Improved Image Reconstruction Using Multi-Energy Information in Spectral Photon-Counting CT

PEI NIU[1], LIHUIWANG[2], BINGQING XIE[1], MARC ROBINI[1], LOIC BOUSSEL[1], PHILLIPPE DOUEK[1], YUEMIN ZHU[1], AND FENG YANG[3]

[1]CREATIS UMR 5220, CNRS, Inserm, INSA Lyon, Université de Lyon, F-69621 Lyon, France
[2]Key Laboratory of Intelligent Medical Image Analysis and Precise Diagnosis of Guizhou Province, College of Computer Science and Technology, Guizhou University, Guiyang 550025, China
[3]National Library of Medicine, National Institute of Health, Bethesda, MD 20894, USA

Corresponding author: Yuemin Zhu (yue-min.zhu@creatis.insa-lyon.fr)

ABSTRACT Spectral photon-counting computed tomography (sCT) appears as a promising imaging technique for clinical applications thanks to its ability to offer low dose and possibility of quantitatively analyzing the composition of materials in a pixel. However, due to the dispatching of photons into different energy bins, the quality of sCT image at each energy bin is considerably degraded. We propose a reconstruction method for sCT images by combining multi-energy information. The method is based on clustering pixels containing similar material compositions, performing linear fitting within each class for all the energy images two-by-two, projecting the pixel values of the images at other energy bins to the pixel of the image at the current energy bin, and combining the original pixel value and projected pixel values. The results on both simulation and real data demonstrated the effectiveness of the proposed method, in terms of both image reconstruction quality and material decomposition accuracy.

INDEX TERMS Image reconstruction, multi-energy information, data clustering, linear fitting, projection data, denoising.

I. INTRODUCTION

As a novel kind of computed tomography (CT), spectral photon-counting CT (sCT) generates usually three or more sets of data with one scanning owing to the use of photon-counting detectors, which have the ability to count separately the number of photons having different energies. Such energy discrimination ability enables sCT to produce not only at least different anatomical images referred to different energy bins but also functional images indicating the composition of materials in the object, including the category and concentration [1].

However, more energy bins also mean smaller photon numbers within one energy bin for a given energy range,

The associate editor coordinating the review of this manuscript and approving it for publication was Aamir Younis Raja[ID].

which will inevitably degrade image reconstruction quality. Therefore, various methods were developed for improving the reconstruction quality of sCT images. An intuitive manner to improve sCT images would be to use directly existing methods devoted to conventional CT images. In the projection domain, Hsieh developed an adaptive filtering according to the noise property to reduce artifacts [2]. La Riviere proposed a penalized likelihood smoothing model on the sinogram [3]. Wang *et al.* presented a penalized weighted least-squares approach to reduce sinogram noise [4]. However, because of the difficulty to accurately model the noise in CT system, some operations will bring in unexpected artifacts in the reconstruction, which could in certain cases be avoided in the image domain. In image domain, iterative methods are still the most widely used in the reconstruction [5], ranging from the original algebraic reconstruction method [6] to the

methods with empirical knowledge [7], [8]. More specifically for sCT reconstructions, Gao *et al.* proposed a compressive sensing approach that utilizes a prior rank, intensity and sparsity model (PRISM) for reconstruction [9]. Zhang *et al.* developed a reconstruction method by making use of the sparsity and similarity in sCT images [10]. Rigie and La Rivière proposed a constrained total nuclear variation method for joint reconstruction of multi-energy reconstruction [11]. Niu *et al.* utilized a total image constrained diffusion tensor for sCT reconstruction [12].

In fact, in different images corresponding to different energy bins of sCT, at the same pixel (position), the relationship between image pixel values is directly linked to material attenuation coefficients at these energy bins. This constitutes an important point exploited by Zhang *et al.* [10], who found the mapping relationship between attenuation coefficients at different energy bins and introduced a piecewise fitting method for regularizing their iterative reconstruction method. Their reconstruction method however works only for materials without K-edge. It is not suitable for materials with K-edge while, in medical applications, contrast agents such as Gd and I often present K-edges within the photon energy range in sCT.

In the present study, we propose a new sCT reconstruction method that allows dealing with materials with K-edge. The method is based on clustering pixels containing similar material compositions, performing linear fitting within each class for all the energy images two-by-two, projecting the pixel values of the images at other energy bins to the pixel of the image at the current energy bin, and combining (median filtering) the original pixel value and projected pixel values. The rest of the paper is organized as follows. In section II, the flowchart and details of the proposed method are described. In section III, experiments and results on both simulation and real data are presented, followed by sections IV and V respectively for discussion and conclusions.

II. METHODS

A. THEORY BASIS OF THE METHOD

The traditional representation of the attenuation coefficient is either effect-based or material-based one [18], [19]. Here, considering the k-edge property of the material, the attenuation coefficient at a pixel (i.e. image pixel value) is represented as the particular mixture of the two effects [20]–[22],

$$\mu(\vec{x}, E) = f_p(\vec{x}) g_p(E) + f_c(\vec{x}) g_c(E) + \sum_{\alpha=1}^{N} \mu_{m\alpha}(E) \rho_\alpha(\vec{x}),$$
(1)

where $g_p(E)$ and $g_c(E)$ are respectively photoelectric absorption and Compton effects, $f_p(\vec{x})$ and $f_c(\vec{x})$ are the constants that are determined by the materials in pixel $\vec{x}$, N is the number of K-edge materials in pixel $\vec{x}$, $\mu_{m\alpha}(E)$ is the mass attenuation coefficient of K-edge material α at

energy E, and $\rho_\alpha(\vec{x})$ is the corresponding density. Eq. (1) can be represented by two parts

$$\mu(\vec{x}, E) = \mu_{No}(\vec{x}, E) + \mu_{K-edge}(\vec{x}, E),$$
(2)

where $\mu_{K-edge}(\vec{x}, E)$ denotes the K-edge effect, and $\mu_{No}(\vec{x}, E)$ the photoelectric absorption and Compton effects. By representing a material as a linear combination of two base materials (water and bone) [10], it follows

$$\mu_{No}^\zeta(\vec{x}, E) = a_1 \mu_{No}^{Water}(\vec{x}, E) + a_2 \mu_{No}^{Bone}(\vec{x}, E),$$
(3)

where a_1 and a_2 are material-dependent constants. As a result, Eq. (4) can be written as

$$\mu(\vec{x}, E) = a_1 \mu_{No}^{Water}(\vec{x}, E) + a_2 \mu_{No}^{Bone}(\vec{x}, E)$$
$$+ \sum_{\alpha=1}^{N} \mu_{m\alpha}(E) \rho_\alpha(\vec{x}).$$
(4)

In the work of Zhang *et al.*, materials without K-edge in humans are classified into water-like and bone-like materials. That is to say, for water-like material, $a_1 \approx 1, a_2 \approx 0$; for bone-like material, $a_1 \approx 0, a_2 \approx 1$. Thus, considering the pixels with K-edge materials, Eq.(4) can further be expressed as

$$\mu(\vec{x}, E) = \mu_{No}^{Water}(\vec{x}, E) + \sum_{\alpha=1}^{N} \mu_{m\alpha}(E) \rho_\alpha(\vec{x}),$$
(5)

or

$$\mu(\vec{x}, E) = \mu_{No}^{Bone}(\vec{x}, E) + \sum_{\alpha=1}^{N} \mu_{m\alpha}(E) \rho_\alpha(\vec{x}),$$
(6)

For a given pixel, $\vec{x}$ is known and Eq. (5) becomes

$$\mu(E) = \mu_{No}^{Water}(E) + \sum_{\alpha=1}^{N} \mu_{m\alpha}(E) \rho_\alpha.$$
(7)

Thus, if a pixel only contains one material with K-edge, which is the most common case in medical diagnosis, the attenuation coefficient is actually the linear function of concentration within each energy bin. For the same pixel at all the energy bins, as the concentration is fixed and the attenuation of water changes little (which can thus be regarded as a constant), μ is also a linear function of the concentration if there is only one material in the pixel.

The flowchart of our method is illustrated in Fig. 1. First, primary images at different energy bins are reconstructed using basic reconstruction methods (FBP for simulation and SART for physical CT in the present study). The next step is to find the image pixels having similar features (attenuation coefficient values and their variations at each energy bin) and label them as the same class by clustering in the 3D patches. Here, the common k-means algorithm is utilized. Note that the clustering of patches carries both spatial and energy information, and that the pixels at the same position for each energy bin are labelled as the same class. Finally,
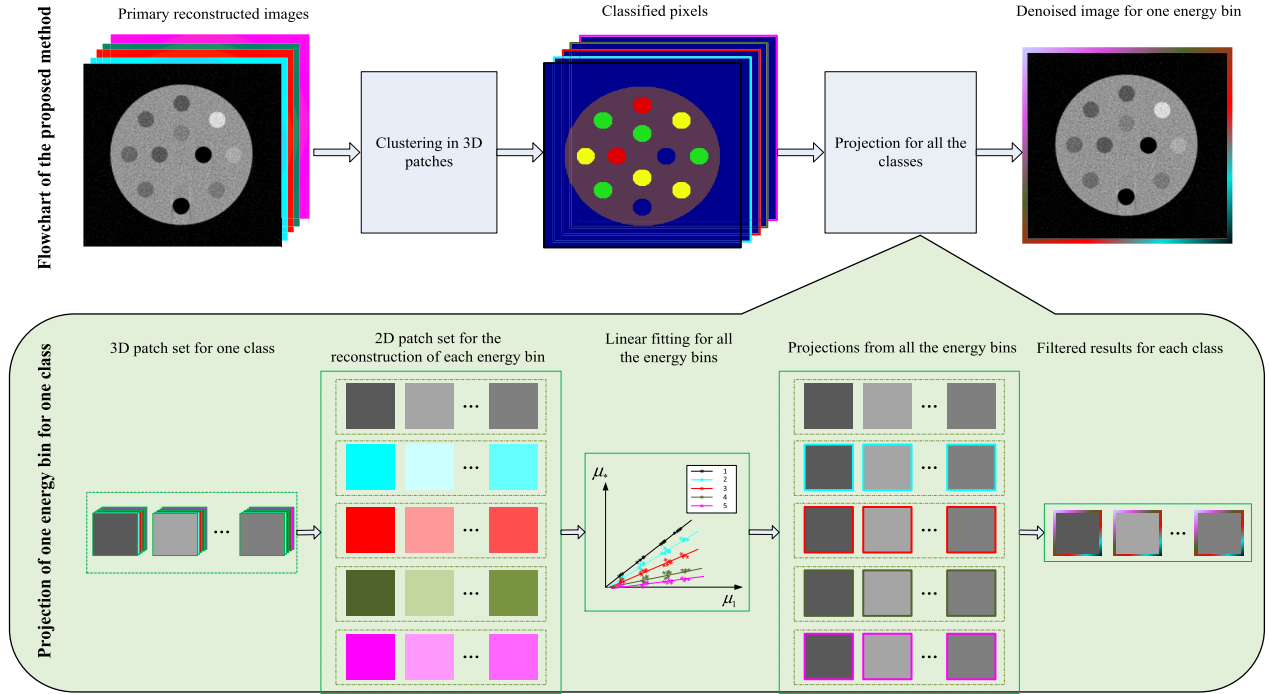
FIGURE 1. Flowchart of the proposed method for one energy-bin projection.

the pixels in one energy bin for all the classes are projected into all the other energy bins and form the final reconstructed images. The projection procedure is implemented by the following steps. The patches (corresponding to the center pixels) in the same class are arranged according to the energy bin. Then, linear fitting is made for all the patches labelled as the same class in any two energy bins separately, thus yielding the fitting coefficients for any two energy bins for each class. By using such image value fitting two-by-two, an image pixel value at one energy bin can be estimated from the image pixel values at all the other energy bins. That is, for a sCT system producing five images corresponding to five energy bins, a pixel will have five attenuation coefficients (i.e. pixel values): the current (original) image value and the other four image values mapped (estimated) from the other four energy bins. A filter is applied to the five values to obtain a single value that presents reduced noise, thus improving the quality of the reconstructed image. In the next paragraphs, the main elements of the framework will be presented in detail.

B. CLUSTERING FOR 3D PATCHES

As linear relationship exists between pixels having similar composition, different material compositions obey different fitting. The first step then is to classify the pixels of the reconstructed images. The classification is operated on the 3D image, which is formed by stacking the 2D images at different energy bins. In view of the influence of noise and spatial correlation between pixels, we perform the classification patch by patch. To this end, 2D patches in primary reconstructed images at all bins are arranged to form a set of 3D patches, the third dimension of which represents energy dimension.

Denoting the primary reconstruction image as

$$U = [\mu(1); \mu(2); ...; \mu(m); ...\mu(M)], \quad (8)$$

where M is the total number of energy bins, and $\mu(m)$ is the reconstructed image for the mth energy bin, which is shaped as

$$\mu(m) = \begin{bmatrix} \mu_{11}(m) & \cdots & \mu_{1N_I}(m) \\ \vdots & \ddots & \vdots \\ \mu_{N_I 1}(m) & \cdots & \mu_{N_I N_I}(m) \end{bmatrix}, \quad m = 1, 2, \ldots, M, \quad (9)$$

where $N_I \times N_I$ is the image size. Then the cluster unit X is represented by

$$X_{ij} = \left[x_{ij}(1); x_{ij}(2); \ldots; x_{ij}(M); \right], \quad (10)$$

$$x_{ij}(m) = \begin{bmatrix} \mu_{\left(i-\frac{P+1}{2}\right)\left(j-\frac{P+1}{2}\right)}(m) & \cdots & \mu_{\left(i-\frac{P+1}{2}\right)\left(j+\frac{P-1}{2}\right)}(m) \\ \vdots & \mu_{ij}(k) & \vdots \\ \mu_{\left(i+\frac{P-1}{2}\right)\left(j-\frac{P+1}{2}\right)}(m) & \cdots & \mu_{\left(i+\frac{P-1}{2}\right)\left(j+\frac{P-1}{2}\right)}(m) \end{bmatrix},$$

$$m = 1, 2, \ldots, M, \mu_{ij}(m) \in \mu(m) \quad (11)$$

where $x_{ij}(m)$ is the 2D patch at energy bin m centered at the position (i, j), $\mu_{ij}(m)$ is the primary reconstruction at the position (i, j) for the mth energy bin, X_{ij} denotes the 3D patch to be clustered centered at the position (i, j), $P \times P$ is the 2D patch size for one energy bin, of which P is usually selected as an odd number. Thus, each 3D patch is of size $P \times P \times M$.

The key point will be to extract the features of 3D patches for clustering. To better bring out the feature of each material, the attenuation coefficient of water is first removed from the

spatial image (at each energy bin) by subtracting from the latter the pure water attenuation coefficient. The processed cluster is then expressed as

$$x'_{ij}(m) = x_{ij}(m) - x_{\text{water}}(m), \quad m = 1, 2, \ldots, M, \quad (12)$$

$$X'_{ij} = \left[x'_{ij}(1); x'_{ij}(2); \ldots; x'_{ij}(M); \right] \quad (13)$$

For different patches, the parameters such as mean, first-order differential and second-order differential of attenuation coefficients and its mean are combined into a vector as the feature for classification.

$$V_{ij} = \left[\overline{X'_{ij}}; \nabla X'_{ij}; \overline{\nabla X'_{ij}}; \nabla^2 X'_{ij}; \overline{\nabla^2 X'_{ij}} \right], \quad (14)$$

where $\overline{X'_{ij}}$ is the mean of X'_{ij}, $\nabla X'_{ij}$ is the first-order differential of X'_{ij} in the 3rd dimension (namely the energy bin dimension), $\overline{\nabla X'_{ij}}$ is the mean of $\nabla X'_{ij}$, $\nabla^2 X'_{ij}$ is the second-order differential of X'_{ij} in the 3rd dimension, and $\overline{\nabla^2 X'_{ij}}$ is the mean of $\nabla^2 X'_{ij}$.

K-means algorithm is then chosen to cluster the features. The number of classes is the most important parameter in k-means clustering. If K is too small, pixels containing different materials will be classified into the same class, which will result in missing classes and reduce fitting accuracy within the class. If K is too large, pixels containing the same material will be classified into different classes and the number of pixels in one class will be small, which will negatively impact fitting results. K is material-dependent; ideally, the number of classes should be set as the same number of materials. However, in practice, accounting for the influences of material edges and noise in the image, we heuristically set the number of classes as twice the number of materials.

Meanwhile, k-means algorithm is sensitive to the selection of initial points. To cope with this problem, in the clustering process, several repeated k-means algorithms with random initials are applied on the 3D patches. If the patches in two continuous operations are clustered to the same class, they will be regarded as reliable patches and the class label is kept for further fitting. Otherwise, they will be regarded as unreliable patches and will not be considered in the next fitting step.

C. LINEAR FITTING BETWEEN BINS

For the patches in the same class, fitting is made between each two energy bins. Let $\{V_{k1}, V_{k2}, \ldots, V_{kN_m}\}$ be the patch set for class k, where N_k is the total number of patches for class k, $k = 1, 2, \ldots, K$, and K is the total number of classes. Considering that a point is often polluted by noise, if fitting is made point by point, the fitting result will likely be affected by the noise. Therefore, the $P \times P$ points in the 2D patch at each bin are utilized for fitting. Meanwhile, to preserve the edges of the image, the weighted mean of the points at each bin is employed, of which the weight is defined as

$$W = \frac{1}{P^D}, \quad (15)$$

where D is the City distance from the neighbor position to the central pixel within the 2D patch. In our method, least square error method (LSE) is adopted to perform the fitting between every two bins. Thus, for each class, $M \times M$ group fitting coefficients will be generated, which will be utilized for further mapping.

The fitting results from bin p to bin q is calculated by

$$\mu^p_{ij}(q) = a_{pq}(k)\, \bar{\mu}_{ij}(p) + b_{pq}(k),$$
$$k = 1, 2, \ldots, K, p, q \in 1, 2, \ldots, M, \ if \ \mu_{ij} \in k, \quad (16)$$

where $\mu^p_{ij}(q)$ denotes the fitted attenuation coefficient from bin p to bin q, $\bar{\mu}_{ij}(p)$ is the weighted average within the 2D patch for the energy bin p, $a_{pq}(k)$ and $b_{pq}(k)$ designate the fitting coefficients for the kth class. Thus, after traversing all the patches in the K class, for all the positions in each energy bin, $M - 1$ fitting results will be obtained from the other bins. If counting the bin itself, M reconstructions will be utilized for the next operation (post-denoising).

D. REFINEMENT OF FITTED IMAGES

After the fittings, we obtain M reconstructed images corresponding to the M energy bins. The next step is to improve the current reconstructed image by making use of the fitting results (i.e. the $M - 1$ fitted images generated from the fittings). The first idea would be to average the M images to reduce noise that is the principal problem in sCT. To this end, weighted averaging is adopted because it has the advantage of producing stable denoising results. However, considering the characteristic of the data acquired from each energy bin, targeted filter is selected to get better denoising performance. If a point in the reconstructed image at one energy bin suffers from noise, the fitting results (i.e. fitted images) from this bin to all the remaining bins will also be noised because the fitting involves all the energy bins. Therefore, in the present study, we choose median filter to achieve the averaging operation, which takes the current reconstructed image and all the fitted images into account. It is expressed as

$$\tilde{\mu}_{ij}(q) = Med\left\{ \mu^1_{ij}(q), \mu^2_{ij}(q), \ldots, \mu^K_{ij}(q) \right\},$$
$$q = 1, 2, \ldots, M \quad (17)$$

Actually, with respect to the original reconstructed image, the fitted images are relatively far away from their respective original image. This can be caused by the mixture (which has complex material compositions) or wrong clustering. Therefore, to make the filtered images more reliable, we measure the relative error between the fitted values and original values. If the error is larger than a threshold, the fitted values are abandoned and the corresponding pixels keep the original values, namely

$$\tilde{\mu}'_{ij}(q) = \begin{cases} \tilde{\mu}_{ij}(q), & if \ \left|\tilde{\mu}_{ij}(q) - \mu_{ij}(q)\right| / \left|\mu_{ij}(q)\right| \\ & < threshold \\ \mu_{ij}(q), & if \ \left|\tilde{\mu}_{ij}(q) - \mu_{ij}(q)\right| / \left|\mu_{ij}(q)\right| \\ & \geq threshold. \end{cases} \quad (18)$$

where $\tilde{\mu}'_{ij}(q)$ is the final estimation of the attenuation coefficient for bin q at position (i, j). After the operation in Eq. (18) we obtain the final refined image.

III. EXPERIMENTAL AND RESULTS

A. SIMULATION

1) PHANTOM AND EVALUATION METRICS

The projection data of sCT was simulated using INSA software Virtual X-ray Imaging (VXI) [16]. The sCT system consists of a PCD with 700 pixels arranged in line. The spectrum was simulated based on Birch & Marshall model for tungsten target material without filtration. The target angle was 17°, with tube voltage of 100kVp, tube current 100 μA and spectrum resolution 1 keV. The detector has energy-resolving ability of 5 energy bins. The energy bin ranges are 30-39, 40-49, 50-59, 60-69 and 70-79 keV. In our experiments, the phantom was scanned with 600 projection views from 0° to 360° of equal interval, and the reconstructed image size is 380×380. The 3D patch size was set as $3 \times 3 \times 5$, and the threshold as 0.3.

The simulation phantom was designed as illustrated in Fig, 2(a). It is a big cylinder made up of PMMA of radius 130 mm, which was digged out 12 holes to fill up water, gadolinium (Gd) and iodine (I). For Gd, the concentrations were separately set as 2, 4, 6, 8 mg/cc, and for I as 1, 3.5, 7 and 14 mg/cc. The 12 holes are regarded as the regions of interest (ROIs), and the evaluations are made on the ROIs.

To evaluate the performance of the reconstruction, root mean square errors (RMSE), contrast-to-noise ratio (CNR) and signal-to-noise ratio (SNR) of the 12 ROIs are utilized, which are respectively defined as

$$RMSE\,(m)$$
$$= \sqrt{\sum_{\vec{x}} \left(\mu_{\vec{x}}(m) - \mu_{\vec{x}}^{GT}(m) \right)^2 \Big/ N_I^2},$$
$$m = 1, 2, \ldots, M, \tag{19}$$

$$CNR\,(m)$$
$$= \frac{|\overline{\mu}_{ROI}(m) - \overline{\mu}_{BG}(m)|}{\sqrt{\sigma_{ROI}^2(m) + \sigma_{BG}^2(m)}}, \quad m = 1, 2, \ldots, M, \tag{20}$$

$$SNR\,(m)$$
$$= 10 \log \left(\sum_{\vec{x}} \left\| \mu_{\vec{x}}^{GT}(m) \right\|^2 \Big/ \sum_{\vec{x}} \left\| \mu_{\vec{x}}(m) - \mu_{\vec{x}}^{GT}(m) \right\|^2 \right). \tag{21}$$

where $\mu_{\vec{x}}(m)$ and $\mu_{\vec{x}}^{GT}(m)$ are respectively the reconstructed and ground-truth at position $\vec{x}$ and energy bin m, $\overline{\mu}_{BG}(m)$ (or $\overline{\mu}_{ROI}(m)$) represents the mean attenuation coefficient of background (or ROI), and $\sigma_{BG}(m)$ (or $\sigma_{ROI}(m)$) the standard deviation of the latter.

The material decomposition results were obtained using the L1-norm method [25]. To assess the decomposition results, the obtained material images (a material image is a
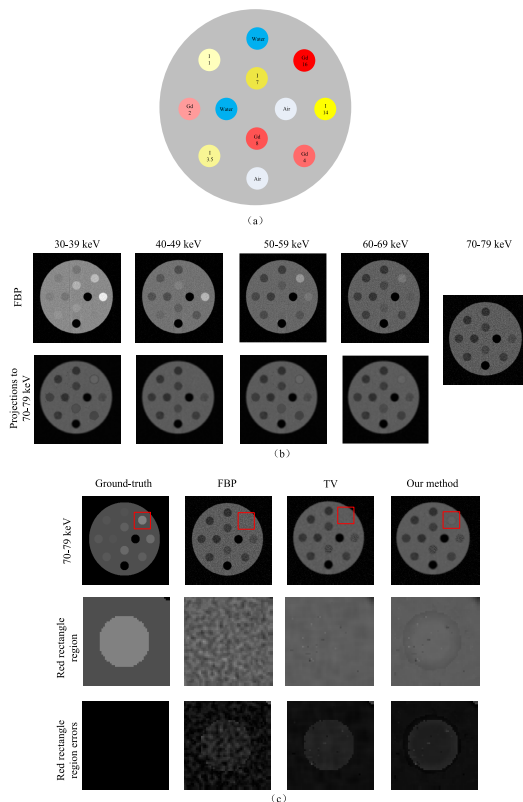

FIGURE 2. Results of reconstruction at energy bin 70-79 keV. (a) The phantom; (b) The first row shows the primary reconstruciton by FPB for each energy bin, and the second row shows the projection for the reconstruction of energy bin 70-79 keV from each other energy bin; (c) the comparison of reconstruction for the energy bin 70-70 keV.

spatial image that contains only one single material) are first visually compared with the ground-truth. Then, the values within the ROIs are compared with the ground-truth and the standard deviations.

2) RECONSTRUCTION COMPARISON

Fig. 2(b) illustrates the original reconstruction by FBP for five energy bins and the fitting results from the four bins to the energy bin 70-79 keV. Fig. 2(c) shows the final reconstruction from Fig. 2(b) at energy bin 70-79 keV, together with the ground-truth and reconstructions from FBP and TV. It can be observed that the results of our method are the closest to ground-truth.

More quantitatively, TABLE 1 gives RMSEs and SNRs of the original image and our fitted image for each energy bin. Clearly, our method has the smallest RMSE and the largest SNR. The TV method gives large RMSE and small SNR. With respect to FBP, the improvement in RMSE of our method for different energy bins is respectively 1.53%, 6.00%, 4.50%, 9.07%, 14.36%, and the improvement in SNR is respectively 1.01%, 4.38%, 3.48%, 7.18%, 12.34%. In other words, the proposed method improves RMSE and SNR for all the energy bins. Concerning CNR, our method shows better performance compared to FBP, but is not as good as TV. This is because the materials within the phantom

TABLE 1. Comparison of results on RMSEs, SNRs and CNRs.

		30-39	40-49	50-59	60-69	70-79
RMSE	FBP	0.046	0.042	0.040	0.040	0.041
	TV	0.082	0.068	0.062	0.057	0.054
	Our method	0.045	0.039	0.038	0.036	0.035
	Improvement wrt FBP	2.17%	7.14%	5.00%	10.00%	14.63%
SNR	FBP	13.685	12.644	12.073	11.365	10.389
	TV	6.149	5.985	5.8341	5.852	5.866
	Our method	13.823	13.198	12.493	12.181	11.67
	Improvement wrt FBP	1.01%	4.38%	3.48%	7.18%	12.34%
CNR	FBP	1.194	1.259	1.332	1.257	1.216
	TV	1.255	1.400	1.491	1.465	1.500
	Our method	1.227	1.298	1.352	1.346	1.349
	Improvement wrt FBP	2.76%	3.04%	1.51%	7.03%	10.94%

are averagely distributed in the ROIs; TV has better ability to smooth the region, but at the same time may alter actual values of reconstructed image and consequently impact the subsequent material decomposition.

3) DECOMPOSITION COMPARISON

As material decomposition is one of the most important applications of sCT, reconstructions are compared in terms of decomposition results using L1-norm decomposition method. The reconstructed images are decomposed into Water, PMMA, I and Gd, as shown in Fig. 3. It is observed that the decomposition results using our method contain less noise and are closer to the ground-truth.
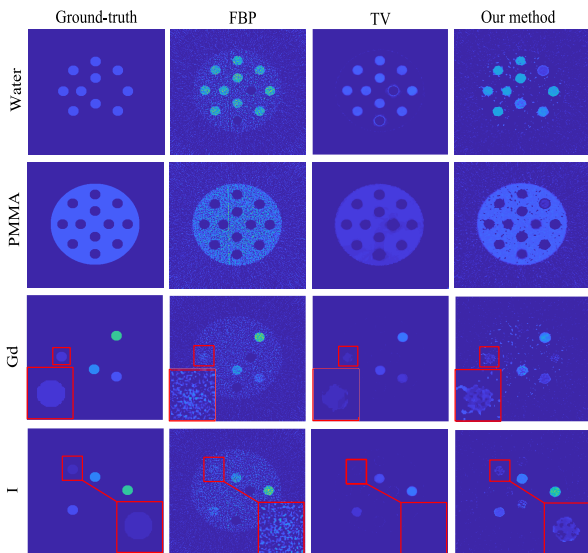

FIGURE 3. Results of material decomposition based on the images reconstructed using FBP, TV and the proposed method.

Furthermore, when comparing the fitting lines plotted in Fig. 4, we observe that although the results of our method were not as smooth as those of TV, the fitting line of our method is closer to the ground-truth compared to the other methods.
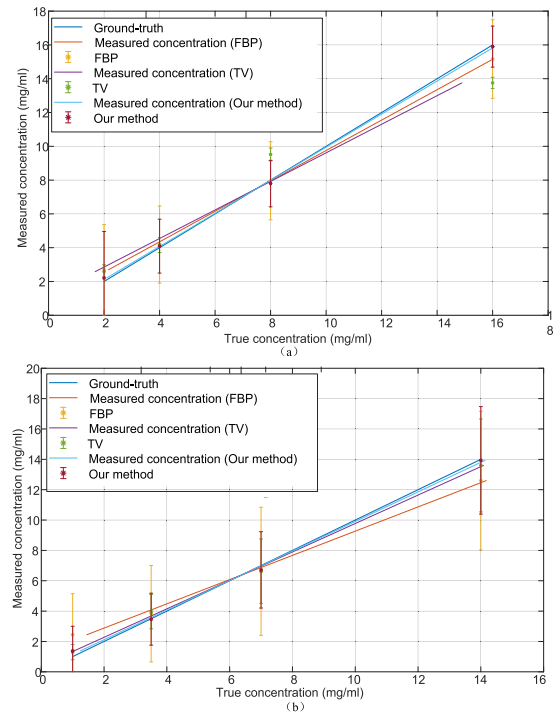

FIGURE 4. Decomposition results evaluation of FBP, TV and our method. (a) Gd; (b) I.

B. REAL DATA
1) ON PHYSICAL PHANTOM

Two physical phantoms (Fig. 5) respectively filled with Gd (or I) and water were utilized to evaluate the performance of the proposed method. The phantom is made of PMMA with the diameter of 100mm. The concentration of Gd was set from 0.1 mg/cc to 15 mg/cc, and that of I from 0.1 mg/cc to 15 mg/cc, as marked by the number in each disk. The phantom was scanned by Philips sCT prototype with the source current 220 mA, source voltage 120 keV and projection views 2400 uniformly distributed over 360°. The sCT system contains 924 detectors, each of which contains 5 energy bins, with thresholds set as 30, 51, 62, 72 and 81 keV. The performance is evaluated in terms of both image reconstruction and material decomposition.
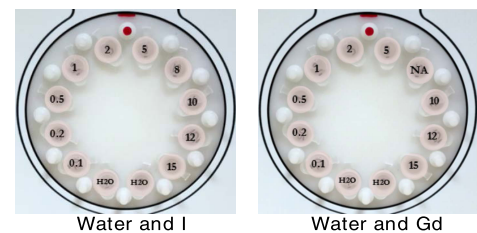

FIGURE 5. Physical phantoms.

Fig. 6 shows the reconstruction results for each energy bin using SART, TV and our method. The SART and TV methods were implemented using RTK toolkit [17]. It is observed that after utilizing information from other energy bins, the artifacts of the reconstruction for each energy bin are visually reduced.

TABLE 2. Reconstruction comparison in terms of CNR.

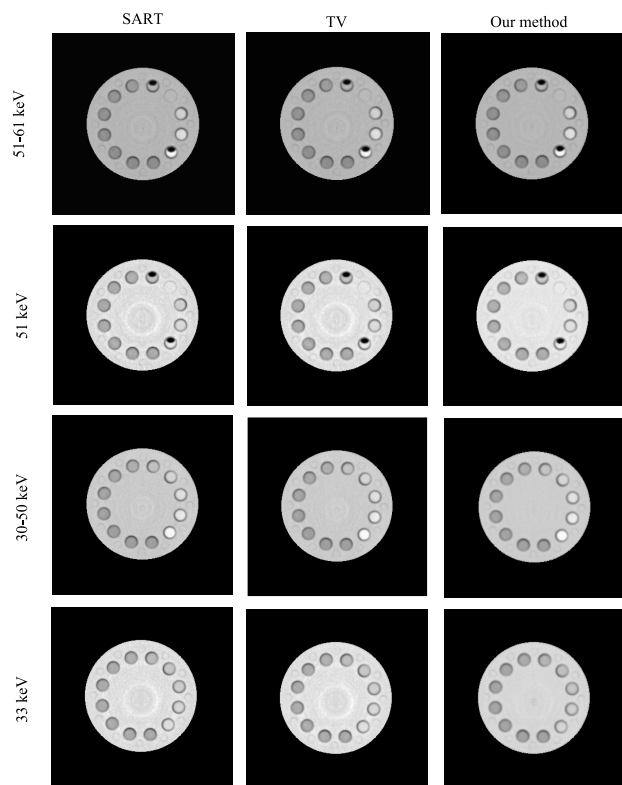|		30-50	51-61	62-71	72-80	81-150
Gd	SART	1.088	1.273	1.128	0.973	0.912
	TV	1.118	1.307	1.157	1.003	0.919
	Our method	1.125	1.303	1.167	1.017	0.956
	Improvement (SART)	3.40%	2.36%	3.46%	4.52%	4.82%
I	SART	1.500	1.370	1.229	1.223	1.307
	TV	1.548	1.387	1.282	1.293	1.374
	Our method	1.551	1.432	1.283	1.296	1.422
	Improvement (SART)	3.40%	4.52%	4.39%	5.97%	8.80%

FIGURE 6. Reconstruction results for different energy bins with K-edge and monoE reconstruction of the K-edge. Rows 1 and 2: reconstruction for Gd. Rows 3 and 4: reconstruction for I.

FIGURE 7. Decomposition results for physical phantom with Gd and I, respectively. The right column for each image is the region with low concentrations (yellow and green rectangle) and of special shape (the red rectangle) in the blown-up form. The images for the same region for all the methods are displayed in the same scale.

FIGURE 8. Decomposition value comparison: the axis is the logarithm of the concentration; each star represents the mean value within the ROI and the line segment through each star is the responding absolute standard error. (a) Gd; (b) I.

Generally, there lacks the ground-truth for real data (except physical phantoms). To maintain the consistency in the evaluation of the real data, the CNRs of the ROIs are calculated to perform comparison, as given in TABLE 2. Our method clearly improves the CNR with respect to SART and shows most of the time better performance than TV.

Fig. 7 displays the ground-truth of the physical phantom and the material decomposition results derived from the images reconstructed using SART, TV and proposed methods. The ROIs with low concentrations (yellow and green rectangles) and special shape (red rectangle) are also zoomed up. We observe that our method is more effective for the ROIs both with low concentrations and of special shape. It distinguishes more pixels of low concentration and the ROI regions are more fully filled. For other non-ROI regions, our method also reduces substantially artifacts.
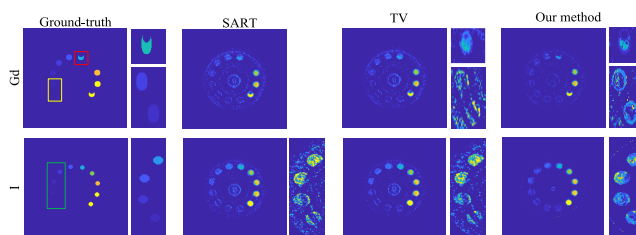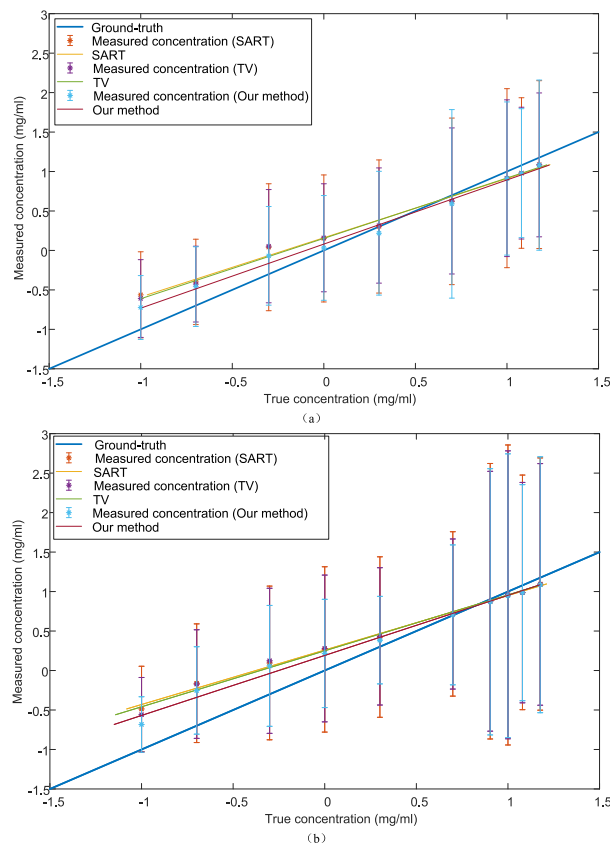
More quantitatively, material concentration and standard error for each ROI are plotted in Fig. 8. The axes are graduated as the logarithm of concentration to better visualize the difference, as the concentration of materials in the phantom has mainly small values. The ground-truth is the line passing through the origin of coordinate system. As observed, the results of our methods are most of the time the nearest to the ground-truth and have the smallest standard errors.

2) ON RABBIT

The rabbit injected with Gd was fixed in the container scanned by the physical sCT. Four tubes filled with Gd with concentrations of 10, 5, 2 and 0 mg/cc (red rectangle
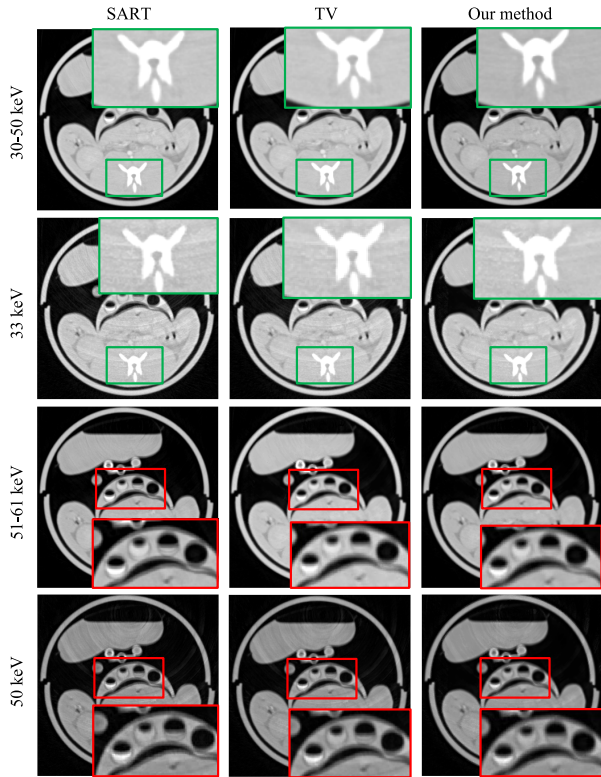
FIGURE 9. Reconstruction on rabbit data at energy bin with K-edge (rows 1 and 3) and monoE reconstruction at K-edge (rows 2 and 4).

FIGURE 10. Decomposition results of the rabbit injected by Gd. The tubes in the red rectangle are filled with Gd as reference, with concentration of 10, 5, 2 and 0 mg/cc from left to right, respectively. The yellow and green rectangles are organs detected on the rabbit. Globally, our method reduces some artifacts compared with the other two methods.

in Fig. 10) were fixed in the container and scanned together with the rabbit to act as a reference. The system parameters of sCT were set the same as those for the physical phantom described before. Fig. 9 shows the reconstructions from the sinograms of 30-50 keV that contains the K-edge energy of I and those of 51-61 keV that contains the K-edge energy of Gd. For the reconstruction at one energy bin, it is not easy to observe the difference since the width of the energy bin is too large. In contrast, in the monoE reconstruction, which is reconstructed from the energy of K-edge, it is observed that the reconstructions with our method visibly reduce artifacts compared to the other two methods.

Now, we decompose the images into water, Gd and I. Note that water-like material (such as plastic) will be decomposed into water, and that I-like material will be decomposed into I (such as the bone whose main material is Ca). It is observed that our method suffers fewer artifacts compared to the other two methods, especially for the material I (second row).

IV. DISCUSSION

We have proposed an improved reconstruction method for sCT images by combining multi-energy information. Based on the similarity of the object composition, material-like pixels are clustered into the same class for fitting. The fitted images at other energy bins are then filtered (median filtering) to obtain the final reconstructed image at the current energy bin, which enables the reconstruction and decomposition more accurate.
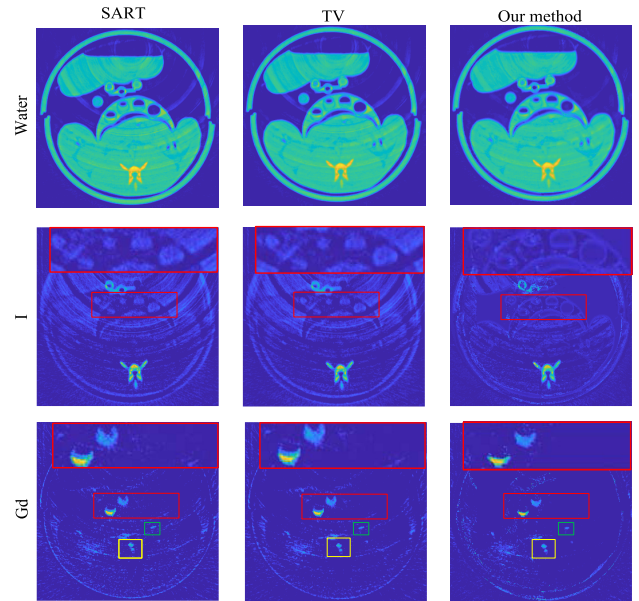
The results on both simulation and real data showed that our method works well not only in terms of reconstruction quality but also in terms of material decomposition. To quantitatively validate our method, we utilize materials with accurate concentrations for simulation. In the simulation, our method operated on the primary results obtained using FBP. The reconstruction results are not as smooth as the TV method. This is due to the fact that our method does not uniformly smooth neighboring pixels. Nevertheless, the reconstruction results of our method are closer to the ground-truth because it revises pixel values at each energy bin through fitting mechanism. This improves reconstruction and material decomposition on real data that contains more complex noise and artifacts. That explains why our method reduces substantially artifacts both in the case of physical phantom and in the case of rabbit.

We treated water as a constant and subtracted it from the raw attenuation value. That preserves the linearity when projecting values from one energy bin into another, but may affect the quantification of multiple soft tissue materials like fat and water. Further improvements would be necessary to be able to describe the features of all the materials.

The performance of our method is affected by the clustering of pixels during reconstruction. As the linear relationship of attenuation coefficients at each energy bin exists between the pixels having similar material composition, correct classification of pixels is an important element of the proposed method. For this reason, feature selection for each pixel is an interesting issue that can be further studied.

The fitting process is another important aspect of the proposed method. Due to the difficulty to model actual noise in

reconstruction, median filter was adopted to obtain the final pixel value from all the fitted values corresponding to all the energy bins. However, other type of filters could also be envisaged.

We did not compare our method with reconstructions directly from projections, because of the problems of resource and computation time. Indeed, for real spectral CT system, it is not always easy to get and correctly decode raw projection data. In contrast, for the simulations, we generated projection data and analyzed it. To reconstruct an image of size 780*780, for 3600 projections, the system matrix is about 10 Gb, which requires large memory and computation power. In the future, if conditions permit (parsed data from physical machine, available resources of computation and memory), it would be very interesting to make comparison with reconstructions from projection data, such as PRISM. We expect to further improve our method with respect to PRISM.

Beam hardening was not accounted for in the present study. It might however be present in large scans, perhaps even with narrow energy bins. In our future work, we can correct beam hardening for each energy bins before applying the proposed reconstruction method. It would also be interesting to assess the effects of beam hardening correction for different widths of energy bins.

The proposed reconstruction method could also been seen as a denoising method like those proposed in [23] and [24], in the sense that it operates in the spatial image domain. Note however that the manner in which the denoising involved in the present study is very different from traditional denoising methods. This is because spatial images at different energy bins were not processed independently; instead, intrinsic correlation between them exploited.

V. CONCLUSION

We have proposed a reconstruction method for images in sCT by exploiting multi-energy information. At each energy bin, we reconstruct the current image through using the images from the other energy bins. To achieve this, we first classify similar pixels using the k-means algorithm. Then, we make linear fitting between two images corresponding to two energy bins within each class. These linear fittings are made two-by-two for all the images corresponding to all the energy bins, which leads, for a current (original) image at a given energy bin, to several other (estimated) images corresponding to the other energy bins. At the end, for each given pixel, a filtering (median) in the energy direction generates the final reconstructed image that presents better quality. The results on both simulated and real data demonstrated the effectiveness of the proposed method, in terms of both reconstruction quality and material decomposition.

REFERENCES

[1] C. O. Schirra, B. Brendel, M. A. Anastasio, and E. Roessl, "Spectral CT: A technology primer for contrast agent development," *Contrast Media Mol. Imag.*, vol. 9, no. 1, pp. 62–70, Jan. 2014.

[2] J. Hsieh, "Adaptive streak artifact reduction in computed tomography resulting from excessive X-ray photon noise," *Med. Phys.*, vol. 25, no. 11, pp. 2139–2147, Nov. 1998.

[3] P. J. La Rivière, "Penalized-likelihood sinogram smoothing for low-dose CT," *Med. Phys.*, vol. 32, no. 6, pp. 1676–1683, May 2005.

[4] S. Tang and X. Tang, "Statistical CT noise reduction with multiscale decomposition and penalized weighted least squares in the projection domain," *Med. Phys.*, vol. 39, no. 9, pp. 5498–5512, Aug. 2012.

[5] G. Wang, H. Yu, and B. De Man, "An outlook on X-ray CT research and development," *Med. Phys.*, vol. 35, no. 3, pp. 1051–1064, Feb. 2008.

[6] R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography," *J. Theor. Biol.*, vol. 29, no. 3, pp. 471–481, Dec. 1970.

[7] K. Mechlem, S. Ehn, T. Sellerer, E. Braig, D. Munzel, F. Pfeiffer, and P. B. Noel, "Joint statistical iterative material image reconstruction for spectral computed tomography using a semi-empirical forward model," *IEEE Trans. Med. Imag.*, vol. 37, no. 1, pp. 68–80, Jan. 2018.

[8] M. Beister, D. Kolditz, and W. A. Kalender, "Iterative reconstruction methods in X-ray CT," *Phys. Medica*, vol. 28, no. 2, pp. 94–108, Apr. 2012.

[9] H. Gao, H. Yu, S. Osher, and G. Wang, "Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM)," *Inverse Problems*, vol. 27, no. 11, pp. 1–30, 2011.

[10] Y. Zhang, Y. Xi, Q. Yang, W. Cong, J. Zhou, and G. Wang, "Spectral CT reconstruction with image sparsity and spectral mean," *IEEE Trans. Comput. Imag.*, vol. 2, no. 4, pp. 510–523, Dec. 2016.

[11] D. S. Rigie and P. J. La Rivière, "Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization," *Phys. Med. Biol.*, vol. 60, no. 5, pp. 1741–1762, Mar. 2015.

[12] S. Niu, Z. Bian, D. Zeng, G. Yu, J. Ma, and J. Wang, "Total image constrained diffusion tensor for spectral computed tomography reconstruction," *Appl. Math. Model.*, vol. 68, pp. 487–508, Apr. 2019.

[13] T. H. Lee and B. C. Song, "De-noising algorithm using sparse 3D transform-domain collaborative filtering and adaptive soft thresholding," in *Proc. Int. Symp. Consum. Electron. (ISCE)*, 2011, vol. 16, no. 8, pp. 128–131.

[14] 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, Jan. 2005.

[15] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Math.*, vol. 76, no. 2, pp. 167–188, Apr. 1997.

[16] P. Duvauchelle, N. Freud, V. Kaftandjian, and D. Babot, "A computer code to simulate X-ray imaging techniques," *Nucl. Instrum. Methods Phys. Res. B, Beam Interact. Mater. At.*, vol. 170, nos. 1–2, pp. 245–258, Sep. 2000.

[17] S. Rit, M. V. Oliva, S. Brousmiche, R. Labarbe, D. Sarrut, and G. C. Sharp, "The reconstruction toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the insight toolkit (ITK)," *J. Phys., Conf. Ser.*, vol. 489, no. 1, p. 4, 2014.

[18] R. E. Alvarez and A. Macovski, "Energy-selective reconstructions in X-ray computerised tomography," *Phys. Med. Biol.*, vol. 21, no. 5, p. 733, 1976.

[19] C. Maaß, E. Meyer, and M. Kachelrieß, "Exact dual energy material decomposition from inconsistent rays (MDIR)," *Med. Phys.*, vol. 38, no. 2, pp. 691–700, Jan. 2011.

[20] E. Roessl and R. Proksa, "K-edge imaging in X-ray computed tomography using multi-bin photon counting detectors," *Phys. Med. Biol.*, vol. 52, no. 15, p. 4679, 2007.

[21] T. Su, V. Kaftandjian, P. Duvauchelle, P. Douek, and Y. Zhu, "Quantitative material decomposition method for spectral CT imaging," in *Proc. 7th Conf. Ind. Comput. Tomogr.*, 2017, pp. 1–7.

[22] B. Xie, P. Niu, T. Su, V. Kaftandjian, L. Boussel, P. Douek, F. Yang, P. Duvauchelle, and Y. Zhu, "ROI-wise material decomposition in spectral photon-counting CT," *IEEE Trans. Nucl. Sci.*, vol. 67, no. 6, pp. 1066–1075, Jun. 2020.

[23] M. Salehjahromi, Y. Zhang, and H. Yu, "A spectral CT denoising algorithm based on weighted block matching 3D filtering," *Proc. SPIE*, vol. 10391, Sep. 2017, Art. no. 103910G.

[24] Y. Zhang, M. Salehjahromi, and H. Yu, "Tensor decomposition and non-local means based spectral CT image denoising," *J. X-Ray Sci. Technol.*, vol. 27, no. 3, pp. 397–416, Jul. 2019.

[25] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

• • •