata, Y. Kimura, and H. Hontani, "Dynamic PET image reconstruction using nonnegative matrix factorization incorporated with deep image prior," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 3126–3135.
[43] V. Lempitsky, A. Vedaldi, and D. Ulyanov, "Deep image prior," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 9446–9454.

SHUANGLIANG CAO received the B.Eng. degree from the Hebei University of Science and Technology, in 2012. He is currently pursuing the Ph.D. degree with the Guangdong Provincial Key Laboratory of Medial Image Processing, School of Biomedical Engineering, Southern Medical University. His research interests include dynamic PET imaging and parametric imaging.

YURU HE received the B.Eng. degree from Southern Medical University, in 2018, where she is currently pursuing the master's degree in engineering with the School of Biomedical Engineering. Her research interest includes PET/MRI denoising.

HONGYAN ZHANG received the B.Eng. degree from Southern Medical University, in 2018, where she is currently pursuing the master's degree with the Guangdong Provincial Key Laboratory of Medical Image Processing, School of Biomedical Engineering. Her research interest includes dynamic PET factor analysis.

WENBING LV received the Ph.D. degree in biomedical engineering from Southern Medical University, China, in 2020. She is currently a Postdoctoral Research Fellow with the Guangdong Provincial Key Laboratory of Medical Image Processing, School of Biomedical Engineering, Southern Medical University. Her research interests include PET image analysis and radiomics.

LIJUN LU received the Ph.D. degree in biomedical engineering from Southern Medical University, in 2012. He is currently an Associate Professor with the Guangdong Provincial Key Laboratory of Medial Image Processing, School of Biomedical Engineering, Southern Medical University. His research interests include PET imaging methods, medical image processing, and radiomics.

WUFAN CHEN received the B.S. and M.Sc. degrees from Beihang University, in 1975 and 1981, respectively. He is currently a Full Professor with the Guangdong Provincial Key Laboratory of Medial Image Processing, School of Biomedical Engineering, Southern Medical University. His research interests include biomedical imaging principle and image processing.