IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Low-Rank Tensor Train and Self-Similarity Based Spectral CT Reconstruction

Jie Guo[1], Xiaohuan Yu[1], Shaoyu Wang[1], Ailong Cai[1], Zhizhong Zheng[1], Ningning Liang[1], Lei Li[1] and Bin Yan[1]

[1]Henan Key Laboratory of Imaging and Intelligent Processing, PLA Strategic Support Force Information Engineering University, Zhengzhou, China

Corresponding author: Bin Yan (e-mail: ybspace@hotmail.com).

This work was supported by the National Natural Science Foundation of China under Grant 62271504, Grant 62201616, and Grant 62101596. This work was also supported by the National Key Research and Development Project of China under Grant 2020YFC1522002 and the China Postdoctoral Science Foundation under Grant 2019M663996.

ABSTRACT Spectral computed tomography (CT) has a wide range of applications in material discrimination, clinical diagnosis and tissue representation. However, the photon counting detector measurements are subject to serious quantum noise caused by photon starvation, photon accumulation, charge sharing, and other factors, which will lead to a decrease in the quality of the reconstructed image and make clinical diagnosis more difficult. To tackle with this problem, this paper proposes a spectral CT reconstruction technique that exploits the spatial sparsity of inter-channel images and the high correlation of images between different energy channels. Specifically, similar patches from spatial and spectral domains are extracted to form the low-rank tensors. Then the tensor-train rank, which is derived from a well-balanced matricization technique, is adopted to depict the high correlation among different energy channels. To capture the self-similarity of the low-rank tensors, the L0-norm of the image gradient is employed for image smoothing. An efficient algorithm is devised to solve the reconstruction model utilizing the Alternating Direction Method of Multipliers. For the sake of testing and verifying the effectiveness of the proposed algorithm, numerical simulations and real data experiments are conducted. Qualitatively, the designed method demonstrates a clear advantage in image quality over the existing state-of-the-art algorithms. For instance, when taking the full energy bin image as an example, the proposed method reduces the Root Mean Square Error (RMSE) by 52.07%, 38.69%, 35.13%, 12.67%, respectively, compared to the competing methods. Quantitative and qualitative assessment indices have revealed that the suggested method has excellent noise suppression, artifact elimination, and image detail preservation properties.

INDEX TERMS Spectral computed tomography, image reconstruction, non-local similarity, tensor train decomposition, alternating direction method of multipliers.

I. INTRODUCTION

With the advent of photon-counting detector (PCD) [1], Spectral computed tomography (CT) has a vast potential for medical diagnosis, such as tumor detection [2], [3], material discrimination [4], [5] and automated bone removal [6]. Spectral CT can acquire comprehensive material information by taking advantage of the difference in attenuation coefficients of materials. Furthermore, spectral CT imaging provides high tissue contrast and patients are exposed to low-dose radiation. [7], [8]. However, the number of photons allotted to each channel through the established energy threshold is inadequate, and the quantum noise of the projection is particularly pronounced [9], [10]. There are numerous classical image reconstruction techniques available to reduce noise, such as the iterative algorithm based on image sparsity regularization in projection domain or image domain such as total variation (TV) and its modifications [11-13], tight frame (TF) [14], dictionary learning (DL) [15], [16], tensor factorization [17] and the non-local similarity [18]. The PRISM (prior rank, intensity and sparsity) algorithm has been successfully applied to low-dose CT reconstruction [19], yielding superior results. While these algorithms based on compressed sensing [20] are partly effective in restoring image structure, they fail to capture the image correlation. To address this issue, several tensor-based algorithms have been proposed in succession to characterize the multi-dimensional features of spectral CT images. In [21], the authors proposed a spectral CT image

reconstruction algorithm that leverages tensor dictionary learning and the L0-norm to effectively extract features and maintain image edge details. Furthermore, by incorporating the adaptive spectral prior information, the framework in [21] was extended in [22]. Subsequently, based on tensor expansion, the authors derived the tensor form of PRISM algorithm in [23], which achieved superior quality by fully exploiting the sparsity.

For spectral CT imaging, the low-rankness and high correlation of tensor are advantageous priors, yet they are not easy to handle or to illustrate clearly. To approximate the low-rankness, the tensor nuclear norm has been widely used in image processing [24], [25]. To capture the high correlation, tensor decomposition has been developed to optimize reconstruction by combining the tensor nuclear norm. Tucker decomposition [26] was adopted to explore the global correlation between cerebral perfusion CT images in [27] and the dimensions reduction in [28]. Candecomp/Para-fac (CP) decomposition [29] was employed to enhance sparse reconstruction based on self-similarity in [30]. Moreover, the authors developed an intrinsic tensor sparsity regularization model based on Tucker decomposition and CP decomposition to characterize the low-rankness [31],[32] and sparsity of tensor [33]. To explore the correlation of tensors more comprehensively, the work in [34], [35] constructed low rank tensor model by clustering similar patches and then decomposed it into low rank tensor and sparse tensor. This approach enabled multi-channel image reconstruction to achieve remarkable results. However, according to the work in [31], [32], the CP decomposition cannot accurately depict the low-rankness of the subspace formed by the modules of tensors, while the Tucker decomposition was difficult to evaluate the sparsity of the core tensor in theory. In [36], the authors established the perfect tensor singular value decomposition(t-SVD) theory through tensor-tensor product. The first and second mode correlation may be characterized by t-SVD, but the third mode's correlation are encoded by embedded circular convolution, which means that the tensor nuclear norm lacks a direct measurement of the third dimension [37], [38]. It is essential to make full use of the high correlated statistical properties of different channels to remove noise and artifacts, which is the core problem of this study. Therefore, it is necessary to seek a suitable tensor decomposition to describe this high correlation between different energy channels to improve the reconstruction images.

In recent decades, tensor train (TT) rank has attracted much attention due to its well-balanced matricization technique. This technique involves flattening the tensor along permutations of modes, which has been effectively used in tensor completion and image restoration as it can better capture the global correlation between different tensor modes [39], [40]. Moreover, a low-order tensor can be augmented in order by employing a suitable block addressing technique, which makes it easier to depict the correlation between various modes [41],[42]. Furthermore, the work [43] constructed high order low-rank tensor groups by exploiting the non-local self-similarity, and made an exhaustive theoretical analysis.

Taking inspiration from the previous work, we present a spectral CT reconstruction technique that makes good use of the low-rank tensor train rank and the self-similarity to preserve details. To begin, similar patches from the initial images are extracted and low-rank tensor groups are constructed. The spectral CT reconstruction model is then developed employing the tensor train nuclear norm (TTNN), which approximates the tensor train rank, and the L0-norm of image gradient. To the best of our knowledge, this is the first time that the TTNN was used in spectral CT imaging, allowing us to better investigate the high spatial-spectral correlation and the non-local self-similarity of spectral CT images due to its well-balanced matrix expansion techniques. Additionally, for the TTNN, we adopted the direct decomposition algorithm, which has a fast decomposition speed and high calculation efficiency. In comparison to other iterative least square algorithms, such as CP decomposition, the direct decomposition consumes less time and less physical memory. Finally, using k-means++clustering, we create low-rank tensor units that effectively encoded non-local self-similarity while reducing computational complexity.

The following is the organization structure of this study. Part II briefly introduces the preliminary knowledge of tensors, and presents the model and its optimization of low rank tensor train applied to spectral CT reconstruction. In part III, the proposed method is evaluated through simulation experiments and actual data experiments. Both quantitative results and property assessment show that the presented method is better than the competing algorithms. The last part is discussion and conclusion.

II. METHODS

A. PRELIMINARIES ON TENSOR

Multi-channel image data could be regarded as a 3rd-order tensor, including two spatial dimensions and one spectral dimension. In general, an *N*-order tensor could be expressed as $\mathcal{X} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$. Tensor and matrix can be transformed by unfolding or folding operator [26]. A tensor can be expanded into a matrix along the *k*-th dimension

$$\text{unfold}_k(\mathcal{X}) = X^k \in \mathbb{R}^{J_k \times (J_1 \cdots J_{k-1} J_{k+1} \cdots J_N)}. \quad (1)$$

The correlation of different channel images could be described by tensor rank. However, tensor does not have a definite rank compared with matrix. With different expansion pattern, the expression of tensor rank is different. Tucker rank [26] of an N -order tensor $\mathcal{X} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ is denoted as

$$\text{rank}_{tc}(\mathcal{X}) := (\text{rank}(X^1), \text{rank}(X^2), \cdots, \text{rank}(X^N)), \quad (2)$$

Where X^n is the mode-n expansion of $\mathcal{X}$ and the $rank(X^i)$ provides the mean of the correlation between a single mode

and other modes of the tensor. Both the matrix rank and the tensor rank are powerful prior to characterize the correlation and the non-local self-similarity. However, minimizing the matrix rank is a difficult problem. In general, the nuclear norm of the matrix could approximat the matrix rank defined as follows [44].

$$\left\| X_{n_1 \times n_2} \right\|_* = \sum_i \sigma_i (X_{n_1 \times n_2}), \tag{3}$$

$\sigma_i (X_{n_1 \times n_2})(i = 1, 2, \cdots, \min(n_1, n_2))$ is the i-th singular value of the matrix $X_{n_1 \times n_2}$. Considering that the nuclear norm of a matrix is the approximation of matrix rank, the sum of the nuclear norms $\sum_i \left\| X^i \right\|_*$ is used as the approximation of the sum of the rank $\sum_i rank(X^i)$. Similarly, according to the tensor nuclear norm introduced in the work [36], we obtain the expression of the tensor nuclear norm $\left\| \mathcal{X} \right\|_* = \sum_{k=1}^{n_3} \left\| X^{(k)} \right\|_*$,where $X^{(k)}$ is the k-th frontal slice of tensor.

B. TENSOR-TRAIN NUCLEAR NORM

Tensor-train decomposition, which has attracted a lot of attention in recent years, is a method of expanding a tensor in a relatively balanced way. For an N-order tensor $\mathcal{X} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$, its tensor-train decomposition is as follows

$$X_{[k]} \in \mathbb{R}^{p_k \times q_k} (p_k = \prod_{i=1}^{k} J_i, q_k = \prod_{i=k+1}^{N} J_i), k = 1, 2, \cdots, N. \tag{4}$$

The TT rank is denoted by vector $\boldsymbol{r} = (r_1, r_2, \cdots, r_{N-1})$, where the $r_k = rank(X_{[k]})$ is the rank of matrix $X_{[k]}$ in (4). In light of this, the TTNN is defined as a convex combination of nuclear norms of different matrices generated via tensor expansion by (4), namely

$$\left\| \mathcal{X} \right\|_{\text{TTNN}} = \sum_{k=1}^{N} \alpha_k \left\| X_{[k]} \right\|_*, \tag{5}$$

where the weights $\{\alpha_k\}_{k=1}^{N}$ satisfying $\sum_{k=1}^{N} \alpha_k = 1, \alpha_k > 0$ represents the weight of $X_{[k]}$. Setting different weights is due to the following considerations. The importance of tensor train rank of different modes is different. The adaptive weighting process establishes the weight α_k [45].

Appendix A demonstrates that the tensor rank derived from the tensor expansion in (1) only accounts for the correlation between a single mode and the other modes (one mode versus the rest), while ignoring the global correlation between the modes. While tensor train rank based on tensor train expansion in (4) takes into account the correlation between some modes and others more fully. Because of the well-balanced matricization technique [40], it can describe the global correlation among the tensor entries in great detail. To demonstrate that the TTNN acquired by tensor train decomposition as described in (4) is superior to the nuclear norm obtained by Tucker decomposition and the nuclear norm obtained by tensor singular value decomposition [46]. We conducted simulation experiments in comparison to other norms such as the TNN (tensor nuclear norm) [36], PSTNN (partial sum of the tubal nuclear norm) based on tensor tube rank [37], SNN (sum of nuclear norm) based on Tucker rank [47]. Gaussian noise of 0.001 was added to the original data. The denoising experiment results were shown in figure 1, which presents the better denoising performance by TTNN. The peal signal to noise ratio (PSNR) and structural similarity (SSIM) of all denoised results are calculated in table 1. The PSNR and SSIM confirm the consistent results for SNN, TNN and PSTNN in [38],[42].

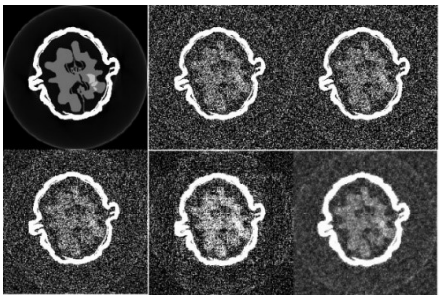

FIGURE 1. **The images in row 1 are ground truth, noised image, denoised image by SNN, respectively. Row 2 are the denoised images by PSTNN, TNN, TTNN. The display window of row 1 and the first image in row 2 are [0.001 0.2]. The rest in row 2 are [0.02 0.1] and [0.009 0.15], respectively.**

TABLE I THE PSNR AND SSIM GAINED BY METHODS

Index	SNN	PSTNN	TNN	TTNN
PSNR	21.1407	22.2526	22.0489	26.1324
SSIM	0.1720	0.3614	0.1939	0.4260

The problem of optimizing the TTNN is expressed as

$$\min_{\mathcal{X}} \frac{\mu}{2} \left\| \mathcal{X} - \mathcal{Y} \right\|^2 + \left\| \mathcal{X} \right\|_{\text{TTNN}}, \tag{6}$$

where $\mathcal{Y}$ is an observed noised tensor. Substituting (5) into (6), the following equation can be obtained

$$\min_{X_{[k]}} \frac{\mu}{2} \left\| X_{[k]} - Y_{[k]} \right\|_{\text{F}}^2 + \sum_{k=1}^{N} \alpha_k \left\| X_{[k]} \right\|_*, \tag{7}$$

where $X_{[k]}, Y_{[k]}$ are the matrices obtained by tensor train expansion of $\mathcal{X}, \mathcal{Y}$ according to (4). The closed-form solution of (7) can be acquired by employing the Singular Value Thresholding (SVT) in [44],

$$\hat{X}_{[k]} = \sum_i \max(\sigma_i - \frac{\alpha_k}{\mu}, 0) u_i v_i^{\text{T}}. \tag{8}$$

where σ_i are the singular values and u_i, v_i the singular vectors of the matrix $Y_{[k]}$.

C.LOW-RANK TENSOR-TRAIN MODEL FOR SPECTRAL CT RECONSTRUCTION

Spectral CT images are usually expressed as 3rd -order tensor $\mathcal{X} \in \mathbb{R}^{n_h \times n_w \times n_s}$, where n_h, n_w and n_s are the spatial dimension and the spectral dimension of the image, respectively. Similarly, spectral CT projections can also be written as a 3rd-order tensor $\mathcal{Y} \in \mathbb{R}^{n_d \times n_v \times n_s}$, where n_d, n_v are the number of

detector elements and projections, respectively. Spectral CT images can be regarded as the sequence of traditional CT images. If extended along some channel, the usual linear system can be obtained

$$y_s = Ax_s (s = 1, 2, \cdots, S), \tag{9}$$

where the system matrix $A \in \mathbb{R}^{(n_d \cdot n_v) \times (n_h \cdot n_w)}$, x_s is the vectorization results of two-dimensional image $X_s \in \mathbb{R}^{n_h \times n_w}$, that is the attenuation coefficient we calculated, and y_s is the vectorization results of two-dimensional projection $Y_s \in \mathbb{R}^{n_d \times n_v}$. The tensor nuclear norm is used to approximate a tensor's low rank, and L0-norm of image gradient is commonly used to describe the spatial-spectral sparsity. To fully encode the low rank and sparsity of images, the spectral images are pre-treated with a similar block matching operation. The noise image block of a specific channel is first selected, and then the similar block is matched. Finally, similar blocks are extracted from different channels at the corresponding positions to form small low-rank tensor unit. The block extraction and clustering methods used in the work [34] will result in inconsistent clustering after repeated iterations, increasing the calculation's complexity. In this study, the non-local self-similarity of images is investigated using the k-means++ clustering method [48]. This preprocessing operation on tensors is recorded as $\mathcal{R}_c(x) \in \mathbb{R}^{J_1 \times J_2 \times J_3}$, where $c = 1, 2, \cdots, C$ C denotes the total number of tensor units.

The spectral CT reconstruction model is constructed as follows

$$\min_{x_s} \frac{1}{2}\sum_{s=1}^{S}\left\|Ax_s - y_s\right\|_2^2 + \lambda\sum_{s=1}^{S}\left\|\nabla x_s\right\|_0 + \rho\sum_{c=1}^{C}\left\|\mathcal{R}_c(\mathcal{X})\right\|_{\mathrm{TTNN}}. \tag{10}$$

The first regularization term is intended to describe the sparsity of the gradient image in each energy channel, which can have a positive effect on noise suppression. The second regularization term aims to depict the correlation between multi-channel images, which can improve the model's ability to protect the details of the image. The regularization parameters λ, ρ will balance the above terms. In order to solve (10), we introduce the auxiliary variables $\mathcal{M}_c = \mathcal{R}_c(\mathcal{X})$ and $f_s = x_s$. We can gain the unconstrained objective function.

$$\begin{aligned}\underset{x_s, \mathcal{M}_c, f_s, u_s, \mathcal{L}}{\arg\min} &\sum_{s=1}^{S}\left(\frac{1}{2}\left\|Ax_s - y_s\right\|_2^2 + \lambda\left\|\nabla f_s\right\|_0 + \left\langle u_s, f_s - x_s\right\rangle + \frac{\beta}{2}\left\|f_s - x_s\right\|_{\mathrm{F}}^2\right) \\ &+\rho\sum_{c=1}^{C}\left\|\mathcal{M}_c\right\|_{\mathrm{TTNN}} + \left\langle \mathcal{L}_c, \mathcal{M}_c - \mathcal{R}_c(\mathcal{X})\right\rangle + \frac{\gamma}{2}\sum_{c=1}^{C}\left\|\mathcal{M}_c - \mathcal{R}_c(\mathcal{X})\right\|_{\mathrm{F}}^2\end{aligned} \tag{11}$$

where $u_s, \mathcal{L}_c$ are Lagrange multipliers and β, γ are regularization parameters. The function can be optimized by alternative direction minimization method (ADMM) [49], and $\mathcal{M}_c$ can be updated with fixed $x, \mathcal{L}_c, \mathcal{R}_c^t(\mathcal{X})$, then obtain the closed-form solution by referring to (8).

$$\mathcal{M}_c^{t+1} = \underset{\mathcal{M}}{\arg\min}\ \rho\sum_{c=1}^{C}\left\|\mathcal{M}_c\right\|_{\mathrm{TTNN}} + \frac{\gamma}{2}\sum_{c=1}^{C}\left\|\mathcal{M}_c - \mathcal{R}_c^t(\mathcal{X}) + \frac{\mathcal{L}_c}{\gamma}\right\|_{\mathrm{F}}^2. \tag{12}$$

Similarly, updating f_s with fixed x_s as follows:

$$f_s^{t+1} = \underset{f_s}{\arg\min}\ \lambda\sum_{s=1}^{S}\left\|\nabla f_s\right\|_0 + \sum_{s=1}^{S}\frac{\beta}{2}\left\|f_s - x_s^t + \frac{u_s^t}{\beta}\right\|_{\mathrm{F}}^2. \tag{13}$$

For (13), it is a typical L0-norm minimization problem, which can be solved by approximation algorithm of l_0-smoothing [50]. For the following x_s-subproblem

$$\begin{aligned}x_s = \underset{x_s}{\arg\min}\ &\frac{1}{2}\sum_{s=1}^{S}\left\|Ax_s - y_s\right\|_2^2 + \frac{\beta}{2}\sum_{s=1}^{S}\left\|x_s - f_s^{t+1} + \frac{u_s^t}{\beta}\right\|_{\mathrm{F}}^2 \\ &+\frac{\gamma}{2}\sum_{c=1}^{C}\left\|\mathcal{M}_c^{t+1} - \mathcal{R}_c^{t+1}(\mathcal{X}) + \frac{\mathcal{L}_c^t}{\gamma}\right\|_{\mathrm{F}}^2.\end{aligned} \tag{14}$$

it is going to be updated, and the elements of $\mathcal{M}_c, \mathcal{R}_c(\mathcal{X})$ and $\mathcal{L}_c$ are vectorized along mode-1, mode-2 and mode-3, and record them as $\boldsymbol{m}_c, R_c(x)$ and Λ_c respectively. The x_s-subproblem can be rewritten as

$$\begin{aligned}x_s = \underset{x_s}{\arg\min}\ &\frac{1}{2}\sum_{s=1}^{S}\left\|Ax_s - y_s\right\|_2^2 + \frac{\beta}{2}\sum_{s=1}^{S}\left\|x_s - f_s^{t+1} + \frac{u_s^t}{\beta}\right\|_{\mathrm{F}}^2 \\ &+\frac{\gamma}{2}\sum_{c=1}^{C}\left\|\boldsymbol{m}_c^t - \boldsymbol{R}_c(x) + \frac{\Lambda_c^t}{\gamma}\right\|_{\mathrm{F}}^2.\end{aligned} \tag{15}$$

Note that each term in (15) is a quadratic form, and it can be solved by the separate quadratic surrogate (SQS) method to update x_s as

$$x_s^{t+1} = x_s^t - \frac{A^{\mathrm{T}}(Ax_s^t - y_s) + P(x_s^t)}{A^{\mathrm{T}}A\mathbf{1} + \gamma\sum_{c=1}^{C}\boldsymbol{R}_c^{\mathrm{T}}\boldsymbol{R}_c\mathbf{1} + \beta}, \tag{16}$$

where $P(x_s^t) = \gamma\sum_{c=1}^{C}\boldsymbol{R}_c^{\mathrm{T}}\left(\boldsymbol{R}_c(x^t) - \frac{\Lambda_c^t}{\gamma} - \boldsymbol{m}_c^{t+1}\right) + \beta\left(x_s^t - f_s^{t+1} + \frac{u_s^t}{\beta}\right)$.

The calculation of formula (16) is performed in the element-wise manner. Finally, the Lagrange multipliers in (11) can be updated by gradient descent:

$$\mathcal{L}_c^{t+1} = \mathcal{L}_c^t + \gamma(\mathcal{M}_c^{t+1} - \mathcal{R}_c^{t+1}(\mathcal{X})) \tag{17}$$

$$u_s^{t+1} = u_s^t + \beta(f_s^{t+1} - x_s^{t+1}). \tag{18}$$

The flow chart of the devised method is given in Algorithm 1.

Algorithm 1 The workflow of proposed algorithm.

Input: the projection y_s and the parameters $\lambda, \beta, \rho, \gamma, C$.
Initialization: the system matrix A and $\mathcal{X}^0, \mathcal{M}^0, f^0$.
While not meet the stop criterion
For $t = 1$: T **Do**
Updating $\mathcal{M}_c$ by the SVT via (12);
Updating f_s via (13);
Updating x_s by the SQS method via (15), (16);
Updating $\mathcal{L}_c, u_s$ by the gradient descent via (17), (18);
$t = t + 1$
End
Output: The reconstructed image x_s.

III. EXPERIMENTS AND RESULTS

A. EXPERIMENTAL DATA

In this part, we first test the feasibility of the designed method, and then evaluate its behavior through the numerical mouse simulations and real data experiments. For the numerical simulations, the scan protocol and energy threshold setting are same as [30]. The distances between the X-ray source and the object and the PCD are 132 mm and 180 mm respectively. The PCD has 512 units and each of them is 0.1 mm length. A polychromatic 50kVp X-ray spectrum is generated and the intervals for the X-ray energy are divided into 8 channels as [16-22) keV, [22-25) keV, [25-28) keV, [28-31) keV, [31-34) keV, [34-37) keV, [37-41) keV and [41-50) keV. Figure 2(a) demonstrates the ground truth image for the simulation experiments and figure 2 (b) demonstrates the standardized X-ray spectrum with eight energy channels. The scanning angle covers 360° with an interval of 2° or 4°, namely we get 180 projection views or 90 projection views. The simulation of image noise was realized by adding the Poisson noise to projections as follows:

$$I_p = \text{Poisson}(I_0 \exp(-y_s)), \tag{19}$$

Where I_0 denotes the number of X-ray photons and y_s denotes the noise-free projection. In this work, I_0 was set to 5×10^3 and 1×10^6, respectively.

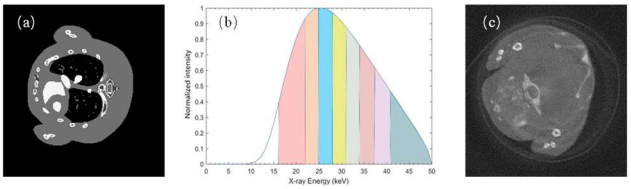

FIGURE 2. **Simulated and real experimental data. (a) The truth image for simulation experiments; (b) Standardized X-ray spectrum with eight energy channels; (c) Real mouse data reconstructed image of channel 1 by FBP.**

To represent the effectiveness of the suggested method, several spectral CT image reconstruction methods are selected as competing algorithms, including SART, LRTV(Low-Rank TV)[28], FTNN(Framelet Tensor Nuclear Norm) [24] and SISTER(Spectral-Image Similarity-based Tensor with Enhanced-sparsity Reconstruction) [30]. The goal of the proposed algorithm is to characterize the low-rankness and high correlation between multi-channel image and perform high-quality spectral CT reconstruction with low-dose radiation. The initial image is reconstructed by FBP, and the number of iterations were put to be 100 for the simulation experiment and 20 for the real data study. We use the sequential subset synchronous iterative reconstruction with a limit of 10 subsets to accelerate the convergence of the image. The parameters in the comparison algorithms are adjusted to the optimum level in order to evaluate the algorithm fairly. In numerical simulation experiments, the iterative methods are stopped at iteration 45 in the LRTV and the FTNN, and 100 in the SART and the SISTER. The regularization parameters in the LRTV are set to be $\beta = 1.9, \rho = 0.0012$. While the regularization parameter $\beta = 100$ and patch size and similar patch numbers are set to 6 and 50 in the FTNN, respectively. The selection of the parameters in SISTER are similar to that in reference [30], and then slightly modified according to the experimental results. In real data experiments, the iterative methods are stopped at iteration 15 in the FTNN, the SART and the SISTER, 20 in the LRTV. We adjusted the according regularization parameters based on the simulation experiments.

In both simulation experiment and real data experiment of LRTT, the parameters were set $\lambda = 1.8\times10^{-5}$, $\beta = 1.1$, $\gamma = 5.6$, $\rho = 22$. The size of the small tensor in each channel is 8×8, and the group number in k-means++clustering method was set to be 152 for the simulation experiment and 100 for the read data study. In this study, we choose the region of interest (ROI) containing complex texture structure for magnification and comparison. In the simulation experiment, the root means square error (RMSE), the PSNR and the SSIM of all reconstruction results are calculated.

Note: *For the convenience of writing, the proposed algorithm in this paper is abbreviated as LRTT (Low-Rank Tensor Train) in the following illustrations and documents.*

B. NUMERICAL SIMULATIONS

To test the feasibility of the devised algorithm, we conducted a preliminarily test with the results shown in figure 3 under low-level noise $I_0 = 1\times10^6$ with size of $256\times256\times8$. It is obvious that the LRTT method is capable of reducing noise and demonstrates its benefits. We performed the simulation experiments with different projection views and high-level noise $I_0 = 5\times10^3$ with size of $512\times512\times8$ to further validate the designed algorithm in figure 4 and 6, which show the results of 180 projections and 90 projections, respectively.

The reconstructed images from different algorithms differ significantly, as shown in figures 4 (a)-(f) and 6 (a)-(f), and images produced by the LRTT method outperform the competing algorithms in fine structure and clear edge. Figures 4 (b1) -(b4) and 6 (b1) -(b4) show that the images reconstructed by SART yielded the lowest image quality and were blurred by severe noise artifacts in four energy bins. The LRTV-based algorithm effectively suppresses noise and produced the cleaner image, but with the visible blocky artifacts shown in figures 4 (c1) -(c4) and 6 (c1) -(c4). In comparison to SART and LRTV, the FTNN and SISTER results provide much structures and reduce a significant amount of noise. However, FTNN and SISTER lose the finer structure of the object and blur the edges, as shown in figures 4 (d1)- (e4) and 6 (d1) -(e4), respectively. Furthermore, the two algorithms were unable to clearly reconstruct the specific region denoted by the red box, including the detail information in figure 4 (a1). LRTT, on the other hand, produces impressive image results, suppresses the noise-induced artifacts, preserves the small structure and results in the remarkable visual effect shown in figure 4 (f1) - (f4). To see the results of different methods more clearly, the ROI "A" and "B" in figures 3 and 4 and the corresponding gradient images are

exhibited in figure 5. In SART, the bone structure in "A" was obscured by noise, blurred by blocky artifacts in LRTV, and well preserved in FTNN and SISTER. The ROI "A" and gradient images show that the LRTT method has a significantly clear contour profile and detail structure. Furthermore, the finer details highlighted by the red arrow could be seen in the LRTT but not in other comparison methods.

Figures 4 and 6 show that the quality of the reconstructed image improves as the number of projection views increases, implying that the more projection views, the higher the image quality. Notably, the image structure of the soft tissue and the bone structure become more distinct.

The line profiles along the yellow dashed line in figure 4 (a1) were drawn in figure 7. It is obvious to see that the line profiles of SART and LRTV fluctuate significantly, with a significant deviation from the true value. However, while the FTNN and SISTER achieve better results, some areas cannot be completely restored (as seen in the area marked by the red arrow in figure 7). LRTT line profiles, on the other hand, are closest to the true value line and have the smallest deviation from the ground truth. Furthermore, the LRTT is more accurate than other algorithms, particularly in areas with complex structures (marked by the red arrows). It is more visible in the magnified images, as illustrated in figure 7 (a2) and (b2), the LRTT agrees with the ground truth with the least amount of error.

To accurately measure the overall performance of various algorithms, RMSE, SSIM and PSNR are displayed in figures 8 and 9. The represented method clearly outperforms the other competing algorithms. Notably, among all the competing methods for all channels, the LRTT algorithm achieves the highest SSIM and the lowest RMSE. When considering the full energy bin image, for instance, the LRTT method reduced the RMSE by 52.07%, 38.69%, 35.13%, 12.67%, respectively, when compared to the SART, LRTV, FTNN and SISTER. According to the SSIM and RMSE results, the LRTT can effectively recover the internal structure and reduce the error with the ground truth, which is consistent with the previous analysis. As can be seen from the curves of the evaluation indicators in figure 8 and 9, the noise increases as the projection views decrease, and the evaluation indicators of the competing algorithm deteriorate, while the LRTT remains relatively unaffected, indicating that the LRTT algorithm has excellent anti-noise capability.

The LRTT algorithm yields the lowest RMSE and a rate of convergence that is almost as fast as the SISTER. Furthermore, when compared to the SISTER, the LRTT significantly enhances the calculation efficiency and decreases caculation time and physical memory, as shown in Table III and IV. The LRTT takes much less time and less memory resources than the SISTER, which confirms our conclusion. The SISTER requires an average of 45 seconds and 49 seconds for each iteration with 180 projection views under high and low noise, respectively. The LRTT algorithm, on the other hand, takes an average of 12.2 seconds and 11.7 seconds. The LRTT requires less processing power than that the SISTER, owing to a more direct tensor train decomposition and lower computational complexity, while the SISTER needs a large number of alternating least squares iterative operations in CP decomposition [30].

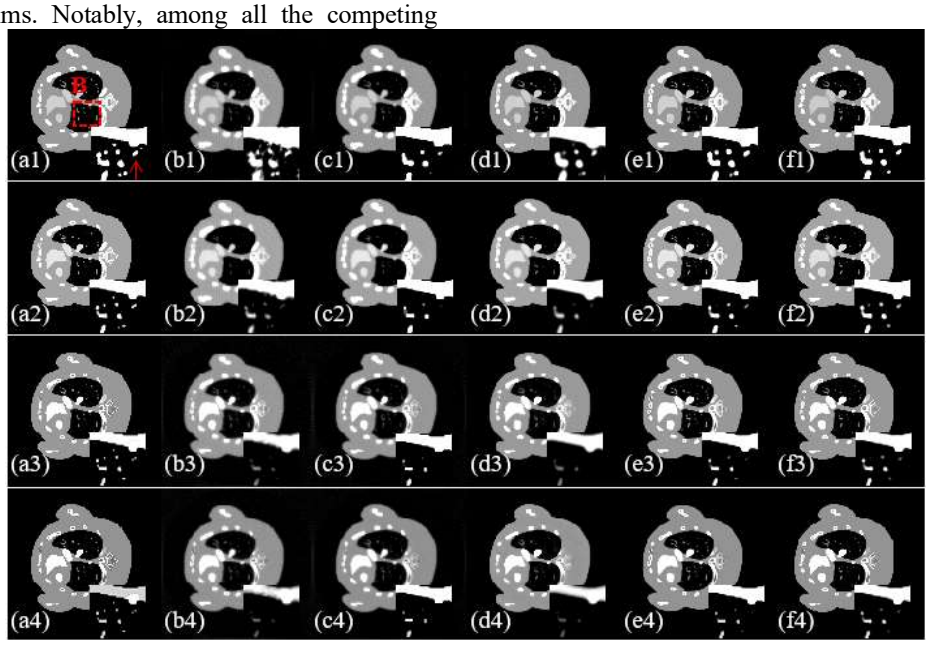

FIGURE 3. **Denoising results of different algorithms from 180 projections and $I_0 = 1 \times 10^6$. Rows 1 to 4 denote the 1th, 4th, 7th and 8th energy channel and the display windows are [0.01 0.19], [0.009 0.065], [0 0.055], [0 0.045], respectively. Column (a) shows the reference images, and columns (b) to (f) indicate the reconstruction images by SART, LRTV, FTNN, SISTER, LRTT method, respectively.**

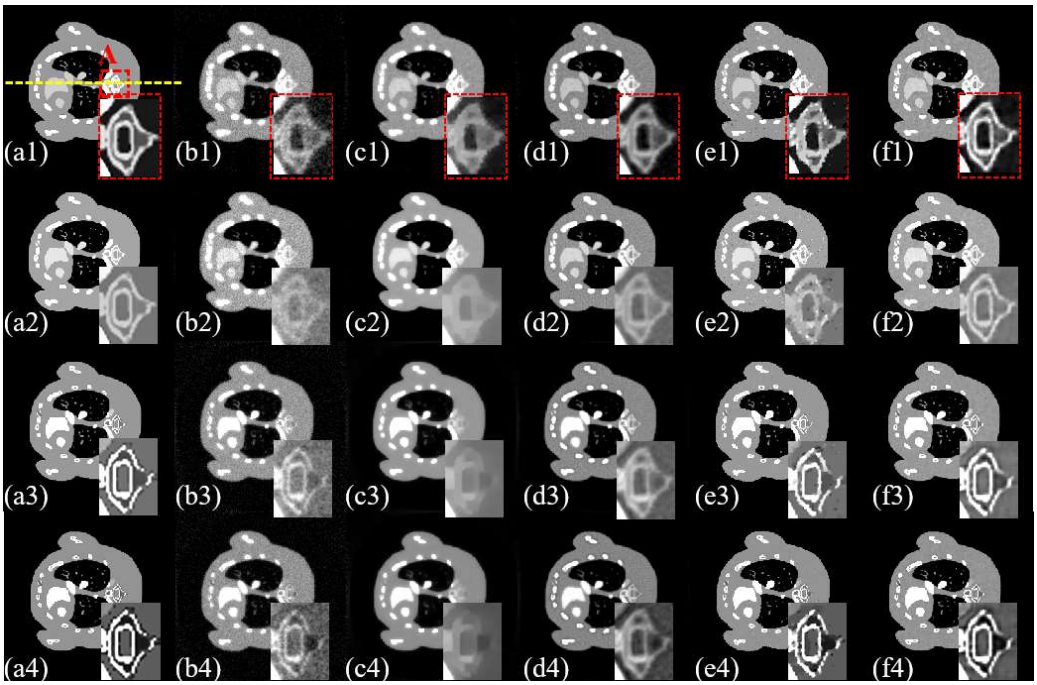

FIGURE 4. **Reconstruction images of different algorithms from 180 projections. Rows 1 to 4 denote the 1th, 4th, 7th and 8th energy channel and the display windows are set to [0.01 0.19], [0.009 0.065], [0 0.055], [0 0.045], respectively. Column (a) shows the reference images, and columns (b) to (f) indicate the reconstruction images by SART, LRTV, FTNN, SISTER, LRTT method, respectively.**

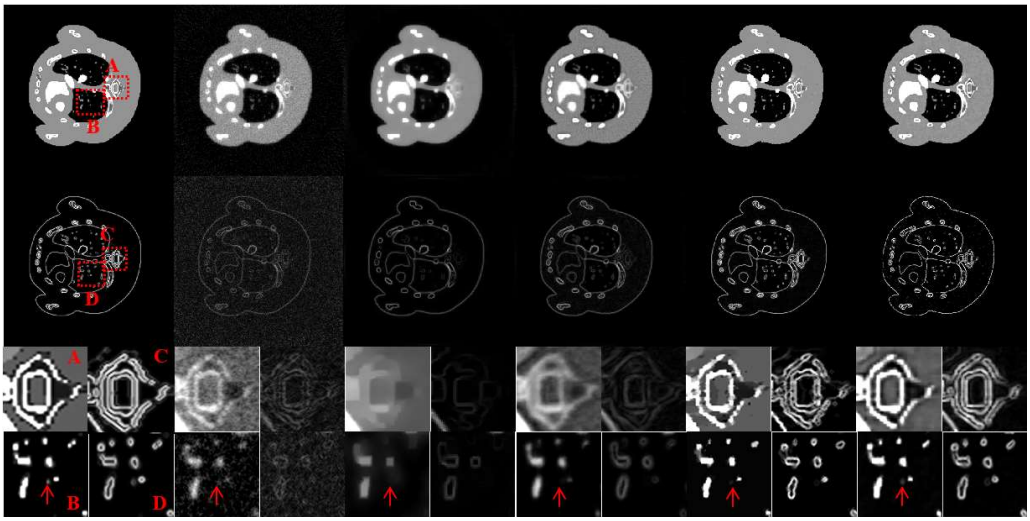

FIGURE 5. **Gradient images of the 8th energy channel from the different algorithms consistent with Figure4. Row 3 is the results of magnifying the ROI A, B, C and D in the corresponding image. The display windows are set to [0.0008 0.025] for the gradient image, [0.003 0.01] for ROI of the gradient image, and [0.007 0.025] for ROI of the reconstruction image.**

TABLE II MEAN VALUES AND STDS OF ROI A, B, E, F, G OF CHANNEL1 IN FIGURE 11

Metric	ROI	SART	LRTV	FTNN	SISTER	LRTT
Mean values ± STD	A	0.4465 ± 1.33e-2	0.4756 ± 3.79e-3	0.4908 ± 3.68e-3	0.5129 ± 4.07e-3	0.4705 ± 3.08e-3
	B	0.5003 ± 2.21e-2	0.4863 ± 6.31e-3	0.5035 ± 6.35e-3	0.5233 ± 7.02e-3	0.4826 ± 5.01e-3
	E	0.4761 ± 5.36e-3	0.4719 ± 1.38e-3	0.4885 ± 1.17e-3	0.5083 ± 1.24e-3	0.4690 ± 6.71e-4
	F	0.6380 ± 8.10e-2	0.5748 ± 1.41e-2	0.5936 ± 1.43e-2	0.6189 ± 1.71e-2	0.5696 ± 1.29e-2
	G	0.5792 ± 2.52e-2	0.5651 ± 6.71e-3	0.5854 ± 6.61e-3	0.6100 ± 7.58e-3	0.5616 ± 5.76e-3

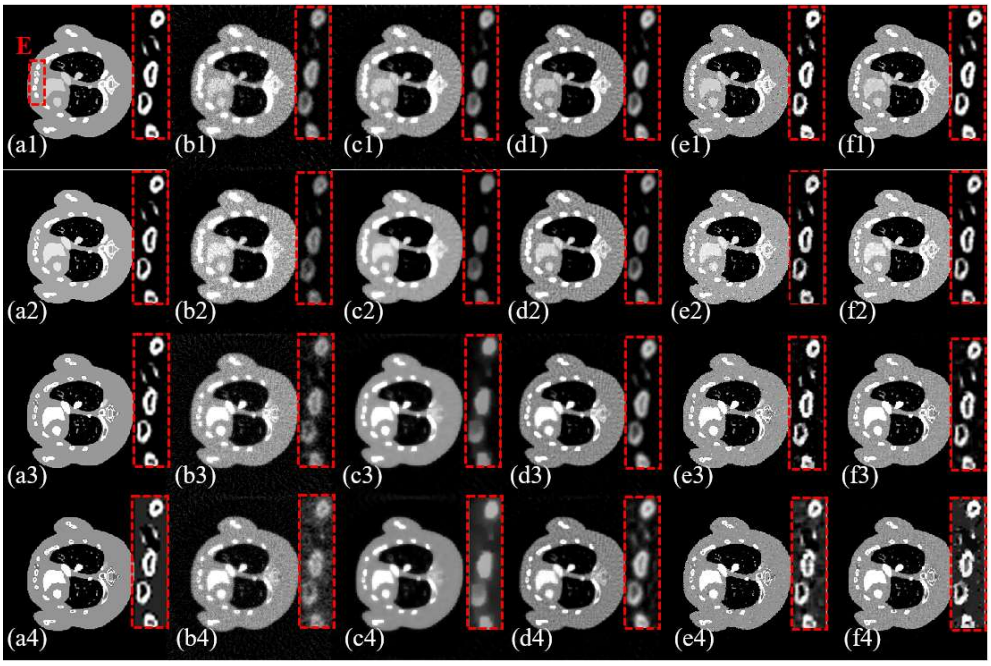

FIGURE 6. **Denoising results of different algorithms from 90 projections. Rows 1 to 4 denote the 1th, 4th, 7th and 8th energy channel image and the display windows are set to [0.01 0.16], [0.009 0.06], [0 0.05], [0 0.045], respectively. Column (a) shows the reference images, and (b) to (f) represent the reconstructions of SART, LRTV, FTNN, SISTER, LRTT method, respectively.**

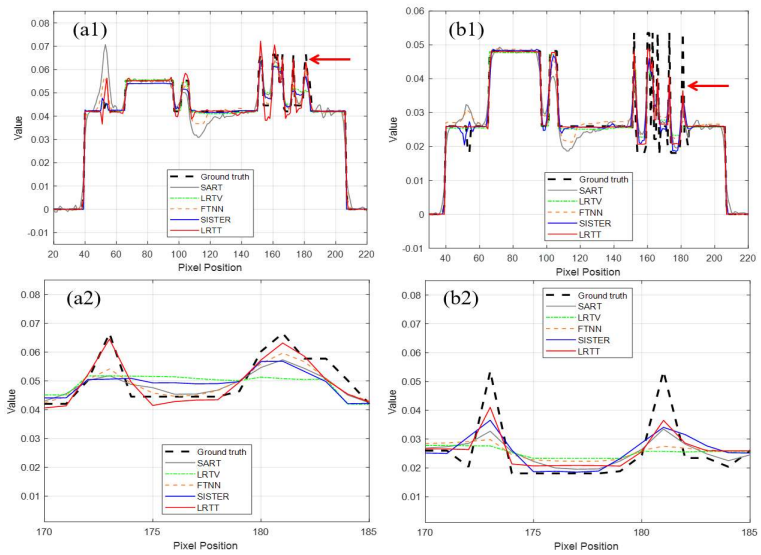

FIGURE 7. **Pixel values of denoising results with different algorithms. column(a) denotes the profile results of channel 4 and (b) is channel 8. Row 2 is the magnification of the corresponding area of red row**

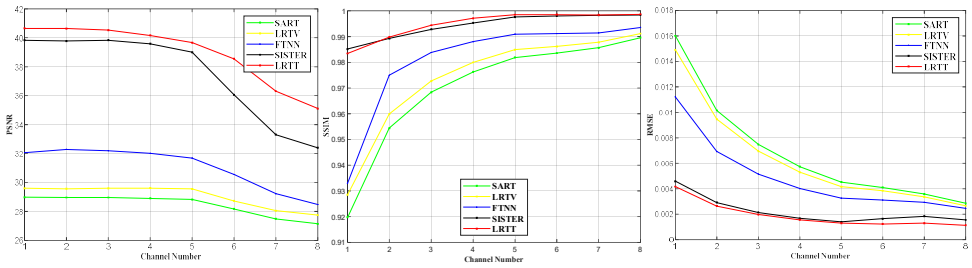

FIGURE 8. **Assessment of image quality for 180 projections and $I_0 = 5 \times 10^3$ with different algorithms. From left to right are the evaluation indicators of PSNR, SSIM and RMSE of different algorithms.**

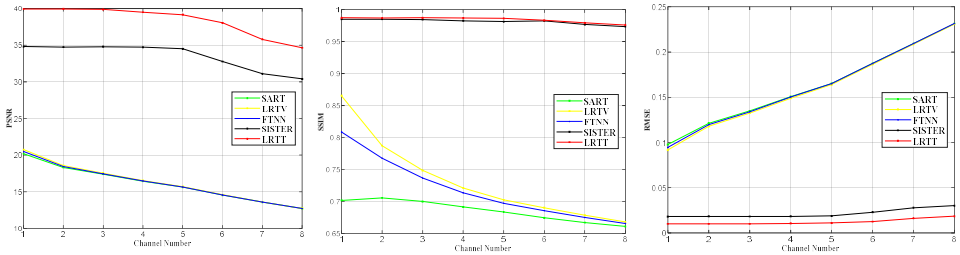

FIGURE 9. **Assessment of image quality for 90 projections and $I_0 = 5 \times 10^3$ with different algorithms. From left to right are the evaluation indicators of PSNR, SSIM and RMSE of different algorithms**

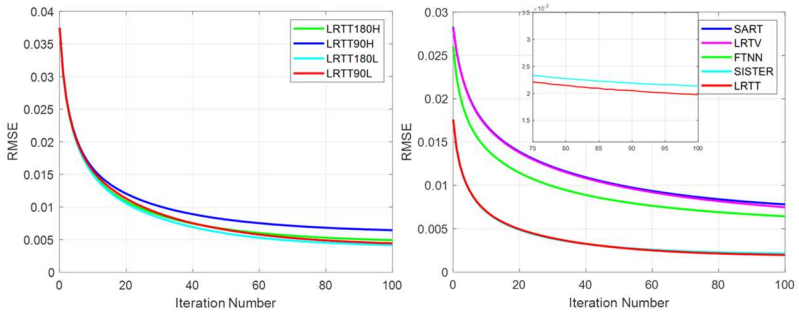

FIGURE 10. **Convergence curve in terms of RMSE. The left is the convergence curve of LRTT for the difference noises and views. The right is the convergence curve of RMSE for the different methods.**

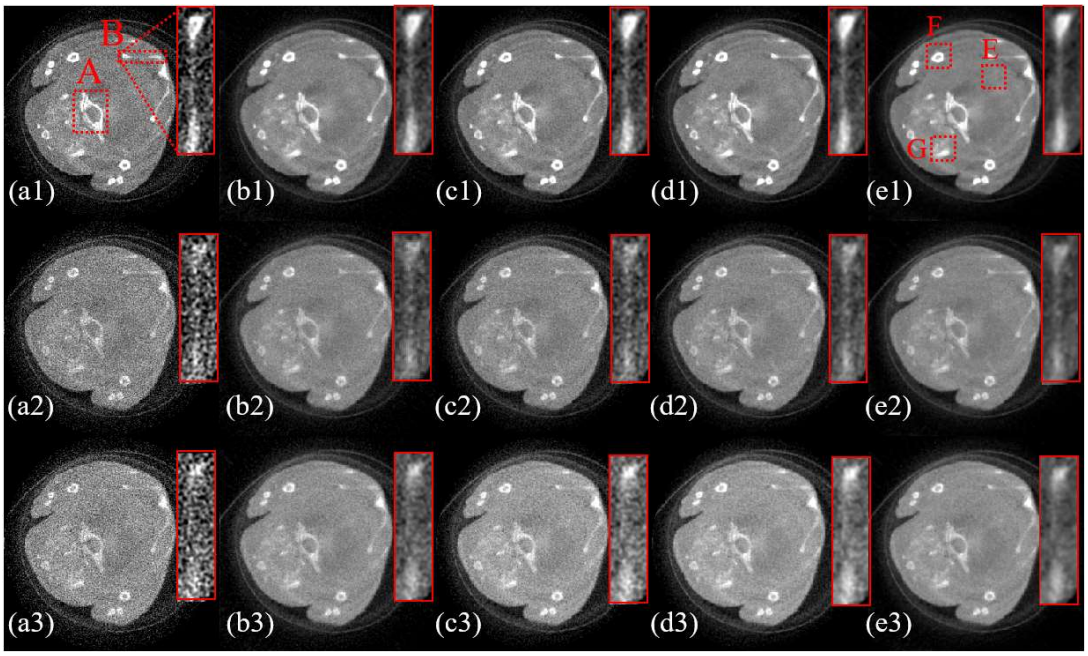

FIGURE 11. **Real mouse data reconstruction results of 270 projection views. Columns from left to right represent the images of SART, LRTV, FTNN, SISTER and LRTT method, respectively. Rows 1 to 3 are the reconstructions of channel 1, 8 and 13, respectively.**

Figure 10 depicts a thorough analysis of the convergence of the various algorithms. The LRTT algorithm's convergence curves under different projection views and noise levels are shown on the left. It is possible to conclude that the LRTT has superior convergence performance regardless of the noise level and the projection views. The convergence curves of the LRTT algorithm and the competing algorithms are displayed on the right. The LRTT and SISTER algorithms both exhibit excellent convergence, with the LRTT achieving the lowest RMSE value. All of these findings support the claim that the proposed approach is highly stable and resistant to noise.

C.REAL DATA EXPERIMENT

The real projection data were gained from a mouse injected with gold nanoparticles (GNPs) by MARS micro- CT as exhibited in [34]. The PCD system was performed at 120 kVp and 175 mA and the source had a minimum focal spot of 75 μm . The distances from the X-ray source to object and PCD

TABLE III THE COMPUTATIONAL COST OF METHODS

		SART	LRTV	FTNN	SISTER	LRTT
90Views	$I_0 = 5\times10^3$	0.1	9.0	18.1	41.0	9.8
	$I_0 = 1\times10^6$	0.1	8.4	18.2	48.7	9.9
180Views	$I_0 = 5\times10^3$	3.9	17.1	18.9	45.0	12.2
	$I_0 = 1\times10^6$	3.2	16.3	18.3	49.0	11.7

TABLE IV THE MEMORY USAGE OF METHODS

Physical Memory (MB)	SART	LRTV	FTNN	SISTER	LRTT
	15034	15564	17412	14505	12384

are 158 mm and 255 mm, respectively. The PCD system has 512 detector elements and each of them is of 0.11 mm. The spectrum of X-ray source is divided into 13 channels. The size of image is 512×512 for each channel. Figure 2 (c) demonstrates the FBP reconstruction image.

Figures 11 and 14 illustrate the results of reconstruction images of 270 and 180 projection views, respectively, using different methods. The SART method produced images that were heavily impaired by noise, making it difficult to distinguish finer bone structure and soft tissue. The LRTV algorithm managed to reduce the noise through

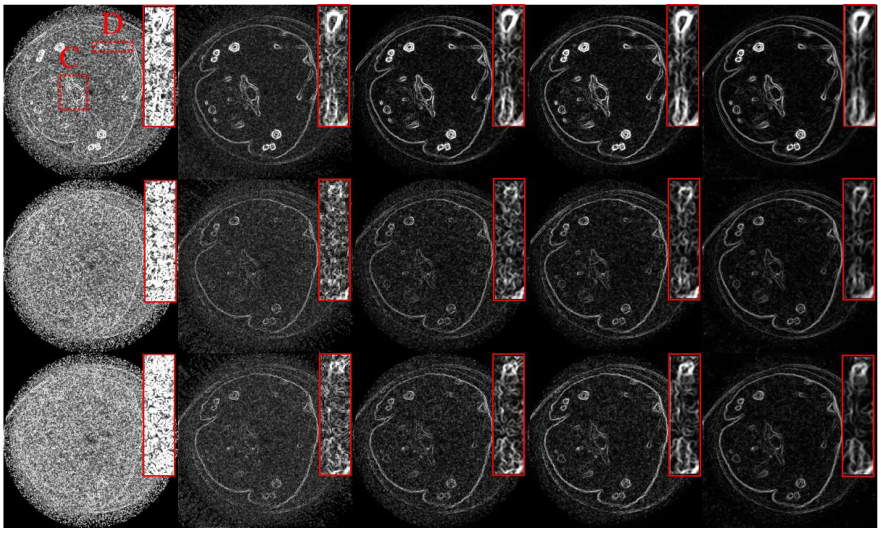

FIGURE 12. **Real data gradient image results of five algorithms. Columns from left to right show the images of SART, LRTV, FTNN, SISTER and LRTT method, respectively. Rows 1 to 3 are the reconstructions of channel 1, 8 and 13, respectively.**

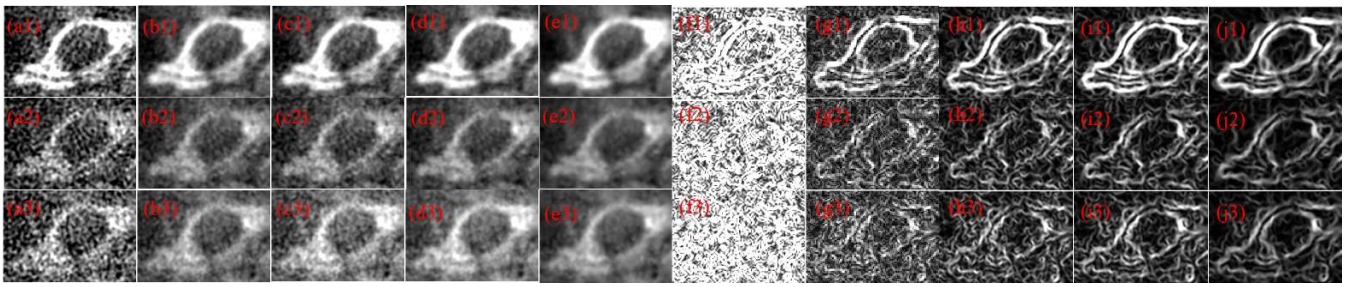

FIGURE 13. **Real data experiment ROI results of five algorithms. Columns a to e denote SART, LRTV, FTNN, SISTER and LRTT images, respectively. Rows f to j are the images and gradient images from channel 1, 8 and 13, respectively.**

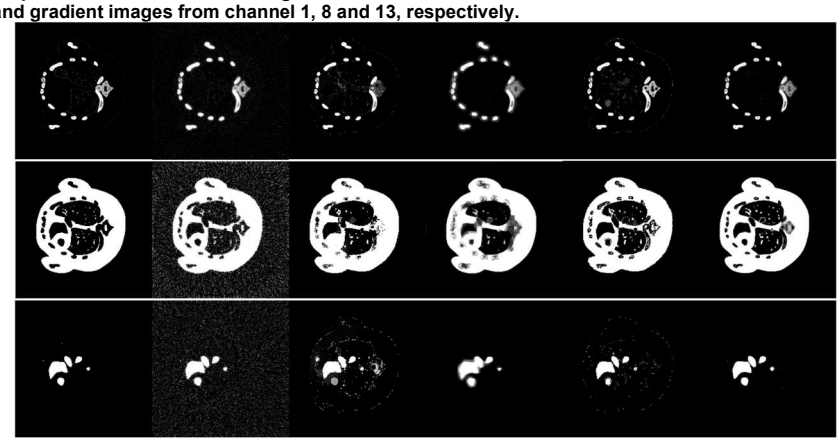

FIGURE 14. **Decomposition results based on the reconstruction images from simulation experiments by different algorithms. Rows 1 to 3 represent the basis material image of bone, tissue, and iodine, respectively. Columns from left to right denote the decomposition results based on the reference and the reconstructed images with the SART, LRTV, FTNN, SISTER and LRTT method, respectively. The display windows of rows 1 to 3 are [0.01 0.551], [0.05 0.65], and [0.25 0.85], respectively.**

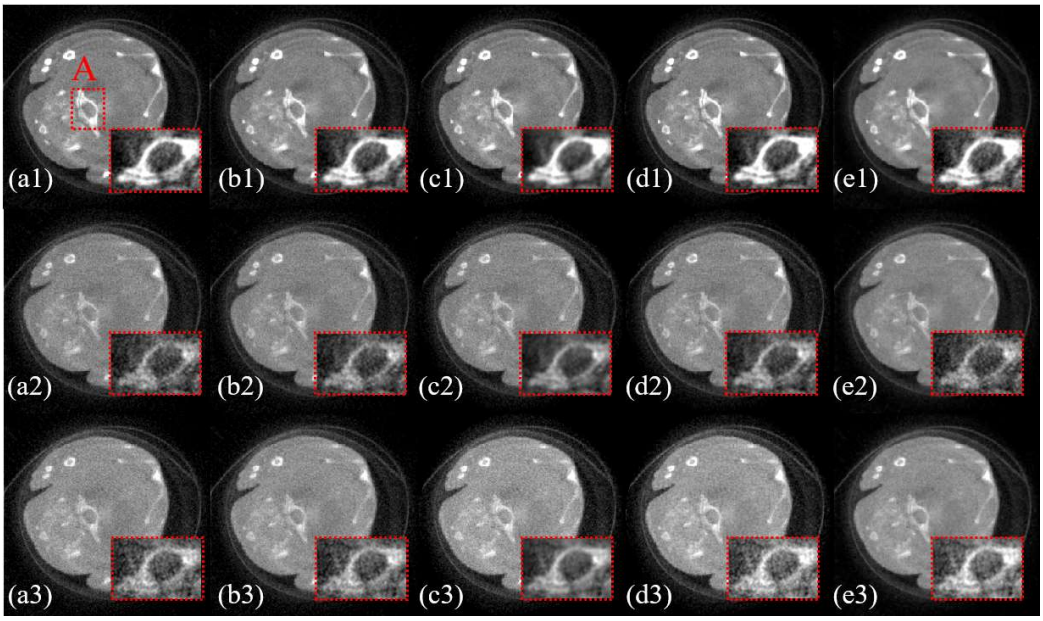

FIGURE 15. **Real data experiment results of 180 projection views. Columns from left to right denote the images of SART, LRTV, FTNN, SISTER and LRTT method, respectively. Rows 1 to 3 are the reconstructions of channel 1, 8 and 13, respectively.**

regularization, but there are still blocky artifacts and some structural details are still lost. The FTNN and SISTER algorithms, as seen in the third and fourth columns, had a significant advantage over the LRTV and SART algorithms in terms of details recovery and edges preservation. However, compared to the LRTT, the two methods still had a noticeable gap in restoring image features and preserving edges.

From the magnified ROIs in figures 11 and 14 and the gradient images in figure 12, it can be seen that it is challenging to differentiate the contour profile in "B" and "D" using SART, LRTV and FTNN, particularly in the high energy channels (e.g.,13^{th}). However, SISTER and LRTT can provide distinct bone contours. The bone structure in "B" and "D" reconstructed by SISTER is connected in low energy channels, but is fragmented in the high energy channels, while the counterpart reconstructed by LRTT could restore the complete bone structure.

To further illustrate the superiority of the LRTT, the ROIs with more intricate structures, namely "A" and "C", are extracted. Upon magnification, it is evident that the images obtained by the SART are heavily corrupted by noise, and the contour structure is barely visible. The LRTV and FTNN have improved the images, but the edges are blurred and hard to make out. Generally, the SISTER and LRTT offer better image quality, with the LRTT providing the most complete and clear image structures, as can be seen from the corresponding gradient images.

The ROI areas A and C depicted in figures 11 and 12 are magnified in figure 13. Upon closer inspection of the magnified ROI images, it is evident that the LRTT algorithm is highly effective in preserving the image edge information and detail structure. For quantitative evaluation of the real data experiments, we have presented mean values and standard deviations, as shown in table II. The mean values address the accuracy of the reconstruction results, while the standard deviations measure the performance of noise suppression for different algorithms. To calculate the evaluation index, we selected textured regions (A, B, G) and smooth regions (E, F) from figure 11 (a1 and e1). SART had the worst noise suppression effect. Compared to FTNN and SISTER, LRTV achieved smaller STDs, but lost detailed structures. From the quantitative evaluation, the LRTT gained the promising results in terms of reconstruction results and noise suppression.

The decomposition results of spectral CT images in the simulation experiment are shown in figure 14. As can be seen from row 1 of figure 14, many soft tissue and iodine contrast agent pixels are incorrectly calssified by SART. LRTV algorithms are more likely to mistake iodine contrast agents for bone structures. Compared to SART and LRTV methods, the decomposition results obtained by FTNN and SISTER are improved, but there are still some structures that are relatively blurred. In contrast, the LRTT result can produce clearer bone and iodine contrast agent maps. Although the fine textures are still slightly blurred.

IV.DISCUSSION AND CONCLUSIONS

This study presented a novel image reconstruction approach based on the global correlation and spatial sparsity of the image, which was intended to address the serious noise and degradation of image quality of photon counting spectral CT

image. Numerical simulation and a real data analysis were employed to validate the efficiency of the designed method. The global correlation of the spectral CT image was investigated further in this study, and the tensor train rank was adopted to describe the global correlation. The relatively balanced expansion of the tensor train can better capure the high correlation between different channel images, effectively reduce image noise, and improve image quality. In order to reconstruct the image with distinct texture features, the energy channel spatial sparsity was also described using the L0 norm of the TV whilepreserving the structure details.

Tensor train rank offers a superior ability to capture hidden correlations among different energy channels of spectral CT images, thanks to its well-balanced matricization strategy. Unlike other tensor ranks like Tucker rank, which cannot appropriately capture the global correlation of a tensor, tensor train rank provides a mean of the correlation between a few modes and the rest modes, while Tucker rank provides the mean of the correlation between a single mode and the rest modes. Therefore, it is a more robust tensor decomposition technique for spectral CT reconstruction.

The tensor train decomposition contrubutes to the reconstruction of spectral CT images by exploring the global correlation between multi-channel images and realizing communication between high and low energy channel information. In other words, by using high-quality channel images to improve the detail retention capabilities of low-quality channel images, the technique improved the expressiveness of the model to spectrum CT image.

In this study, the tensor train decomposition is carried out directly, which reduced the calculation complexity and time consumptio, as shown in the table III and IV, compared with the least square iterative operation of CP decomposition used in the SISTER. This advantage over the competing methods was obtained while maintaining superior performance. The patch extraction operation employed in the SISTER took much longer than the k-mean++clustering in the LRTT. Intuitively, the LRTT algorithm is not significantly superior to the SISTER, but it could be concluded from Table III that the calculation efficiency of the LRTT has been greatly improved, and the time consumption has been reduced by about 75%.

In this paper, a spectral CT reconstruction method based on the low rank tensor train was proposed. Based on the high correlation among channels and the sparsity of spectral CT image, two regularization terms were developed. To begin, k-means++clustering was employed to extract similar blocks from various energy channel and form low rank tensors. Because of the benefit of maintaining image detail structure, the L0-norm of TV was utilized. In additional, the effective ADMM algorithm was adopted to optimize the model. Finally, numerical simulations and real data experiments were conducted to validate the effectiveness of the LRTT and the competing algorithms, demonstrating that the proposed method outperformed the contrast algorithms in terms of noise suppression and image detail structure preservation.

In spectral CT imaging, material decomposition is a crucial component of spectrum CT imaging, as it enables material discrimination and information extraction. While our proposed algorithm outperformed the competing methods in image reconstruction, the real data experiments showed slight noise, leading to suboptimal material decomposition results in our numerical simulation experiments, as illustrated in figure 14. Consequently, designing an efficient and reliable multi-material decomposition algorithm is a key focus for the next stage of our project.

As shown in table III and IV, our proposed LRTT algorithm requires less computation time and RAM due to its direct tensor train decomposition on low-rank tensor unit comparing with the competing algorithms. However, as demonstrated by the results of our numerical simulation and real data experiments, the reconstruction images of LRTT are not significantly better than those of the contrast algorithms. This may be due to the adoption of tensor train rank and L0-norm of gradient image, which only partially describe the global correlation and the non-local similarities of spectral CT. To address this limitation, we plan to explore the addition of neural network regularization in the next stage as in reference [51]-[54]. This approach will leverage the powerful feature extraction capabilities of neural network to more fully express the rich prior information of spectral CT.

APPENDIX A

For completeness, we decribe the concept of von Neumann entropy in [40] to describe the correlation between subspace and its complementary subspace. An N-order tensor $\mathcal{X} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ can be expanded into a matrix along the k-th dimension

$$X_{(k)} \in \mathbb{R}^{m \times n}, m = J_k, n = \prod_{l=1, l \neq k}^{N} J_l, \tag{A1}$$

$X_{(k)}$ represents a composite system and the two subspaces are $X_A \in \mathbb{R}^m, X_B \in \mathbb{R}^n$, respectively. The von Neumann entropy can be calculated by the following formula

$$S = -\sum_{i=1}^{r_k} \sigma_i^2 \log_2 \sigma_i^2, \tag{A2}$$

where r_k is the rank of $X_{(k)}$ and non-zero singular values σ_i fulfilling $\sum_{i=1}^{r_k} \sigma_i^2 = 1$. According to (A2), if the singular value decreases quickly, the entropy will be very small, indicating that the correlation between the two subsystems is, and vice versa. If $X_{(k)}$ is obtained by expanding a tensor using formulation (4), the singular value of $X_{(k)}$ will be realistically balanced and decreases slowly, resulting in a high entropy, indicating that the correlation between two sub-spaces is strong.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant 62271504, Grant 62201616, and Grant 62101596. This work was also supported by the National Key Research and Development Project of China under Grant 2020YFC1522002 and the China Postdoctoral Science Foundation under Grant 2019M663996.

REFERENCES

[1] P. M. Shikhaliev, T. Xu, and S. Molloi, "Photon Counting Computed Tomography: Concept and Initial Results," *Med Phys*, vol. 32, no. 2, pp. 427-436, Feb. 2005.

[2] R. Symons, A. Pourmorteza, V. Sandfort, M. A. Ahlman, T. Cropper, and M. Mallek, "Feasibility of Dose-Reduced Chest CT with Photon Counting Detectors: Initial Results in Humans," *Radiology*, vol. 285, no. 3, pp. 980-989, Mar. 2017.

[3] W. A. Kalender, D. Kolditz, C. Steiding, V. Ruth, F. Luck, and E. Wenkel, "Technical Feasibility Proof for High-Resolution Low-Dose Photon-Counting CT of the Breast," *European radiology*, vol. 27, no.3, pp. 1081-1086, Mar. 2017.

[4] S. Tilley, W. Zbijewski, and J. W. Stayman, "Model-Based Material Decomposition with a Penalized Nonlinear Least-Squares CT Reconstruction Algorithm," *Physics in Medicine and Biology*, vol. 64, no.3, Jan. 2019.

[5] W. Wu, P. Chen, S. Wang, F. Liu, and H. Yu, "Image-Domain Material Decomposition for Multi-Energy CT Using a Generalized Dictionary Learning," *IEEE Transactions on Radiation and Plasma Medical Science*, vol. 5, no. 4, pp. 537–547, Apr.2021.

[6] A. Graser, T.R.C. Johnson, H. Chandarana, and M. Macari, "Dual energy CT: Preliminary Observations and Potential Clinical Applications in the Abdomen," *Eur.Radiol.*, vol. 19, no.2, pp. 13–23, Aug, 2009.

[7] A. Sawatzky, Q. Xu. C. Schirra and M. Anastasio**,** "Proximal ADMM for Multi-Channel Image Reconstruction in Spectral X-ray CT," *IEEE Trans Med Imaging*, vol. 33, no. 8, pp. 1657-1668, Aug, 2014.

[8] X. Lu, Z. Lu, J. Yin, Y. Gao, X. Chen, and Q, Guo, "Effects of Radiation Dose Levels and Spectral Iterative Reconstruction Levels on the Accuracy of Iodine Quantification and Virtual Monochromatic CT Numbers in Dual-Layer Spectral Detector CT: an Iodine Phantom Study," *Quant Imaging Med Surg*., vol. 9, no. 2, pp. 188-200, Feb, 2019.

[9] W. Wuest, M. May, M. Saake, M. Brand, M. Uder, and M. Lell, "Low-Dose CT of the Paranasal Sinuses: Minimizing X-Ray Exposure with Spectral Shaping," *Eur. Radiol.*, vol. 26, no. 11, pp. 4155-4161, Feb, 2016.

[10] R. Aamir, A. Chernoglazov, C. J. Bateman, P. H. Butler, N.G. Anderson, and S.T. Bell, "MARS Spectral Molecular Imaging of Lamb Tissue: Data Collection and Image Analysis," *J. Instrum.*, vol. 9, no. 2, p. 02005, Feb, 2014.

[11] E. Sidky and X. Pan, "Image Reconstruction in Circular Cone Beam Computed Tomography by Constrained, Total-Variation Minimization," *Phys Med Bio.*, vol. 53, no. 17, pp. 4777-4807, Aug, 2008.

[12] S. Li, D. Zeng, J. Peng, Z. Bian, H. Zhang, and Q. Xie, "An Efficient Iterative Cerebral Perfusion CT Reconstruction via Low-Rank Tensor Decomposition with Spatial–Temporal Total Variation Regularization," *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 360-370, Feb. 2019.

[13] S. Niu, Y. Gao, Z. Bina, J. Huang, Z. Liang, and J. Ma, "Sparse-View X-Ray CT Reconstruction via Total Generalized Variation Regularization," *Physics in Medicine & Biology*, vol. 59, no. 12, p. 2997-3017, May, 2014.

[14] B. Zhao, H. Gao, H. Ding, and S. Molloi, "Tight-Frame Based Iterative Image Reconstruction for Spectral Breast CT," Medical physics, vol. 40, no.3, pp. 8217–8229, Feb, 2013.

[15] T. Zhang, H. Yu, Y. Xi, S. Wang and F. Liu, "Spectral CT Image-Domain Material Decomposition via Sparsity Residual Prior and Dictionary Learning," *IEEE Transactions on Instrumentation and Measurement*. Vol. 72, Art no.4501313. pp. 1-13, 2023.

[16] Y. Zhang, X. Mou, G. Wang, and H. Yu, "Tensor-Based Dictionary Learning for Spectral CT Reconstruction," *IEEE Trans Med Imaging*. vol. 36, no. 1, pp. 142-154, Jan, 2017.

[17] S. Wang, A. Cai, W. Wu, T. Zhang, F. Liu and H. Yu, "IMD-MTFC: Image-domain Material Decomposition via Material-image Tensor Factorization and Clustering for Spectral CT," *IEEE Transactions on Radiation and Plasma Medical Sciences*. Doi:10.1109/TRPMS. 2023.3234613.

[18] Y. Zhang, S. Morteza, and H. Yu, "Tensor Decomposition and Non-Local Means based Multi-Energy CT Image Denoising," *Journal of X-Ray Science and Technology*, vol. 27, no. 3, pp. 397-416, Feb, 2019.

[19] H. Gao, H. Yu, S. Osher, and Wang G, "Multi-Energy CT based on a Prior Rank, Intensity and Sparsity Model (PRISM)," *Inverse Probl.*, vol. 27, no. 11, Nov, 2011, Art. no. 115012.

[20] E. Candès, J. Romberg, and T. Tao., "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information," *IEEE Trans Inf Theory*, vol. 52, no. 2, pp. 489-509, 2006.

[21] W. Wu, Y. Zhang, Q. Wang, F. Liu, P. Chen, and H. Yu, "Low-Dose Spectral CT Reconstruction Using Image Gradient *L*0-Norm and Tensor Dictionary," *Appl Math Model*. vol. 63, pp. 538-557, Nov, 2018.

[22] S. Wang, W. Wu, J. Feng, F. Liu, and H. Yu, "Low-Dose Spectral CT Reconstruction Based on Image-Gradient L0-Norm and Adaptive Spectral PICCS," *Phys Med Biol*. vol. 65, no. 24, p. 245005. Dec, 2020.

[23] L. Li, Z. Chen, G. Wang, J. Chu, and H. Gao, "A Tensor PRISM Algorithm for Multi-Energy CT Reconstruction and Comparative Studies," *Journal of X-ray science and technology*, vol. 22, no. 2, pp. 147-163, Jan, 2014.

[24] X. Yu, A. Cai, L. Wang, Z. Zheng, Y. Wang, Z. Wang, et al, "Framelet Tensor Sparsity with Block Matching for Spectral CT Reconstruction," *Medical Physics*, vol. 49, no. 4, pp. 2486-2501, Feb, 2022.

[25] Y. Zhang, D. Zeng, Z. Bian, H. Lu, and J. Ma, "Weighted Tensor Low-Rankness and Learnable Analysis Sparse Representation Model for Texture Preserving Low-Dose CT Reconstruction," *IEEE Trans. on Compute. Imaging*, vol. 7, pp. 321-336, Nov, 2021.

[26] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *SIAM Rev*., vol. 51, no. 3, pp. 455–500, 2009.

[27] S. Wang, H. Yu, Y. Xi, C. Gong, W. Wu, and F. Liu, "Spectral-Image Decomposition with Energy-Fusion Sensing for Spectral CT Reconstruction," *IEEE Transactions on Instrumentation and Measurement*. vol. 70, pp. 1-11. May, 2021.

[28] N. Renard, S. Bourennane and J. Talon. "Denoising and Dimensionality Reduction Using Multilinear Tools for Hyperspectral Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp.138-142, Feb, 2008.

[29] R. Tucker. "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no,3, pp. 279-311, 1996.

[30] D. Hu, W. Wu, M. Xu, Y. Zhang, J. Liu, R. Ge, et al, "SISTER: Spectral-Image Similarity-Based Tensor with Enhanced-Sparsity Reconstruction for Sparse-View Multi-Energy CT," *IEEE Transactions on Computational Imaging*. Vol. 6, pp. 477-490, Nov, 2019.

[31] Q. Xie, Q. Zhao, D. Meng, and Z. Xu, "Kronecker-Basis-Representation Based Tensor Sparsity and Its Applications to Tensor Recovery," *IEEE transactions on pattern analysis and machine intelligence,* vol. 40, no. 8, pp. 1888-1902. Aug, 2018.

[32] D. Zeng, Q. Xie, W. Cao, J. Lin, H. Zhang, J. Ma, et al, "Low-Dose Dynamic Cerebral Perfusion CT Reconstruction via Kronecker-Basis-Representation Tensor Sparsity Regularization," IEEE Transactions on Medical Imaging, vol. 36, no. 12, pp. 2546-2556. Dec, 2017.

[33] S. Wang, W. Wu, A Cai, Y Xu, Vardhanabhuti V, F Liu, and H Yu, "Image-Spectral Decomposition Extended-Learning Assisted by Sparsity for Multi-Energy Computed Tomography Regularization," *Quant Imaging Med Surg*, 2022. Doi: 10.21037/QIMS-22-235.

[34] W. Xia, W. Wu, S. Niu, F. Liu, J. Zhou, H. Yu, et al, "Spectral CT Reconstruction—ASSIST: Aided by Self-Similarity in Image-Spectral Tensors," *IEEE Transactions on Computational Imaging*. vol. 5, no. 3, pp. 420-436, Mar, 2019.

[35] X. Chen, W. Xia, Y. Liu, H. Chen, J. Zhou, Z. Zha, et al, "FONT-SIR: Fourth-Order Nonlocal Tensor Decomposition Model for Spectral CT Image Reconstruction," *IEEE Trans Med Imaging*. vol. 41, no. 8, pp. 2144-2156, Aug, 2022.

[36] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm," *IEEE Trans Pattern Anal Mach Intell*., vol.42, no. 4, pp. 925-938, Apr, 2020.

[37] T.H. Oh, Y. Tai, J. C. Bazin, H. Kim, and I.S. Kweon, "Partial Sum Minimization of Singular Values in Robust PCA: Algorithm and Applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 4, pp. 744-758, Aug, 2015.

[38] T. Jiang, T. Huang, X. Zhao, and L. Deng, "Multi-Dimensional Imaging Data Recovery via Minimizing the Partial Sum of Tubal Nuclear Norm," *Journal of Computational and Applied Mathematics*, vol. 372, p. 112680, Jul, 2020.

[39] I. Oseledets, "Tensor-Train Decomposition，" *SIAM J. Sci. Comput*，vol. 33, no. 5, pp. 2295-2317, May, 2011.

[40] J. Bengua, H Phien, H. Tuan, and M. Do. "Efficient Tensor Completion for Color Image and Video Recovery: Low-Rank Tensor Train," *IEEE Transactions on Image Processing*. vol. 26, no. 5, pp. 2466-2479, May, 2017.

[41] C. Chen, Z. Wu, Z. Chen, Z. Zheng, and X. Zhang, "Auto-Weighted Robust Low-Rank Tensor Completion via Tensor-Train," *Information Sciences*. vol. 567. pp.100-115, Aug, 2021.

[42] J. Yang, X. Zhao, T. Ji, T. Ma, and T. Huang, "Low-Rank Tensor Train for Tensor Robust Principal Component Analysis," *Applied Mathematics and Computation*. vol. 367, p. 124783, Feb, 2020.

[43] M. Ding, T. Huang, X. Zhao, M. Ng, and T. Ma, "Tensor Train Rank Minimization with Nonlocal Self-Similarity for Tensor Completion," *Inverse Probl. & Imag.*, vol. 15, no. 3, pp. 475–498, Jun, 2021.

[44] J. F. Cai, E. J. Candès, and Z. Shen, "A Singular Value Thresholding Algorithm for Matrix Completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956-1982, Apr, 2008.

[45] C. Chen, H. Qian, W. Chen, Z. Zheng, and H. Zhu, "Auto-Weighted Multi-View Constrained Spectral Clustering," *Neurocomputing,* vol. 366, pp. 1–11, Nov, 2019.

[46] Z. Zhang and S. Aeron. "Exact Tensor Completion Using T-SVD," *IEEE Trans. Signal Process*, vol. 65, no. 6, pp. 1511-1526, Mar, 2017.

[47] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor Completion for Estimating Missing Values in Visual Data," *IEEE Trans. Pattern Anal*, vol. 35, no. 1, pp. 208-220, Jan, 2013.

[48] E. A. Zimichev, N. L. Kazanskii, P. G. Serafimovich, "Spectral-Spatial Classification with K-Means++ Particional Clustering," *Computer Optics*, vol. 38, no. 2, pp. 281-286. 2014.

[49] Powell MJD. *A method for nonlinear constraints in minimization problems*. London, New York: Academic Press, 1969.

[50] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image Smoothing via L0 Gradient Minimization," *ACM Trans. Graph*., vol. 30, no. 6, pp. 1-12, Dec, 2011.

[51] Wu W, Hu D, Niu C, Lieza Vanden, Anthony P.H. Bulter, et al, "Deep learning based spectral CT imaging," *Neurea Network*, vol. 144, pp. 342-358, 2021.

[52] Wu W, Hu D, Niu C, Yu H, V Vandhannbhuti, Wang G, "DRONE: dual-domain residual-based optimization network for sparse-view CT reconstruction," *IEEE Transactions on Medical Imaging,* vol. 40, no.11, pp. 3002-3014, 2021.

[53] Wu W, X Guo, Y Chen, S Wang and J Chen. "Deep embedding-attention-refinement for sparse-view CT reconstruction," *IEEE Transactions on Instrumentation and Measurement,* vol. 72, pp. 1-11, 2023.

[54] Wei Zhang, Z Zhou, Zhi Gao, G Yang, L Xu, and Wu W. "Multiple adversarial learning-based angiography reconstruction for ultra-low-dose contrast medium CT," *IEEE Journal of Biomedical and Health Informatics,* vol. 27, no.1, pp. 409-420, 2022.

Jie Guo received the M.S. degree from the Central China Normal University, and she is currently pursuing the Ph.D. degree from the College of Information Engineering, Information Engineering University, Zhengzhou, China. She is currently an associated professor of the College of Information Engineering. Her research interests including theory and application of imaging, especially on intelligent algorithms for inverse problems, spectral computed tomography, and its applications on biomedical and industrial fields.

Xiaohuan Yu received the M.S. degree from the Zhengzhou University, and she is currently pursuing the Ph.D. degree from the College of Information Engineering, Information Engineering University, Zhengzhou, China. Her research interests including theory and application of imaging, especially on intelligent algorithms for inverse problems, image processing, spectral computed tomography, and its applications on biomedical and industrial fields.

Shaoyu Wang was born in Nanyang, Henan, China, in 1992. He received the M.S. and Ph.D. degree from Chongqing University, Chongqing, in 2018 and 2021. From 2019 to 2020, he was a joint-training Ph.D. Candidate with the University of Massachusetts Lowell. He is currently a lecturer with the College Information Engineering. His research interests include X-ray system design, image reconstruction, and material decomposition. He also achieved the Chinese Scholarship Council Scholarship, in 2019.

Ailong Cai received the B.S., M.S., and Ph.D. degrees from the College of Information Engineering, Information Engineering University, Zhengzhou, China, in 2011, 2014, and 2018, respectively. He is currently an associated professor with the College Information Engineering and the Key Laboratory of Imaging and Intelligent Processing of Henan Province. His research interests including theory and application of imaging, especially on intelligent algorithms for inverse problems, image processing, spectral computed tomography, and its applications on biomedical and industrial fields.

Zhizhong Zheng received B.S., M.S., and Ph.D. degrees from Information Engineering University, Zhengzhou, China, in 2002, 2005, and 2019, respectively. He is currently an associate professor at the College of Information Engineering. His research interests include the theory and technology of image reconstruction, especially on optimal transportation, optimization methods for inverse problems, and its applications in spectral computed tomography.

Ningning Liang was born in Tianjin, China, in 1982. She received the B.S., M.S., and Ph.D. degrees from the Information Engineering University, Zhengzhou, China, in 2004, 2009, and 2015, respectively. She is currently an associated professor with the Information Engineering University and the deputy director of the Key Laboratory of Imaging and Intelligent Processing of Henan Province. Her research interests include X-ray CT system design and intelligent image processing.

Lei Li received the B.S., M.S., and Ph.D. degrees from the College of Information Engineering, Information Engineering University, Zhengzhou, China, in 2003, 2006, and 2016, respectively. He is currently an associated professor of the College of Information Engineering and the Key Laboratory of Imaging and Intelligent Processing of Henan Province. His research interests including theory, technology and application of imaging, especially on X-ray system design, spectral computed tomography, and its applications on biomedical and industrial fields.

Bin Yan graduated from the Chinese Academy of Sciences in 2005 with a physics Ph.D. degree. He is currently a professor with the College Information Engineering, Information Engineering University. His research interests including theory and application of computed tomography imaging, computer vision, and brain computer interaction.