HHS Public Access

Author manuscript

Inverse Probl. Author manuscript; available in PMC 2019 October 01.

Published in final edited form as:

Inverse Probl. 2018 October ; 34(10): . doi:10.1088/1361-6420/aad67b.

Spatial-Spectral Cube Matching Frame for Spectral CT Reconstruction

Weiwen Wu[1,2], **Yanbo Zhang**[2], **Qian Wang**[2], **Fenglin Liu**[1,3,*], **Fulin Luo**[4], and **Hengyong Yu**[2,*]

[1]Key Lab of Optoelectronic Technology and Systems, Ministry of Education, Chongqing University, Chongqing 400044, China

[2]Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell, MA 01854, USA

[3]Engineering Research Center of Industrial Computed Tomography Nondestructive Testing, Ministry of Education, Chongqing University, Chongqing 400044, China

[4]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

Abstract

Spectral computed tomography (CT) reconstructs the same scanned object from projections of multiple narrow energy windows, and it can be used for material identification and decomposition. However, the multi-energy projection dataset has a lower signal-noise-ratio (SNR), resulting in poor reconstructed image quality. To address this thorny problem, we develop a spectral CT reconstruction method, namely spatial-spectral cube matching frame (SSCMF). This method is inspired by the following three facts: *i*) human body usually consists of two or three basic materials implying that the reconstructed spectral images have a strong sparsity; *ii*) the same basic material component in a single channel image has similar intensity and structures in local regions. Different material components within the same energy channel share similar structural information; *iii*) multi-energy projection datasets are collected from the subject by using different narrow energy windows, which means images reconstructed from different energy-channels share similar structures. To explore those information, we first establish a tensor cube matching frame (CMF) for a BM4D denoising procedure. Then, as a new regularizer, the CMF is introduced into a basic spectral CT reconstruction model, generating the SSCMF method. Because the SSCMF model contains an L_0-norm minimization of 4D transform coefficients, an effective strategy is employed for optimization. Both numerical simulations and realistic preclinical mouse studies are performed. The results show that the SSCMF method outperforms the state-of-the-art algorithms, including the simultaneous algebraic reconstruction technique, total variation minimization, total variation plus low rank, and tensor dictionary learning.

Keywords

spectral computed tomography; image reconstruction; image similarity; cube matching frame

*Corresponding author: Fenglin Liu and Hengyong Yu liufl@cqu.edu.cn and hengyong-yu@ieee.org.

1. Introduction

The spectral computed tomography (CT) (also called spectroscopic CT, multi-energy or color CT) has attracted continuous attentions because it has a great potential for tissue characterization, lesion detection and material decomposition [1, 2]. For example, the dual-energy CT (DECT), as the simplest and commercially available clinical spectral CT, has been employed for material differentiation [3], urinary calculi characterization [4], *etc.* Recently, thanks to the emergence and development of photon-counting detectors (PCDs), the multi-energy CT moves one step forward in practical applications. For example, the MARS (Medipix All Resolution System) research group developed a MARS-CT-3 scanner using three pairs of PCDs [5]. However, the charge sharing, Kescape, and fluorescence x-ray emission and re-absorption lead to corrupted spectral CT projection datasets with complicated noises [6]. This can decrease the signal-noise-ratio (SNR) of the projections and further comprise the reconstructed image quality. Thus, a primary task of spectral CT is to reconstruct accurate and clean images from the noisy projections. To address this intractable problem, lots of spectral CT reconstruction algorithms have been developed. In our opinion, all the methods employ some prior knowledge, and the knowledge can be divided into two categories: empirical-knowledge and prior-image-knowledge.

Regarding the empirical-knowledge based methods, they first convert the image into transform domain with a unified and image-independent transformation, and then focus on minimizing an L_0-norm or L_1-norm of the transform coefficients. Specifically, a satisfied image can be obtained by introducing man-made image characteristics as prior knowledge into the basic reconstruction model and optimizing the whole objection function. To address different problems in practical applications, different empirical-knowledge were reported, such as total variation (TV) [7], PRISM (prior rank, intensity and sparsity model) method [8], tensor-based nuclear norm [9], piecewise linear tight frame transform [10], total nuclear variation (TVN) [11], *etc.* Indeed, these reconstruction algorithms can obtain a satisfied result by enforcing the sparsity over a reconstructed image. Furtherly, the patch-based low-rank [1] , tensor nuclear norm (TNN) with TV [12], structure tensor TV [13] methods consider energy channel images similarity to formulate reconstruction model. However, image similarities in non-local spatial space are usually ignored among these methods. It can result in blurred edges, and finer structures are lost in the reconstructed images.

A representative group of prior-image-knowledge based methods are achieved by constructing a redundant dictionary from high quality images. This group can explore both image sparsity and similarity with a dictionary. They can be divided into two steps: dictionary training and image reconstruction. The first step is to train a redundant dictionary using images reconstructed by an analytic (e.g. FBP) or iterative (e.g. SART) algorithm. The second step is to reconstruct spectral images using the trained dictionary. In 2012, Zhao *et al.* applied a dual-dictionary learning (DDL) method for sparse-view spectral CT reconstruction [14]. Considering the image similarity among different energy channels, Zhang *et al.* introduced a tensor dictionary learning (TDL) into the spectral CT reconstruction [15]. To enhance the capability for clinical applications, very recently, Wu *et al.* incorporated an image gradient L_0-norm into the TDL (L_0TDL) framework for low-dose spectral CT [2]. Note that the final results depend on the image quality in the training

process. The higher the quality of training image is, the better the reconstructed image is. Another group prior-image-knowledge based methods are realized by introducing a high quality spectral image into the optimized reconstruction model to constrain solution, such as spectral prior image constrained compressed sensing algorithm (spectral PICCS)[16], TV-TV and total variation spectral mean (TV-SM) methods [17]. In general, both of the two groups need high quality prior images. However, it might be impossible to obtain high quality prior images in practice. In addition, this type of reconstruction method can not consider image similarity within non-local spatial space.

About ten years ago, the Block Matching 3D (BM3D) was proposed for 2D image denoising by considering small image patch similarity within a searching window [18]. Then, it also obtained a great success in image deblurring [19], inpainting [20] and compression [21], *etc*. To address the volumetric image denoising issue, the BM3D was extended to BM4D and then it was applied to limited-angle MRI reconstruction [22]. Because the BM4D can accurately explore the nonlocal similarity of image cubes, it was applied to multispectral [23] and hyperspectral denoising [24] with a great success.

To handle the aforementioned spectral CT reconstruction issues, i.e., image edge blurring and finer structure losing in the empirical-knowledge based methods and difficulty in obtaining high quality prior images in the prior-image-knowledge based methods, we propose a spatial-spectral cube matching frame (SSCMF) to explore the spatial-spectral space similarity for spectral CT. Compared with the previous other tensor-based or multi-channel-based methods [15, 25, 26], the proposed SSCMF not only considers image structure similarity across spectral channels but also explores the non-local similarity of spectral image itself by clustering up a series of similar small cubes. The contributions are threefold. First, we establish a tensor cube matching frame based on the BM4D denoising procedure. Its insight can be easily considered as a regularizer. We thus call it cube matching frame (CMF) for convenience. Such a regularization term not only encodes the sparsity insights delivered by the traditional sparsity constraint, but also explores the similarities implied in a 3rd tensor. Second, we sufficiently analyze the features of spectral images and introduce the CMF regularization term into the spectral CT reconstruction to further explore the image similarity and sparsity. This technique is named as spatial-spectral cube matching frame (SSCMF). Third, an effective denoising technique is employed to solve the 4D group coefficient L_0-norm minimization problem in the SSCMF model [27].

The rest of this paper is organized as follows. In section II, we establish the tensor CMF for BM4D. In section III, we construct the SSCMF model and adopt an effective denoising method for optimization. In section IV, both numerical simulations and real dataset experiments are performed to evaluate the proposed SSCMF method. In section V, we discuss some related issues and draw the conclusions.

2. BM4D Denoising Model

2.1. BM4D Algorithm

The BM4D was proposed as a volumetric nonlocal image denoising technique [22]. It can be considered as a natural extension of the state-of-the-art BM3D algorithm [19]. Similar to the BM3D algorithm, the BM4D algorithm can be implemented in three steps:

i) *Grouping similar volumetric cubes*. The volumetric image cubes within a given image window share similar structures and they can be clustered to one group. These cubes in the same group can be further stacked as a 4D data array. During the course of grouping the cubes, the noise level σ of the volumetric image significantly affects the grouping accuracy[18].

ii) *Filtering*. This consists of hard-thresholding filtering and collaborating filtering. The hard-thresholding filtering is used for image denoising by removing small groupwise spectrum coefficients in a 4D transform domain. An initial image can be recovered by performing the inverse transformation from 4D coefficients. The final denoised volumetric image can be obtained by the collaborating filtering.

iii) *Aggregation*. The final denoised image cubes will be rearranged into the original corresponding locations by adopting a general weighting method for redundancy. To illustrate the effectiveness of data redundancy, the matrix representation of BM3D was introduced in [19]. To explore the volumetric data denoising and realize spectral CT reconstruction, we first provide the tensor cube frame for the BM4D

2.2. Tensor Cube Frame of BM4D

Let $\mathcal{X} \in \mathcal{R}^{N_W \times N_H \times N_D}$ be a 3rd order tensor and $\mathcal{x}_j \in \mathcal{R}^{n_1 \times n_2 \times n_3}$ ($n_1 \ll N_W$, $n_2 \ll N_H$, $n_3 \ll N_D$) be an extracted tensor cube, where $j = 1,2,\ldots,J$ and $J=(N_W-n_1+1)(N_H-n_2+1)\times(N_D-n_3+1)$. Assuming E_j represents the tensor cube extraction operation $\mathcal{x}_j$ can be expressed as

$$\mathcal{x}_j = E_j(\mathcal{X}). \quad \#(1)$$

For a tensor cube $\mathcal{x}_j$, we can determine a tensor searching window $\mathcal{W}_j \in \mathcal{R}^{N_1 \times N_2 \times N_3}$ ($n_1 < N_1 < N_W$, $n_2 < N_2 < N_H$, $n_3 < N_3 < N_D$). The process of determining a search window can considered as a special operation on $\mathcal{X}$. Let $\mathcal{Z}_j$ represents the operation of determining a search window,the searching window $\mathcal{W}_j$ can be expressed as

$$\mathcal{W}_j = \mathcal{Z}_j(\mathcal{X}). \quad \#(2)$$

We extract R tensor cubes $\mathcal{Y}_{j,r}(1 < r \leq R)$ in the searching window tensor $\mathcal{W}_j$. All the extracted sensor cubes are stacked to form a 4th order tensor as the j^{th} For simplicity, we assume that R is fixed.it assumed that,for each volume pixel in the searching window, there is atleast one cube that entersin the group and contains the volumetric pixel. $\mathcal{Y}_{j,r}$ can expressed as

$$\mathcal{Y}_{j,r} = \varepsilon_{j,r}\left(\mathcal{W}_j\right), \quad \#(3)$$

where $\varepsilon_{j,r}$ represents the r^{th} tensor cube extraction operation from the j^{th} searching window.infact,the BM4D modeling can be considered as a special form of the 4d decorrelating transform.it can be realized by a combination of separable 3D intrablock and 1D interblock transforms. Let $\mathbf{\Phi}(\cdot)$and $\mathbf{\Psi}(\cdot)$ be the 3D interblock and 1D intrablock transforms,repectively. The 3D transform coefficient $\mathcal{G}_{j,r}$ of $\mathcal{Y}_{j,r}$ can be written as

$$\mathcal{G}_{j,r} = \mathbf{\Phi}\left(\mathcal{Y}_{j,r}\right). \quad \#(4)$$

There, the 3D coefficients of the j^{th} group can be written as $\mathcal{G}_j = \left(\mathcal{G}_{j,1}, \mathcal{G}_{j,2}, \ldots, \mathcal{G}_{j,R}\right)$. We apply the 1D interblock transformation to $\mathcal{G}_j$ and the 4D spectrum of this group $\mathcal{w}_j$ can be written as

$$\begin{aligned} \mathcal{w}_j &= \mathbf{\Psi}\left(\mathcal{G}_j\right) = \mathbf{\Psi}\left(\mathcal{G}_{j,1}, \mathcal{G}_{j,2}, \ldots, \mathcal{G}_{j,R}\right) = \mathbf{\Psi}\left(\mathbf{\Phi}\left(\mathcal{Y}_{j,1}\right), \mathbf{\Phi}\left(\mathcal{Y}_{j,2}\right), \ldots, \mathbf{\Phi}\left(\mathcal{Y}_{j,R}\right)\right) \\ &= \mathbf{\Psi}\mathbf{\Phi}\left(\mathcal{Y}_{j,1}, \mathcal{Y}_{j,2}, \ldots \mathcal{Y}_{j,R}\right) = \mathbf{\Psi}\mathbf{\Phi}\left(\varepsilon_{j,1}\left(\mathcal{Z}_j(\mathcal{X})\right), \varepsilon_{j,2}\left(\mathcal{Z}_j(\mathcal{X})\right), \ldots, \varepsilon_{j,R}\left(\mathcal{Z}_j(\mathcal{X})\right)\right). \end{aligned} \quad \#(5)$$

Now, let $\mathcal{E}_j$ be $(\varepsilon_{j,1},\varepsilon_{j,2},\ldots,\varepsilon_{j,R})$. Eq. (5) can be written as

$$\mathcal{w}_j = \mathbf{\Psi}\mathbf{\Phi}\left[\left(\varepsilon_{j,1}, \varepsilon_{j,2}, \ldots, \varepsilon_{j,R}\right)\mathcal{Z}_j(\mathcal{X})\right] = \mathbf{\Psi}\mathbf{\Phi}\left(\varepsilon_j\left(\mathcal{Z}_j(\mathcal{X})\right)\right). \quad \#(6)$$

Here, in order to better understand the process of the BM4D, $\mathcal{\psi}$ is denoted as a 4D analysis operator and it equal to $\mathbf{\Psi}\mathbf{\Phi}$. Thus, Eq. (6) can be simplified as

$$\mathcal{w}_j = \mathcal{\psi}\left(\mathcal{E}_j\left(\mathcal{Z}_j(\mathcal{X})\right)\right). \quad \#(7)$$

Note that Eq. (6) only considers a group constructed from the j^{th} searching window. A set of all groupwise coefficients $\mathbf{\Omega}$ of tensor $\mathcal{X}$ can be expressed as following

$$\boldsymbol{\Omega}(\boldsymbol{\mathcal{X}}) = \begin{pmatrix} \boldsymbol{w}_1 \\ \vdots \\ \boldsymbol{w}_J \end{pmatrix}. \quad \#(8)$$

The synthesis transform of BM4D model can be similarly derived First, the inverse intrablock 1D transform and 3D transforms are applied to each tensor group spectrum $\boldsymbol{w}_j$. Then the obtained tensor cubes estimation are rearranged into their corresponding original position. The estimated tensor group from the j^{th} group spectrum can be denoted as $\widehat{\mathcal{Z}}_j(\boldsymbol{\mathcal{X}}) = \boldsymbol{\mathcal{E}}_j^{-1}\boldsymbol{\vartheta}\left(\boldsymbol{w}_j\right)$, where $\boldsymbol{\vartheta}$ is the BM4D synthesis operation and $\boldsymbol{\mathcal{E}}_j^{-1}$ represents the inverse of extraction operation $\boldsymbol{\mathcal{E}}_j$ with appropriate weighting for redundancy. The final image can be further obtained by a weighting strategy i.e., the gropuwise estimation using weights $\alpha_j > 0$ therefore, tensor $\boldsymbol{\mathcal{X}}$ can be recovered as the following expression

$$\widehat{\boldsymbol{\mathcal{X}}} = \boldsymbol{\alpha}^{-1}\begin{pmatrix} \boldsymbol{\alpha}_1\boldsymbol{\mathcal{E}}_1^{-1}\boldsymbol{\vartheta}\left(\boldsymbol{w}_1\right) \\ \vdots \\ \boldsymbol{\alpha}_J\boldsymbol{\mathcal{E}}_J^{-1}\boldsymbol{\vartheta}\left(\boldsymbol{w}_J\right) \end{pmatrix}, \quad \#(9)$$

where

$$\boldsymbol{\alpha} = \sum_{j=1}^{J} \alpha_j \sum_{E_j(\boldsymbol{\mathcal{X}}) \in \boldsymbol{\mathcal{E}}\left(\mathcal{Z}_j(\boldsymbol{\mathcal{X}})\right)} \mathcal{Z}_j{}^T\mathcal{Z}_j$$

is a weighting tensor. if the u^{th} volume element of $\boldsymbol{\mathcal{X}}$ belongs to the j^{th} tensor cube, the value of corresponding location on $\mathcal{Z}_j{}^T\mathcal{Z}_j$ would be 1 and otherwise it woul be 0. Therefore, the $\sum_{E_j(\boldsymbol{\mathcal{X}}) \in \boldsymbol{\mathcal{E}}_{j,r}(\mathcal{Z}_j(\boldsymbol{\mathcal{X}}))} \mathcal{Z}_j{}^T\mathcal{Z}_j$ can indicate the number of cubes in the j^{th} group containing u^{th} tensor element.

2.3 Tensor frame Interpretation of BM4D

For a given group $\boldsymbol{\mathcal{E}}_j\left(\mathcal{Z}_j(\boldsymbol{\mathcal{X}})\right)$, the 4D analysis operator $\boldsymbol{\psi}$ can transform it into the coefficient $\boldsymbol{w}_j$. In contrast, the BM4D synthesis operator $\boldsymbol{\vartheta}$ can transform the group coefficient $\boldsymbol{w}_j$ into the corresponding tensor group $\widehat{\boldsymbol{\mathcal{E}}}_j\left(\mathcal{Z}_j(\boldsymbol{\mathcal{X}})\right)$. The synthesis operator $\boldsymbol{\vartheta}$ can be treated as an inverse transform of the analysis operartor $\boldsymbol{\psi}$, i.e., $\boldsymbol{\mathcal{E}}_j\left(\mathcal{Z}_j(\boldsymbol{\mathcal{X}})\right) = \boldsymbol{\vartheta}\left(\boldsymbol{\psi}\left(\boldsymbol{\mathcal{E}}_j\left(\mathcal{Z}_j(\boldsymbol{\mathcal{X}})\right)\right)\right)$. Again, the product $\boldsymbol{\vartheta}$ and $\boldsymbol{\psi}$ can be expressed as

$$\boldsymbol{I} = \boldsymbol{\vartheta}\boldsymbol{\psi}, \quad \#(10)$$

where $\boldsymbol{I}$ stands for the identify transform. However,an independent tensor group $\mathscr{E}_j\left(\mathscr{Z}_j(\mathscr{X})\right)$ is considered in the above analysis,and our goal is to recover the whole tensor $\mathscr{X}$ from overlapped groups $\left(\mathscr{E}_1\left(\mathscr{Z}_1(\mathscr{X})\right),\ldots,\mathscr{E}_J\left(\mathscr{Z}_J(\mathscr{X})\right)\right)^T$. Considering the weight coefficient tensor $\boldsymbol{\alpha}$, then the tensor $\mathscr{X}$ can be further recovered

$$\widehat{\mathscr{X}}=\boldsymbol{\alpha}^{-1}\begin{pmatrix}\boldsymbol{\alpha}_1\mathscr{E}_1^{-1}\vartheta_{\boldsymbol{\psi}}\mathscr{E}_1\left(\mathscr{Z}_1(\mathscr{X})\right)\\ \vdots\\ \boldsymbol{\alpha}_J\mathscr{E}_J^{-1}\vartheta_{\boldsymbol{\psi}}\mathscr{E}_J\left(\mathscr{Z}_J(\mathscr{X})\right)\end{pmatrix},\quad \#(11)$$

which is in consisitent with Eq. (9).it is necessary to point out that the weight coefficient tensor $\boldsymbol{\alpha}$ can be changed with respect to the image noise level σ. This means different noise may result im dofferent weight coefficient tensors.

3. Spectral CT Reconstruction based on spatial-spectral CMF

3.1. Spectral CT Imaging Model

Considering the noise in the projection datasets, the acquired or observed data forward model of the conventional CT for fan-beam scanning geometry van be approximated as a discretized linear system

$$\boldsymbol{p}=\mathscr{A}\boldsymbol{f}+\boldsymbol{\eta},\quad \#(12)$$

where $\boldsymbol{f}\in\mathscr{R}^N$ ($N=N_w\times N_H$) represents the vectorized 2D image, $\boldsymbol{p}\in\mathscr{R}^T$ ($T=T_1\times T_2$) represents the measurements, T_1 and T_2 are respectively the view and detector numbers, $\mathscr{A}\in\mathscr{R}^{N\times T}$ is the CT projection matrix and $\boldsymbol{\eta}\in\mathscr{R}^T$ stands for projection noise. Because the matrix $\mathscr{A}$ is huge, Eq. (12) can not be solved directly by adopting a matrix inverse technique due to the memory limitation Alternatively, it can be iteratively solved in terms of the following minimization problem

$$\underset{f}{\operatorname{argmin}}\quad \frac{1}{2}\|\boldsymbol{p}-\mathscr{A}\boldsymbol{f}\|_2^2,\quad \#(13)$$

where $\|.\|$ represents the L_2 norm Eq. (13) involves the linear programming problem,which can be updated with the ART or SART methods. Methods [28] To further optimize the result of Eq. (13), a regularization of prior knowledge can be introduced into Eq. (13) Therefore, we have

$$\underset{f}{\operatorname{argmin}}\quad \frac{1}{2}\|\boldsymbol{p}-\mathscr{A}\boldsymbol{f}\|_2^2+\tau\boldsymbol{R}(\boldsymbol{f}).\quad \#(14)$$

Note that Eq. (14) consisits of two terms: data fidelity term $\frac{1}{2}\|\boldsymbol{p}-\mathscr{A}\boldsymbol{f}\|_2^2$ and regularization term $\boldsymbol{R}(f)$, and $\tau > 0$ is a parameter to balance the data fidelity and regularization terms. For the regularization term $\boldsymbol{R}(f)$,it can be chosen as TV [29,30] wavelet frame [31], image gradient $L_{0\text{-norm}}$ [32], low-rank [33],etc.

For the multi-energy CT, multiple projection datasets can be obtained with one scan. This means we can reconstruct multiple energy-dependent images of the same object using different projection datasets. Let u be the number of energy channels. Similar to Eq. (13), spectral CT reconstruction can be modeled as

$$\underset{\mathscr{F}}{\operatorname{argmin}} \quad \left(\sum_{s=1}^{S}\frac{1}{2}\|\boldsymbol{p}_s-\mathscr{A}\boldsymbol{f}_s\|_2^2\right)+\tau\boldsymbol{R}(\mathscr{F}), \quad \#(15)$$

where $\boldsymbol{f}_s \in \mathscr{R}^N$ ($s = 1,2,\cdots,S$) is a vectorized image of the s^{th} energy channel of a 3rd order tensor $\mathscr{F} \in \mathscr{R}^{N_W \times N_H \times S}$, and the projection $\boldsymbol{p}_s \in \mathscr{R}^T$ is a coloumn of $\boldsymbol{P} \in \mathscr{R}^{N \times S}$. The first term of Eq. (15), (i.e., $\sum_{s=1}^{S}\frac{1}{2}\|\boldsymbol{p}_s-\mathscr{A}\boldsymbol{f}_S\|_2^2$) can be considered as data fidelity term for the multi-energy CT and it is formulated with summation of data fidelity error from all energy channels.

3.2. Spatial-spectral CMF for Spectral CT

For the spectral CT, there are some facts: *i*) from the aspects of material decomposition, patients usually consist of two or three basic materials, i.e., soft tissue, bone and water. Thus, the constructed spectral images have a strong sparsity in material composition; *ii*) the same basic material components in a single channel have similar intensity and structures. Different material components within the same energy channel share similar structural information; *iii*) multi-energy projection datasets are collected from the same patient by using different energy windows. Although images reconstructed from different energy channels have different attenuation coefficients, they share the same physical structures; *iv*) charge sharing, K-escape, and fluorescence x-ray emission and re-absorption can cause severe noise in the reconstructed images. Considering the advantages of CMF in characterizing 3D image similarity and reducing noise, the CMF as a regularizer is introduced into the spectral CT reconstruction to generate a spatial-spectral CMF (SSCMF) reconstruction method. Fig. 1 demonstrates the process of grouping which is a key procedure in the SSCMF method. We extract similar tensor cubes within a given spatial-spectral searching window and then stack them as the j^{th} group. To clarify the procedure of SSCMF algorithm, Fig. 2 shows flowchart for the whole iteration. The proposed SSCMF method can be divided into two steps: image reconstruction and CMF constraint. For image reconstruction, we adopt the SART to perform back projection. In the CMF constraint step, first, the images reconstructed by the SART are grouped and clustered as 4D arrays. Then, the noises in these 4D data arrays can be suppressed by the hard-thresholding operation. This process can be considered as the 1st estimation. To enhance the denoising capability of the CMF, we perform the collaborative

filtering on the original signal and initial estimation, which can be called as the 2nd estimation.

3.3. SSCMF Model

Mathematically speaking, the SSCMF model can be constructed as follow

$$\underset{\mathcal{F}}{\operatorname{argmin}} \quad \left\{ \left(\sum_{s=1}^{S} \frac{1}{2} \| \boldsymbol{p}_s - \mathcal{A} \boldsymbol{f}_s \|_2^2 \right) + \tau \| \boldsymbol{\Omega}(\mathcal{F}) \|_0 \right\}, \quad \#(16)$$

where $\|.\|$ represents the L_0 norm regularization term of the spatial –spectral cube 4D transform coefficients. Since this model contains the quasi L_0-norm it makes Eq.(16) an NP-hard problem Here,we employ a strategy in [27] to minimize eq. (16) which can be divided into two steps:

$$\begin{cases} \mathcal{F}^{(k+1/2)} = \underset{\mathcal{F}}{\operatorname{argmin}} \quad \left(\sum_{s=1}^{5} \frac{1}{2} \| \boldsymbol{p}_s - \mathcal{A} \boldsymbol{f}_s \|_2^2 \right) + \frac{\lambda}{2} \| \mathcal{F} - \mathcal{F}^{(k)} \|_F^2, \\ \mathcal{F}^{(k+1)} = \underset{\mathcal{F}}{\operatorname{argmin}} \quad \| \mathcal{F} - \mathcal{F}^{(k+1/2)} \|_F^2 + \tau \| \boldsymbol{\Omega}(\mathcal{F}) \|_0, \end{cases} \quad \#(17a)\#(17b)$$

where λ is acoupling parameter and $\|.\|_F$ represents the tensor frobenius norm Eqs. (17a) and (17b) can be called as observation fidelity and model fidelity steps, respectively. Especially, different forward system matrix $\mathcal{A}$ means different imaging problems. For example, when $\mathcal{A}$ means different imaging problems. For example ,when A corresponds a bluring matrix, Eq. (17a) is to update the current intermediate tensor image $\mathcal{F}^{(k+1/2)}$ so that it can be better agree with observation matrix $\boldsymbol{P}$ in the image debluring problem [27]. The regularization parameter λ is introduced into Eq(17a) to determine the degree of agreement between $\mathcal{F}^{(k+1/2)}$ and $\boldsymbol{P}$. Eq. (17b) achieves an accurate update which can conform with the prior image model or structure as enforced by the penalty function $\|\boldsymbol{\Omega}(\mathcal{F})\|_0$. For the spectral data fidelity term, because the update of each energy channel is independent, Eq (17a) can be rewritten as

$$\mathcal{F}^{\left(k+\frac{1}{2}\right)} = \underset{\mathcal{F}}{\operatorname{argmin}} \quad \sum_{s=1}^{S} \left(\frac{1}{2} \| \boldsymbol{p}_s - \mathcal{A} \boldsymbol{f}_S \|_F^2 + \frac{\lambda}{2} \| \boldsymbol{f}_s - \boldsymbol{f}_s^{(k)} \|_2^2 \right). \quad \#(18)$$

The update of $\mathcal{F}^{\left(\boldsymbol{k}+\frac{\boldsymbol{1}}{\boldsymbol{2}}\right)}$ can be achieved by separetly updating all energy channel image $\boldsymbol{f}_s(s = 1,\ldots,S)$. Therefore, Eq. (18) is equivalent to

$$\boldsymbol{f}_s^{\left(k+\frac{1}{2}\right)} = \underset{\boldsymbol{f}_s}{\operatorname{argmin}} \quad \frac{1}{2}\|\boldsymbol{p}_s - \boldsymbol{\mathcal{A}}\boldsymbol{f}_s\|_2^2 + \frac{\lambda}{2}\|\boldsymbol{f}_s - \boldsymbol{f}_s^{(k)}\|_2^2, \quad (s = 1, \ldots, S). \quad \#(19)$$

To determine the minimization point of f_s, the derivate of Eq. (19) should be equivalent to zero. We have,

$$-\boldsymbol{\mathcal{A}}^T\left(\boldsymbol{p}_s - \boldsymbol{\mathcal{A}}\boldsymbol{f}_s\right) + \lambda\left(\boldsymbol{f}_s - \boldsymbol{f}_s^{(k)}\right) = 0, \quad (s = 1, \ldots, S). \quad \#(20)$$

which can be simplified as

$$\left(\boldsymbol{\mathcal{A}}^T\boldsymbol{\mathcal{A}} + \lambda\boldsymbol{I}\right)\boldsymbol{f}_s = \boldsymbol{\mathcal{A}}^T\boldsymbol{p}_S + \lambda\boldsymbol{f}_S^{(k)}. \quad (s = 1, \ldots, S) \quad \#(21)$$

Eq. (21) is equivalent to

$$\left(\boldsymbol{\mathcal{A}}^T\boldsymbol{\mathcal{A}} + \lambda\boldsymbol{I}\right)\boldsymbol{f}_s = \boldsymbol{\mathcal{A}}^T\boldsymbol{p}_s + \lambda\boldsymbol{f}_s^{(k)} + \boldsymbol{\mathcal{A}}^T\boldsymbol{\mathcal{A}}\boldsymbol{f}_s^{(k)} - \boldsymbol{\mathcal{A}}^T\boldsymbol{\mathcal{A}}\boldsymbol{f}_s^{(k)}, \quad (s = 1, \ldots, S) \quad \#(22)$$

$$\left(\boldsymbol{\mathcal{A}}^T\boldsymbol{\mathcal{A}} + \lambda\boldsymbol{I}\right)\boldsymbol{f}_s = \left(\boldsymbol{\mathcal{A}}^T\boldsymbol{\mathcal{A}} + \lambda\boldsymbol{I}\right)\boldsymbol{f}_s^{(k)} + \boldsymbol{\mathcal{A}}^T\left(\boldsymbol{p}_s - \boldsymbol{\mathcal{A}}\boldsymbol{f}_s^{(k)}\right). \quad (s = 1, \ldots, S) \quad \#(23)$$

Therefore, $\boldsymbol{f}_s^{\left(k+\frac{1}{2}\right)}$ can be updated as

$$\boldsymbol{f}_s^{\left(k+\frac{1}{2}\right)} = \boldsymbol{f}_s^{(k)} + \beta\left(\boldsymbol{\mathcal{A}}^T\boldsymbol{\mathcal{A}} + \lambda\boldsymbol{I}\right)^{-1}\boldsymbol{\mathcal{A}}^T\left(\boldsymbol{p}_s - \boldsymbol{\mathcal{A}}\boldsymbol{f}_s^{(k)}\right), \quad (s = 1, \ldots, S), \quad \#(24)$$

where $\beta = 1$ and it can be any other relaxation factor in the range of (0,2). Eq. (21) indicates that λ only affects the iteration step. Considering Eq (24) as the solution of Eq. (17a), $\lambda = 0$ becomes a legal regularization parameter for a unique solution. In this case, Eq. (24) is degraded as the SART method [34]. Thus, λ is set as zero in this work.

The first term $\|\boldsymbol{\mathcal{F}} - \boldsymbol{\mathcal{F}}^{(k+1/2)}\|_F^2$ of Eq. (17b) can be considered as the data fidelity to balance the final estimation $\boldsymbol{\mathcal{F}}^{(k+1)}$ and the current estimation $\boldsymbol{\mathcal{F}}^{(k+1/2)}$. The regularization term $\|\boldsymbol{\Omega}(\boldsymbol{\mathcal{F}})\|_0$ enforces the L_0-norm sparse constraint over the 4D tensor coefficients to suppress noise by removing small coefficients in the 4D transform domain.considering Eqs. (7) and (8), Eq. (17b) is equivalent to

$$\mathcal{F}^{(k+1)} = \underset{\mathcal{F}}{\operatorname{argmin}} \quad \left\| \mathcal{F} - \mathcal{F}^{\left(k+\frac{1}{2}\right)} \right\|_F^2 + \tau \left\| \psi \begin{pmatrix} \mathcal{E}_1\left(\mathcal{Z}_1(\mathcal{F})\right) \\ \vdots \\ \mathcal{E}_J\left(\mathcal{Z}_J(\mathcal{F})\right) \end{pmatrix} \right\|_0 . \quad \#(25)$$

Substituting Eq. (10) into Eq .(11) and then Eq. (25) becomes

$$\mathcal{F}^{(k+1)} = \underset{\mathcal{F}}{\operatorname{argmin}} \quad \left\| \mathcal{F} - \mathcal{F}^{\left(k+\frac{1}{2}\right)} \right\|_F^2 + \tau \left\| \begin{pmatrix} \psi \mathcal{E}_1 \mathcal{Z}_1 \\ \vdots \\ \psi \mathcal{E}_J \mathcal{Z}_J \end{pmatrix} \mathcal{F} \right\|_0 , \quad \#(26)$$

which can be further written as

$$\mathcal{F}^{(k+1)} = \underset{\mathcal{F}}{\operatorname{argmin}} \quad \left\| \mathcal{F} - \mathcal{F}^{\left(k+\frac{1}{2}\right)} \right\|_F^2 + \tau \| \varrho(\mathcal{F}) \|_0, \quad \#(27)$$

where $\varrho = \begin{pmatrix} \psi \mathcal{E}_1 \mathcal{Z}_1 \\ \vdots \\ \psi \mathcal{E}_J \mathcal{Z}_J \end{pmatrix}$. The problem of Eq (27) can be treated as a denoising with to solve such a problem with an analysis prior defined over the tensor frame shrinkage algorithm [36].In this work, we employ a typical technology (i.e., hard-thresholding method [36] to solve Eq. (27). The solution can be given as

$$\mathcal{F}^{(k+1)} = \alpha^{-1} \varrho^{-1} \left\lfloor \varrho \left(\mathcal{F}^{\left(k+\frac{1}{2}\right)} \right) \right\rfloor_\tau , \quad \#(28)$$

where ϱ^{-1} represents the inverse operator of the ϱ with appropriate weighting for redundancy. The operator $\lfloor \cdot \rfloor_\tau$ denotes the hard thresholding operation and it can be calculated as [36]

$$\left\lfloor \varrho \left(\mathcal{F}^{\left(k+\frac{1}{2}\right)} \right) \right\rfloor_\tau = \begin{cases} 0, & \varrho \left(\mathcal{F}^{\left(k+\frac{1}{2}\right)} \right) < \sqrt{\tau} \\ \varrho \left(\mathcal{F}^{\left(k+\frac{1}{2}\right)} \right), & \varrho \left(\mathcal{F}^{\left(k+\frac{1}{2}\right)} \right) \geq \sqrt{\tau} \end{cases} . \quad \#(29)$$

It is necessary to emphasize that Eq. (29) can be implemented by the BM4D denoising algorithm, which is formulated using the frame notation in [22]. For our proposed

reconstruction algorithm, the frame shrinkage step Eq. (29) is treated as an approximate solution for Eq. (17b). Therefore, Eq. (17b) will be updated by a BM4D-based denoising for spectral CT images. Our algorithm mainly explores the similarity of spatial-spectral information by utilizing the cube matching frame regularization and its workflow is summarized as Algorithm I.

Algorithm I

SSCMF

Input: $\lambda, \tau, \beta, \sigma, \boldsymbol{P}$;

1: Initialization of $\mathscr{F} \leftarrow \mathbf{0}$, $k = 0$

2: **repeat**

3: **for** s: = 1,2, ⋯,S **do**

4: Updating $\boldsymbol{f}_s^{(k+1/2)}$ by SART

5: **end for**

6: Generating frames $\boldsymbol{\gamma}$ and $\boldsymbol{\vartheta}$ using the $\mathscr{F}^{(k+1/2)}$;

7: Updating $\mathscr{F}^{(k+1)}$ using Eq. (28);

8: Applying positive constraint over $\mathscr{F}^{(k+1)}$;

9: $k = k + 1$;

10: **until** convergence

Output: Reconstructed spectral CT image $\mathscr{F}$

In the Algorithm I, the model fidelity 6–7 steps correspond to the BM4D denoising process of image $\boldsymbol{f}_S^{(k+1/2)}$. They are implemented by adopting a highly available online optimized BM4D denoising toolbox [22] (http://www.cs.tut.fi/~foi/GCF-BM3D/). Eq. (28) is implemented with a 4D transform shrinkage followed by the wiener collaborative filtering.Here, the coefficient shrinkage Eq. (29) is implemented by a hard thresholding filtering with the controlling parameter τ in our numerical simulations and preclinical experiments ,the cube size is $4 \times 4 \times 4$ and the sliding step is 3. Other parameter setting and transform selection are the same as those in the BM4D denosing toolbox (http://www.cs.tut.fi/~foi/GCF-BM3D/).

4. Experiments

To evaluate the performance of our SSCMF method, simultaneous algebratic reconctruciton technique (SART), total variation minimization (TVM), total variation plus low rank (TV +LR)[8], and tensor dictionary learning (TDL)[15] algorithms are selected and implemented for comparison. Both numerical simulations and preclinical spectral datasets are employed for stopped after 50 iterations. The relaxation factor β is fixed as 0.03 and other parameters are evaluation. For all the iterative methods, the initial images are set as zero and all of them are optimized and summarized as in Table 1.

4.1. Mouse Thorax Phantom Study

A simulated mouse thorax phantom is employed, containing three basic materials: soft tissue, bone, and iodine (Fig. 3). As shown in Fig. 3, 1.2% iodine is injected into the mouse thorax phantom as contrast agent. A polychromatic 50KVp x-ray source is assumed, and it is divided into eight different energy channels: [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. The PCD used in the experiment contains 512 detector elements, and each element covers a length of $0.1mm$. Regarding the scanning geometry, the distances from source to rotation center and PCD are $132mm$ and $180mm$, respectively. 640 projections are uniformly collected over a full scan. Each reconstructed energy channel image consists of 512×512 pixels, and each pixel covers an area of 0.075×0.075 mm^2. The photon number of each x-ray path is set as 2×10^4, and Poisson noise is added in to the received photon number.

a) Reconstruction Results—To compare the performances of all algorithms in spectral CT reconstruction, parameters in all algorithms are optimized, and their best results are selected for comparison and analysis. Fig. 4 shows the reconstructed images of three representative energy channels (1^{st}, 4^{th} and 8^{th}). One can see that the SSCMF algorithm always obtain higher quality images compared with other competitors. More specifically, the SSCMF method outperforms other algorithms in edge preservation and finer structure recovery. This point is further confirmed by two magnified region of interests (ROIs) A (bony structure) and B (soft tissue) in Fig. 5. It can be observed that images reconstructed by using the SART have the strongest noise. The TVM, TV+LR and TDL methods can reduce noise to a certain extent, but the profiles of reconstructed images are not clearer. In contrast, the SSCMF method has an excellent performance in noise reduction as well as image edge preservation. Fig. 6 plots representative profiles along the yellow line in bony structure A in Fig. 5. From Fig. 6, one can infer that the profiles of SSCMF method are closer to the ground truths compared with the TVM, TV+LR and TDL algorithms.

To further demonstrate the advantages of the SSCMF method, Fig. 7 shows the reconstructed difference images with respect to the references. Note that the reference images are reconstructed by the SART from noise-free projections. It can be seen from Fig. 7 that the reconstructed images utilizing the SART have the largest differences, followed by the TVM, TV+LR and TDL. In contrast, the difference images of SSCMF are smaller than those reconstructed using other competing algorithms, especially for the 1^{th} energy channel.

To quantitatively evaluate the reconstructed image quality, the indexes of root mean square error (RMSE), peak signal-to-noise ratio (PSNR) [37], feature similarity (FSIM) [38] and structural similarity (SSIM) [39] are employed. Table 2 summarizes all the quantitative results which indicate that the SSCMF can obtain the smallest RMSEs for all channels, followed by the TDL and then TV+LR method which has slightly smaller RMSEs than the TVM. The SART has the largest RMSEs in all channels because of no prior knowledge. These conclusions are also confirmed with the indexes of PSNR, FSIM and SSIM.

b) Material Decomposition—The material decomposition of spectral CT images is a nonlinear ill-posed problem. Regarding the material decomposition methods, they mainly

can be classified into two categories, i.e., projection-based and image-based methods. The projection-based methods model the nonlinear material decomposition problem and usually introduce regularization in the material-specific projection domain. Unfortunately, those methods are very sensitive to noise, and the decomposed images are highly affected by noise [40, 41]. Regarding the image-based methods, they first reconstruct all energy bin images independently using a set of different spectra and then obtain basis material images using a matrix inversion. Such an image-based method fails to characterize the real nonlinearity relationship between the polychromatic projections and the decomposed results. This may cause artifacts in the decomposed results [42].

To validate the proposed SSCMF algorithm in terms of material decomposition, the reconstructed all channel images are decomposed into three basis materials (soft tissue ,bone and iodine contrast agent) utilizing an image-based material decomposition method [43]. Fig. 8 shows the proportions of the corresponding three basis materials and the individual merged color images. From the 1st row of Fig. 8, it can be seen that more pixels of iodine contrast agents are easier to be wrongly classified into the bony regions by the SART, TVM, TV+LR and TDL methods than the SSCMF method. For the soft-tissue component decomposition (2nd row), the SSCMF result provides more finer structures than those achieved by other competing algorithms. In terms of the iodine contrast agent decomposition component, the TDL and SSCMF methods have similar high accuracy. To quantitatively evaluate the material decomposition results, the RMSE, PSNR, SSIM and FSIM values of each basis material from all reconstruction methods are listed in Table 3.

c) Convergence Analysis—The SSCMF model contains the L_0-norm minimization of 4D groupwise coefficients, which is NP-hard and makes the convergence complicated. In this study, we only numerically analyze the convergence of the proposed method. Specifically, the averaged RMSE and PSNR *vs.* iteration number are given in Fig. 9. By comparison, the SSCMF algorithm can converge to an optimized solution quickly with a smaller RMSE or a higher PSNR. Because the spectral CT projection datasets contain severe noise, the RMSE values of SART decrease rapidly and then increase slowly. Fig. 9 indicates that the RMSEs can maintain subsequently stable after 40 iterations for all the reconstruction methods.

d) Parameters Selection—The parameters of SSCMF mainly include the estimated image noise level σ and the hard-threshold value τ. σ depends on the estimated noise at current iteration. That is to say, the higher the image noise levels is, the greater the σ value is. In this study, σ is set as a constant. Regarding the hard-threshold value, τ it is used to suppress noise by removing small coefficients in the 4D transform domain. A greater value of τ can enforce the anti-noising capability and lose more image details. Correspondingly, a smaller value of τ can decrease the anti-noising capability and keep more image details. Here, to investigate the performance of the SSCMF method with respect to different parameter settings, we quantitatively compare the RMSEs and SSIMs from all channels with respect a small σ ($\sigma = 0.005$) have large RMSEs and small SSIMs in all energy chanσ nels. A greater σ to one parameter, as shown in Fig. 10. Fig. 10 (a) and (b) show that the reconstructed images with ($\sigma = 0.1$) may lead to similar conclusions. This means an

appropriate σ value can result in smalls RMSEs and large SSIMs. For this case, the optimal is around 0.016, which matches the noise level in the reconstructed image. From Fig. 10 (c) and (d), we can see that a smaller can result in a of SSIM, a greater τ might obtain better results. For this case, the optimal τ is around 11 and it was smaller RMSE in low energy channel and larger RMSE in high energy channels. In terms of the index set as 12.

e) Low-dose Reconstruction—To evaluate the performance of SSCMF in the case of low-dose reconstruction, the photon number of each x-ray path is reduced to 5×10^3. The lower the photon number of each x-ray path is, the higher the Poisson noise in projection is. Fig. 11 shows the reconstructed images. From Fig. 11, it can be seen that the images reconstructed by the TVM, TV+LR and TDL methods loss some finer structures and details, especially in the high-energy channels. Compared with the TVM and TV+LR results, the images reconstructed by the SSCMF algorithm have clear image edges.

f) Contrast Agent Concentration Analysis—To validate the performance of the proposed SSCMF method for different concentrations of iodine, 0.6%, 2.4% and 4.2% iodine concentrations are injected into the mouse thorax phantom, respectively. Fig. 12 demonstrates the reconstructed results, and it can be seen that the proposed SSCMF has a good performance with the selected iodine contrast concentration. The higher the iodine concentration is, the larger the image contrast is. In clinical applications, it is appropriate to select the iodine contrast concentration so that the final material decomposition results can separate three basis materials.

g) Computational Cost Analysis—Regarding the computational costs of all the optimization algorithms, i.e., TVM, TV+LR, TDL and SSCMF methods, they can be divided into two major procedures: backprojection reconstruction and regularization constraint. The computational cost of multi-energy projection datasets reconstruction step mainly depends on the number of energy channels. The higher the number of energy channels is, the larger the time consumption is. Regarding the regularization term, different regularization constraints have different computational costs and they also depend on the parameter settings. For example, the computational cost of CMF depends on the tensor cube size, search tensor cube window size, the cube sliding step, *etc*. In general, the computational cost to update the regularization term is less than that is consumed by the back projection reconstruction. In this study, all the aforementioned methods are programmed by Matlab (2017b) with GPU techniques on a PC (8 CPUs @3.40GHz, 32.0GB RAM, Intel(R) HD Graphics 530). Here, the back projection reconstruction step consumes114.58 seconds in all algorithms. For the TV, TV+LR, TDL and CMF regularization terms, they cost 1.64, 0.52, 10.65, 38.23 seconds, respectively. The CMF regularization term requires more time than the TV and TV+LR regularizations because it is performed on overlapped small image tensor cubes. Because the number of TV iteration is 20 and it needs to operate on each channel image independently, the TV regularization needs more computational costs than the TV+LR method.

4.2. Preclinical Mouse Study

A mouse with injected gold nanoparticles (GNP) was scanned by a MARS micro spectral CT system. The system includes one x-ray source and one PCD. The distances from the x-ray source to the PCD and rotation axis are 255 *mm* and 158 *mm,* respectively. The PCD horizontally consists of 512 elements, and each element covers a length of 0.11 *mm*, resulting in an FOV with a radius of 9.21mm. 371 raw projections are collected in a full scan. A strategy of multiple scans is adopted, and 13 energy channel datasets are obtained. The projections for the central slice in fan-beam geometry Each reconstructed channel image is a matrix of 512×512, covering area of $18.41\times18.41 mm^2$. are extracted and employed to verify the advantages of the SSCMF algorithm in practical applications.

Fig. 13 shows the reconstructed results of four representative energy channels (1^{st}, 5^{th}, 9^{th} and 13^{th}). From the 1^{st} row of Fig. 13, it can be seen that the images reconstructed by the SART are corrupted by severe noises and lose some details. Although the reconstructed images using the TVM and TV+LR methods (2^{nd} and 3^{rd} rows in Fig. 13) contain less noise, the soft tissue structures are blurred to some degree, especially the obvious blocky effects in the TVM reconstruction. In contrast, the TDL and SSCMF methods achieve a great success in recovering finer structures as well as preserving image edges (see the 4^{th} and 5^{th} rows).

To further demonstrate the advantages of the SSCMF method, a bony ROI A and soft tissue ROI B are extracted and magnified as in Fig. 14. For the magnified bony ROI A, one can see that the image structures indicated by the red arrows “1” and “2” can hardly be distinguished in the images reconstructed by the SART, TVM and TV+LR methods. The image structure indicated by the red arrow “2” is slightly blurred in the images reconstructed by the TDL technique. However, the images reconstructed by the SSCMF have clearer edge features. As for the soft tissue ROI B, we can see that the images reconstructed by the SSCMF are smoother with more image details. The region marked with the arrow “3” indicates the unique merits of the TDL and SSCMF in recovering fine structures. The image edge structures indicated by arrow “4” is more clearer using the SSCMF than that obtained from the TDL method.

Regarding the material decomposition, Fig. 15 shows three decomposed basic materials (tissue, bone and GNP) and color rendering images. For the bone material, the profile of SSCMF is clearer than those obtained from the SART, TVM, TV+LR and TDL methods. In addition, the SSCMF can also provide more accurate tissue decomposition results without blocky artifacts. In the rendered color images, the edges of the SSCMF method are more obvious than other competing algorithms.

To further analyze the impacts of the parameters on the reconstructed image quality in practical series of different parameter settings. From Fig. 16, it can be seen that a small σ value (0.005 and 0.01) applications, Figs. 16 and 17 show the reconstructed mouse images using the SSCMF method with a can reduce the algorithm robustness and a greater σ value (0.028, 0.04 and 0.06) can smear finerτ structures and image edges. Regarding the selection of parameter , a similar conclusion can be made.

5. Discussions and Conclusions

Although the SSCMF method reconstructs excellent results for the spectral CT, there are still some issues to be addressed. First, there are numerous parameters in the BM4D denoising procedure to be optimized for Eq. (25), including the tensor cube size, tensor cube searching window size, the number of tensor cubes, *etc.* In this work, we only empirically optimized the main parameters based on extensive experiments. However, the theoretical analysis of optimization for all parameters is still open. Second, different similarity measurements for cubes may generate different groups and further result in different reconstruction results. In this study, the sum of squared differences between two input tensor cubes is considered as the similarity measurement [22], which can be read as

$$d\left(\mathcal{Y}_{j,r}, \mathcal{Y}_{j,l}\right) = \frac{\|\mathcal{Y}_{j,r} - \mathcal{Y}_{j,l}\|_F^2}{L^3},$$

where L^3 is a normalized factor. Of course, there are other methods to quantify such similarity, such as SVM-KNN [44], PCA basis [45], *etc.* In our future work, we will investigate the reconstructed results with respect to different similarity measurements. Third, because the BM4D denoising codes are only implemented on CPU, we will transplant this procedure to GPU to speedup the convergence. Fourth, we only employ a simple strategy to solve Eq. (16). In fact, there are many other available methods, such as split-bregman method [8], alternative direction minimization method (ADMM) [46], *etc.* Fifth, because a volumetric CT image has 3D similar structures, the proposed CMF regularization term can be directly applied for 3D volumetric CT reconstruction. Furthermore, a 4DCT image not only has similarity in 3D volumetric space but also has structural and intensity similarities among different time frames. Moreover, as for the cone-beam spectral CT reconstruction, it is necessary to extend the established CMF to 5D so that it can explore both the 3D volumetric structures and 1D spectral similarities.

In this study, the numbers of energy channel are 8 and 13 in numerical simulation and clinical experiment, respectively. In practice, the number of energy channel may be smaller. To evaluate the proposed method for the case with less channels, the aforementioned spectrum in numerical simulation is divided into four different energy channels: [16, 25) keV, [26, 33) keV, [34, 40) keV, [41, 50) keV. Fig. 18 shows all channel images reconstructed by using the SSCMF method. It can be observed that the SSCMF algorithm still perform well in reconstructing high quality images from less energy channels.

In summary, to take advantages of the similarity of nonlocal spatial-spectral cubes in spectral CT images, we first propose a CMF for BM4D procedure, and then it is introduced into the spectral CT reconstruction model as a new regularizer. By incorporating the CMF into basic spectral CT reconstruction model, the reconstructed image quality is dramatically improved. The advantages of the SSCMF method are twofold. First, it can explore both the nonlocal similarity of an image and structural similarity among different spectral channels by grouping similar spatial-spectral cubes. This makes it a great success for the SSCMF method to preserve image edges and recover finer structures. Second, performing

collaborative filtering on the raw group and 1st estimation can improve the algorithm robustness. Both numerical simulations and realistic preclinical mouse experiments confirm these merits of the proposed SSCMF method by comparing with the SART, TVM, TV+LR and TDL methods.

Acknowledgement

This study is funded in part by the China Scholarship Council (No. 201706050070), National Natural Science Foundation of China (No. 61471070), NIH/NIBIB U01 grant (EB017140), and National Instrumentation Program of China (No. 2013YQ030629). The authors would like to thank the MARS team in New Zealand for providing the preclinical mouse datasets.

Reference:

1. Kim K, Ye JC, Worstell W, Ouyang J, Rakvongthai Y, Fakhri GE, Li Q. Sparse-View Spectral CT Reconstruction Using Spectral Patch-Based Low-Rank Penalty. IEEE Transactions on Medical Imaging. 2015;34:748–60. [PubMed: 25532170]
2. Wu W, Zhang Y, Wang Q, Liu F, Chen P, Yu H. Low-dose spectral CT reconstruction using L0 image gradient and tensor dictionary. arXiv preprint arXiv:180101452. 2017.
3. Johnson TR, Krauss B, Sedlmair M, Grasruck M, Bruder H, Morhard D, Fink C, Weckbach S,Lenhard M, Schmidt B. Material differentiation by dual energy CT: initial experience. European Radiology. 2007;17:1510–7. [PubMed: 17151859]
4. Graser A, Johnson TR, Bader M, Staehler M, Haseke N, Nikolaou K, Reiser MF, Stief CG,Becker CR. Dual energy CT characterization of urinary calculi: initial in vitro and clinical experience. Investigative radiology. 2008;43:112–9. [PubMed: 18197063]
5. Walsh M, Opie A, Ronaldson J, Doesburg R, Nik S, Mohr J, Ballabriga R, Butler A, Butler P. First CT using Medipix3 and the MARS-CT-3 spectral scanner. Journal of Instrumentation. 2011;6:C01095.
6. Taguchi K, Stierstorfer K, Polster C, Lee O, Kappler S, Spatio-energetic cross-talk in photon counting detectors: N× N binning and sub-pixel masking. 2018: Publisher.
7. Xu Q, Yu H, Bennett J, He P, Zainon R, Doesburg R, Opie A, Walsh M, Shen H, Butler A.Image reconstruction for hybrid true-color micro-CT. IEEE transactions on biomedical engineering. 2012;59:1711–9. [PubMed: 22481806]
8. Chu J, Li L, Chen Z, Wang G, Gao H, Multi-energy CT reconstruction based on low rank and sparsity with the split-bregman method (MLRSS). 2012: Publisher.
9. Semerci O, Hao N, Kilmer ME, Miller EL. Tensor-based formulation and nuclear norm regularization for multienergy computed tomography. IEEE Transactions on Image Processing. 2014;23:1678–93. [PubMed: 24808339]
10. Zhao B, Gao H, Ding H, Molloi S. Tight-frame based iterative image reconstruction for spectral breast CT. Medical Physics. 2013;40:031905. [PubMed: 23464320]
11. Rigie DS, Rivière PJL. Joint Reconstruction of Multi-channel, Spectral CT Data via Constrained Total Nuclear Variation Minimization. Physics in Medicine & Biology. 2015;60:1741–62. [PubMed: 25658985]
12. Semerci O, Hao N, Kilmer ME, Miller EL. Tensor-Based Formulation and Nuclear Norm Regularization for Multienergy Computed Tomography. IEEE Transactions on Image Processing. 2014;23:1678–93. [PubMed: 24808339]
13. Zeng D, Gao Y, Huang J, Bian Z, Zhang H, Lu L, Ma J. Penalized weighted least-squares approach for multienergy computed tomography image reconstruction via structure tensor total variation regularization. Computerized Medical Imaging and Graphics. 2016;53:19–29. [PubMed: 27490315]
14. Zhao B, Ding H, Lu Y, Wang G, Zhao J, Molloi S. Dual-dictionary learning-based iterative image reconstruction for spectral computed tomography application. Physics in medicine and biology. 2012;57:8217. [PubMed: 23192234]

15. Zhang Y, Mou X, Wang G, Yu H. Tensor-based dictionary learning for spectral CT reconstruction. IEEE transactions on medical imaging. 2017;36:142–54. [PubMed: 27541628]
16. Yu Z, Leng S, Li Z, McCollough CH. Spectral prior image constrained compressed sensing(spectral PICCS) for photon-counting computed tomography. Physics in Medicine & Biology. 2016;61:6707. [PubMed: 27551878]
17. Zhang Y, Xi Y, Yang Q, Cong W, Zhou J, Wang G. Spectral CT reconstruction with images parsity and spectral mean. IEEE Transactions on Computational Imaging. 2016;2:510–23. [PubMed: 29034267]
18. Dabov K, Foi A, Katkovnik V, Egiazarian KO, Image restoration by sparse 3D transform domain collaborative filtering. 2008: Publisher.
19. Danielyan A, Katkovnik V, Egiazarian K. BM3D frames and variational image deblurring. IEEE Trans Image Process. 2012;21:1715–28. Epub 2011/12/01. [PubMed: 22128008]
20. Li X Image recovery via hybrid sparse representations: A deterministic annealing approach. IEEE Journal of Selected Topics in Signal Processing. 2011;5:953–62.
21. Crandall R, Bilgin A, Lossless image compression using causal block matching and 3dcollaborative filtering. 2014: Publisher.
22. Maggioni M, Katkovnik V, Egiazarian K, Foi A. Nonlocal transform-domain filter forvolumetric data denoising and reconstruction. IEEE transactions on image processing. 2013;22:119–33. [PubMed: 22868570]
23. Xie Q, Zhao Q, Meng D, Xu Z, Gu S, Zuo W, Zhang L, Multispectral images denoising byintrinsic tensor sparsity regularization. 2016: Publisher.
24. Chang Y, Yan L, Zhong S, Hyperspectral image denoising via spectral and spatial low-rank approximation. 2017: Publisher.
25. Niu S, Yu G, Ma J, Wang J. Nonlocal low-rank and sparse matrix decomposition for spectral CT reconstruction. Inverse Problems. 2018;34:024003. [PubMed: 30294061]
26. Gao H, Yu H, Osher S, Wang G. Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM). Inverse problems. 2011;27:115012. [PubMed: 22223929]
27. Wen Y-W, Ng MK, Ching W-K. Iterative algorithms based on decoupling of deblurring and denoising for image restoration. SIAM Journal on Scientific Computing. 2008;30:2655–74.
28. Andersen AH, Kak AC. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm. Ultrasonic imaging. 1984;6:81–94. [PubMed: 6548059]
29. Sidky EY, Pan X. Image reconstruction in circular cone-beam computed tomography byconstrained, total-variation minimization. Physics in medicine and biology. 2008;53:4777. [PubMed: 18701771]
30. Chen Z, Jin X, Li L, Wang G. A limited-angle CT reconstruction method based on anisotropicTV minimization. Physics in Medicine & Biology. 2013;58:2119. [PubMed: 23470430]
31. Dong B, Li J, Shen Z. X-ray CT image reconstruction via wavelet frame based regularizationand Radon domain inpainting. Journal of Scientific Computing. 2013;54:333–49.
32. Yu W, Wang C, Huang M. Edge-preserving reconstruction from sparse projections of limited angle computed tomography using ℓ0-regularized gradient prior. Review of Scientific Instruments. 2017;88:043703. [PubMed: 28456252]
33. Cai J-F, Jia X, Gao H, Jiang SB, Shen Z, Zhao H. Cine cone beam CT reconstruction usinglow-rank matrix factorization: algorithm and a proof-of-principle study. IEEE transactions on medical imaging. 2014;33:1581–91. [PubMed: 24771574]
34. Wu W, Yu H, Gong C, Liu F. Swinging multi-source industrial CT systems for aperiodic dynamic imaging. Optics express. 2017;25:24215–35. [PubMed: 29041367]
35. Goldstein T, Osher S. The split Bregman method for L1-regularized problems. SIAM journalon imaging sciences. 2009;2:323–43.
36. Setzer S Operator splittings, Bregman methods and frame shrinkage in image processing.International Journal of Computer Vision. 2011;92:265–80.
37. Chandler DM, Hemami SS. VSNR: A wavelet-based visual signal-to-noise ratio for natural images. IEEE transactions on image processing. 2007;16:2284–98. [PubMed: 17784602]

38. Zhang L, Zhang L, Mou X, Zhang D. FSIM: A feature similarity index for image qualityassessment. IEEE transactions on Image Processing. 2011;20:2378–86. [PubMed: 21292594]

39. Wang Z, Simoncelli EP, Bovik AC, Multiscale structural similarity for image quality assessment. 2003: Publisher.

40. Ducros N, Abascal JFPJ, Sixou B, Rit S, Peyrin F. Regularization of nonlinear decomposition of spectral x-ray projection images. Medical physics. 2017;44.

41. Liu J, Gao H. Material reconstruction for spectral computed tomography with detector response function. Inverse Problems. 2016;32:114001.

42. Wang Q, Zhu Y, Yu H. Locally linear constraint based optimization model for material decomposition. Physics in Medicine & Biology. 2017;62:8314. [PubMed: 28933350]

43. Granton P, Pollmann S, Ford N, Drangova M, Holdsworth D. Implementation of dual-and triple-energy cone- beam micro-CT for post reconstruction material decomposition. Medical physics. 2008;35:5030–42. [PubMed: 19070237]

44. Zhang H, Berg AC, Maire M, Malik J, SVM-KNN: Discriminative nearest neighbor classification for visual category recognition. 2006: Publisher.

45. Dabov K, Foi A, Katkovnik V, Egiazarian K, BM3D image denoising with shape-adaptive principal component analysis. 2009: Publisher.

46. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statisticallearning via the alternating direction method of multipliers. Foundations and Trends® in Machine learning. 2011;3:1–122.

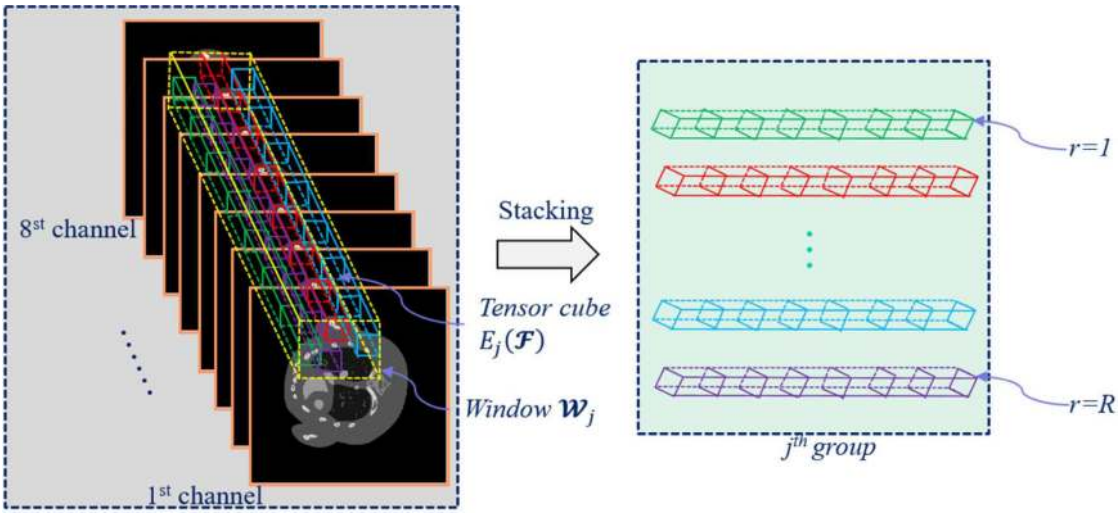

Fig.1.
Illustration of grouping spatial-spectral cubes.

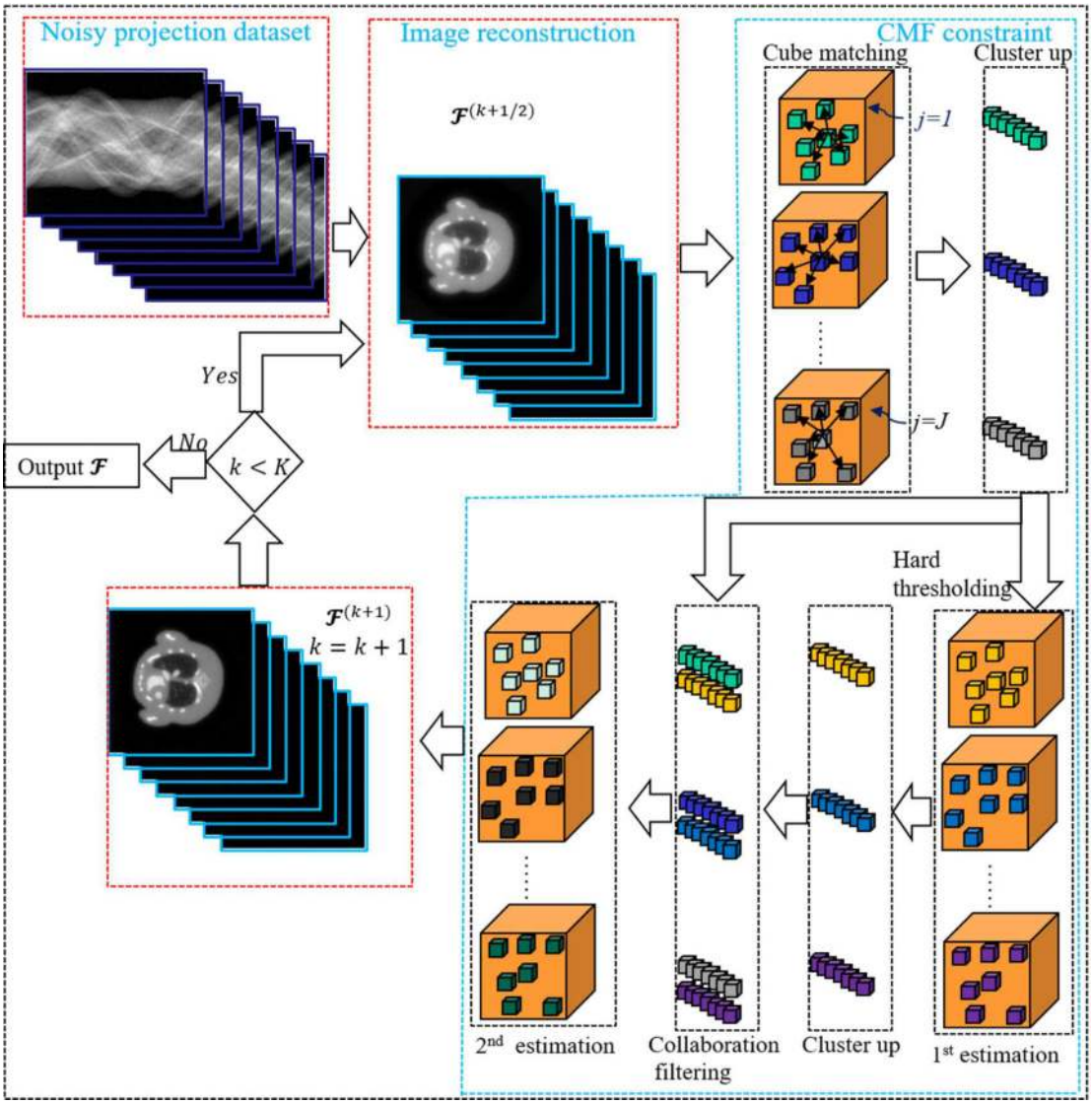

Fig.2.
The flowchart of SSCMF method.

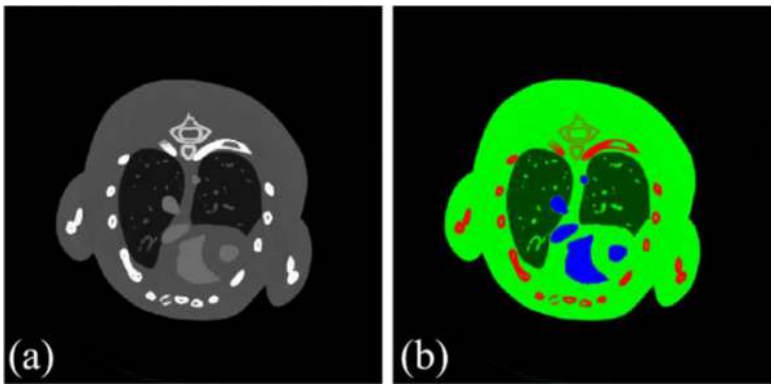

Fig 3.

The mouse thorax phantom. (a) is the 1st channel image in [0, 3] cm^{-1} reconstructed by the SART with full-scan noise-free projections. (b) is the merged color image where green, red and blue stand for soft tissue, bone and iodine, respectively.

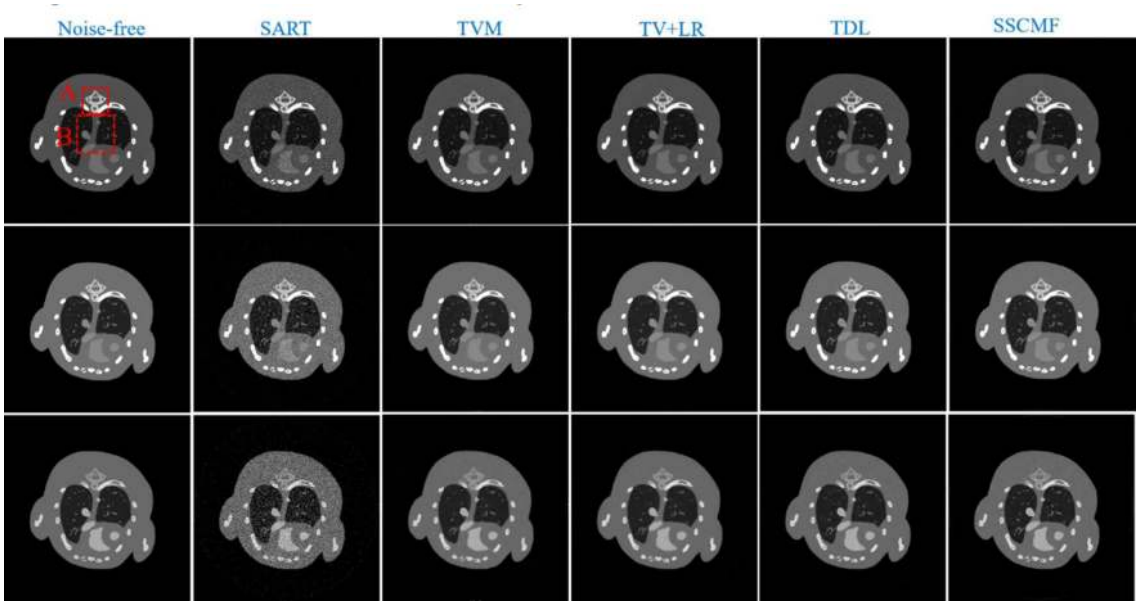

Fig. 4.

Reconstructed results of three representative energy-channel images. The 1st to 3rd row images correspond to the 1st, 4th and 8th channels, and their display windows are [0, 3], [0, 1.2] and [0, 0.8] cm^{-1}, respectively.

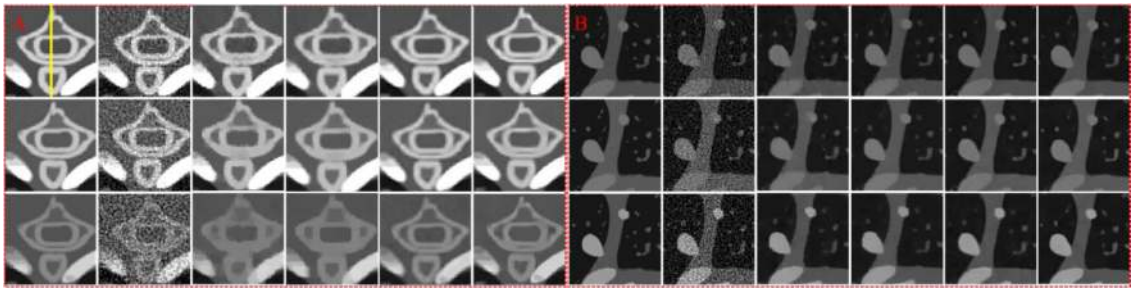

Fig. 5.
Magnified ROIs A (bony structure) and B (soft tissue) in Fig. 4.

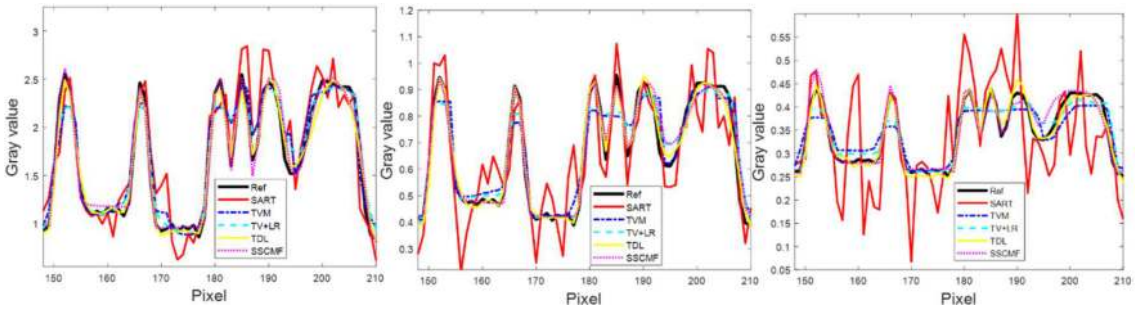

Fig. 6.

Representative profiles along the yellow line in bony ROI A in Fig. 5. (a), (b) and (c) are for 1, 4 and 8 channels, respectively.

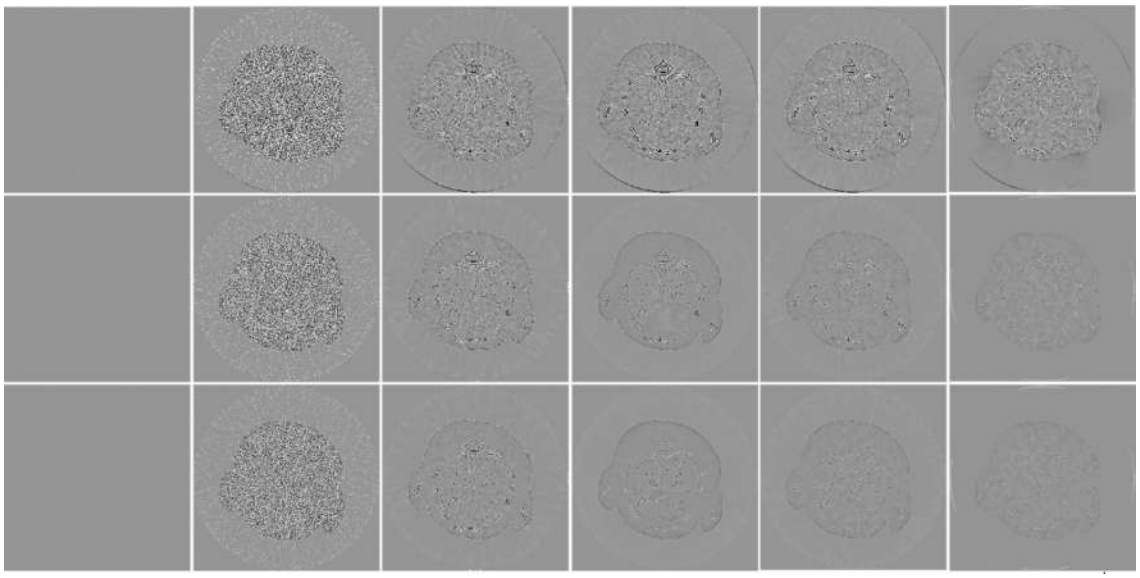

Fig. 7.
Same as Fig. 4 but the difference images with respect to references. The display window is $[-0.2, 0.2]$ cm^{-1}.

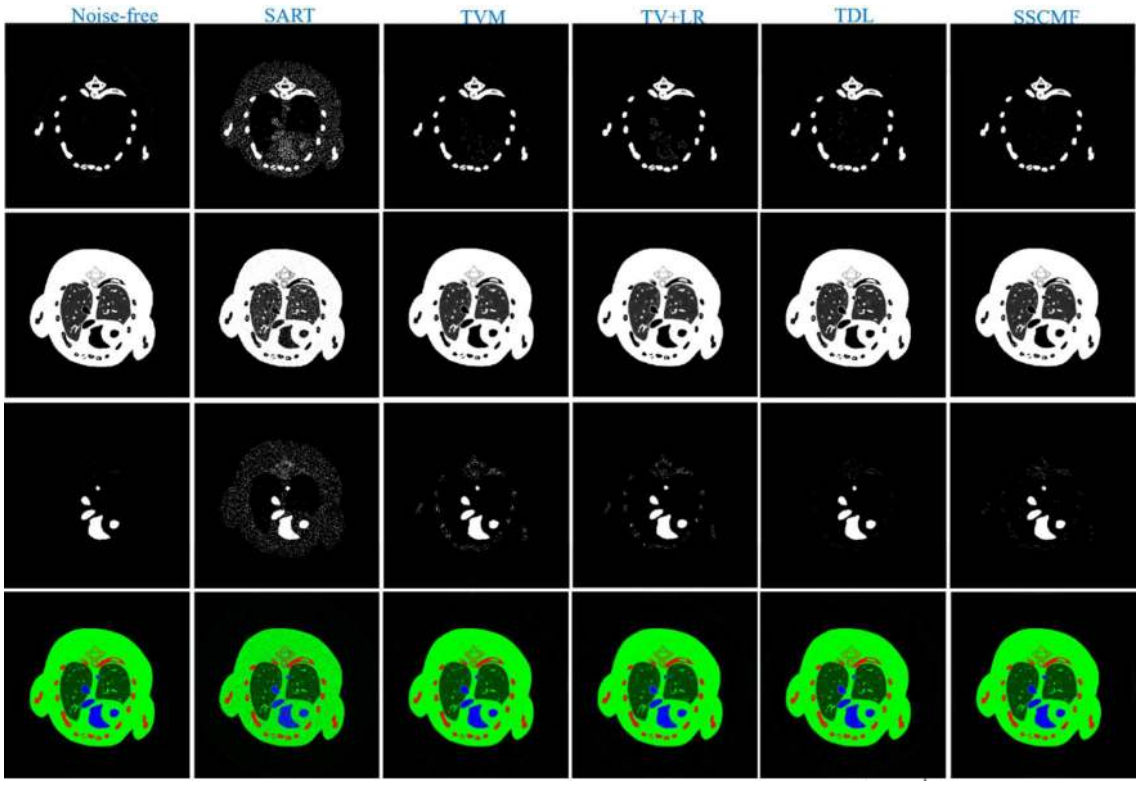

Fig. 8.

Three basic materials decomposed from the reconstructed results in Fig.4. The 1st to 3rd rows represent bone, soft tissue and iodine. The 4th row images are color renderings, where red, green and blue represent bone, soft tissue and iodine, respectively. For the 1st to 3rd rows, the display windows are [0, 0.1], [0.2, 0.8] and [0, 0.5] cm^{-1}, respectively

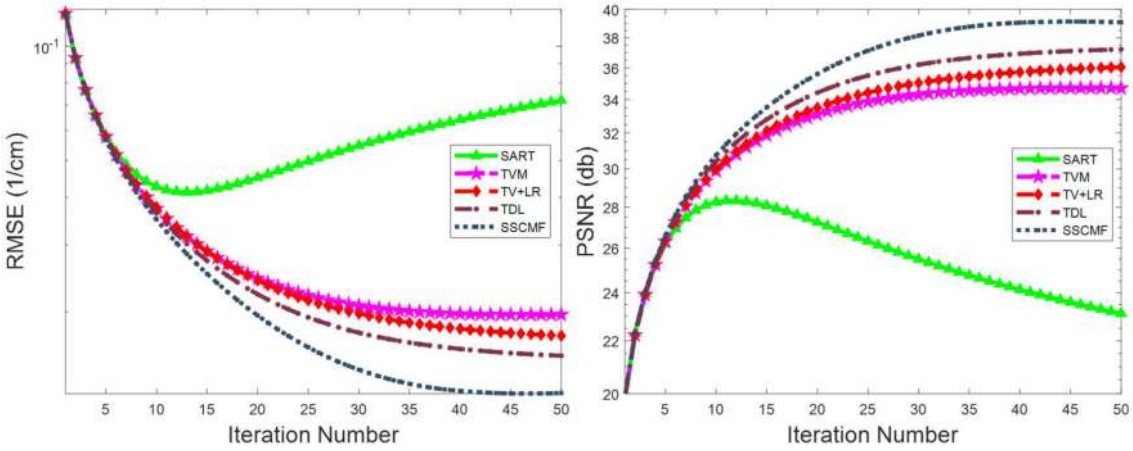

Fig. 9.

Convergence curves of averaged RMSE (a) and PSNR (b) *vs.* iteration number.

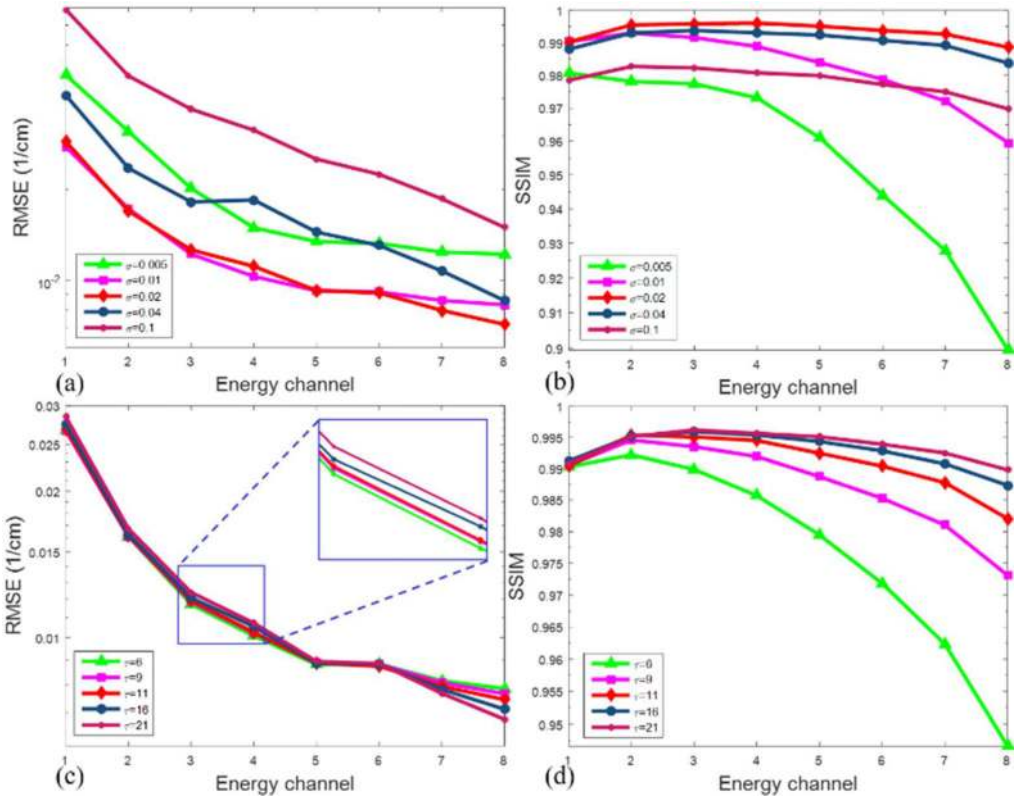

Fig. 10.
Quantitative results of all energy channels with respect to different parameters.

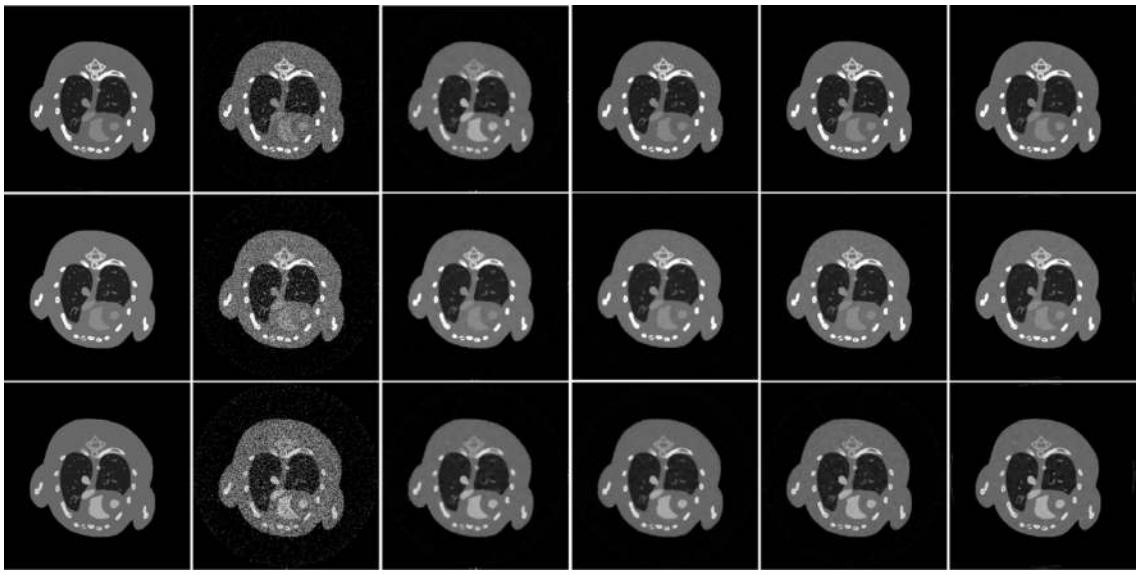

Fig. 11.
Same as Fig.4 but reconstructed with 5×10^3 photons per x-ray path.

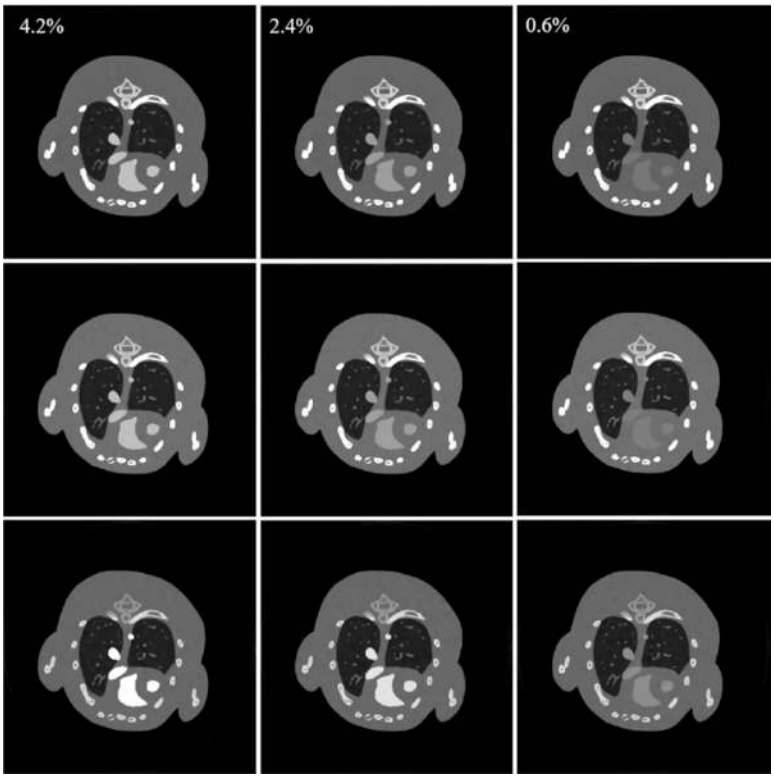

Fig. 12.

Results reconstructed by the SSCMF method with different iodine concentrations. From 1st to 3rd columns, the iodine concentrations are 4.2%, 2.4% and 0.6%, respectively. The 1st to 3rd row images correspond to the 1st, 4th and 8th channels, and their display windows are [0, 3], [0, 1.2] and [0, 0.8] cm^{-1}, respectively

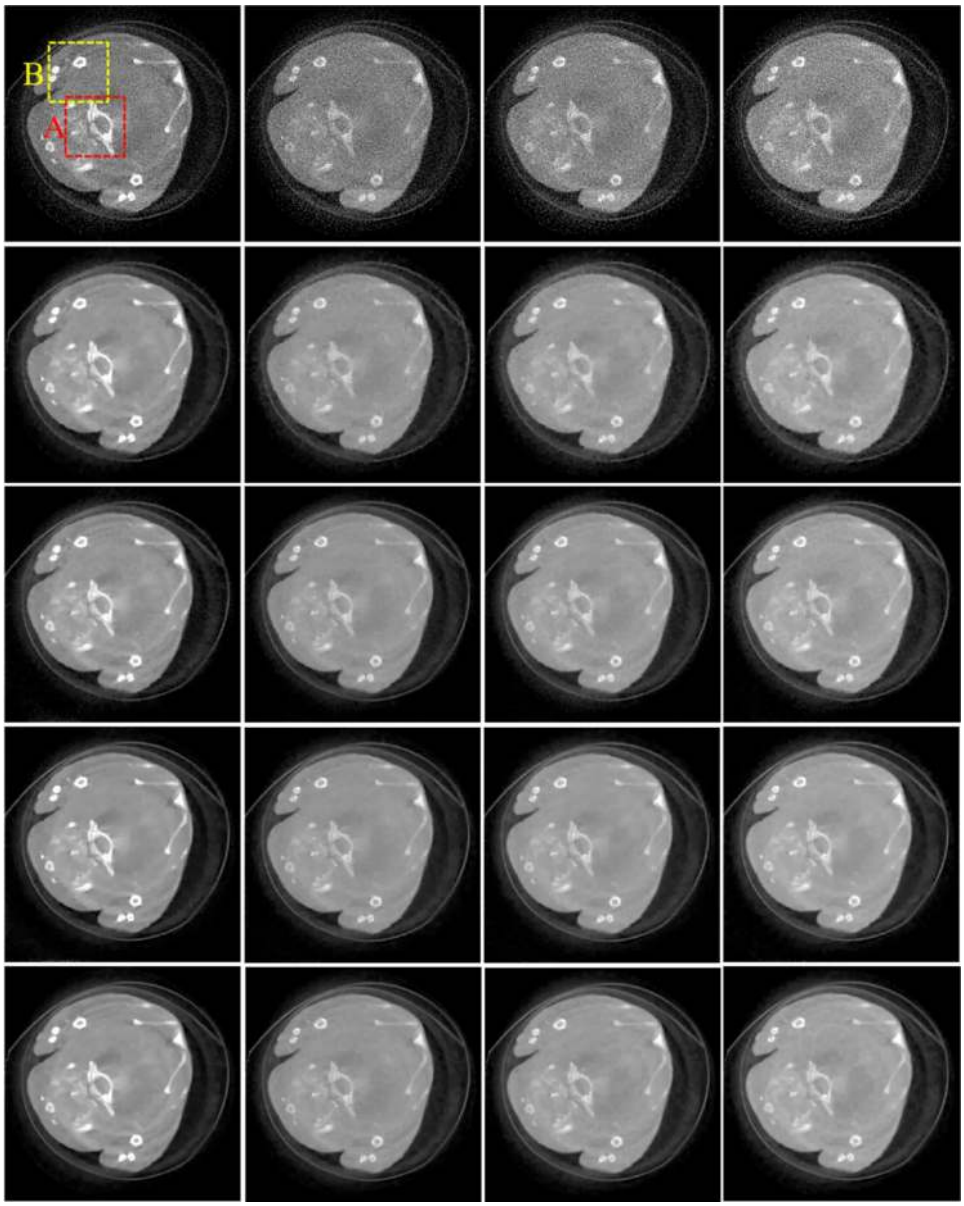

Fig. 13.

Reconstructed results of four representative energy-channels. From the 1st to 4th columns, the images are for different energy-channels (1st, 5th, 9th and 13th). From the 1st to 4th rows, the images are reconstructed using the SART, TVM, TV+LR and SSCMF methods, respectively. The display window for all the reconstructed images are [0, 0.8] cm^{-1}.

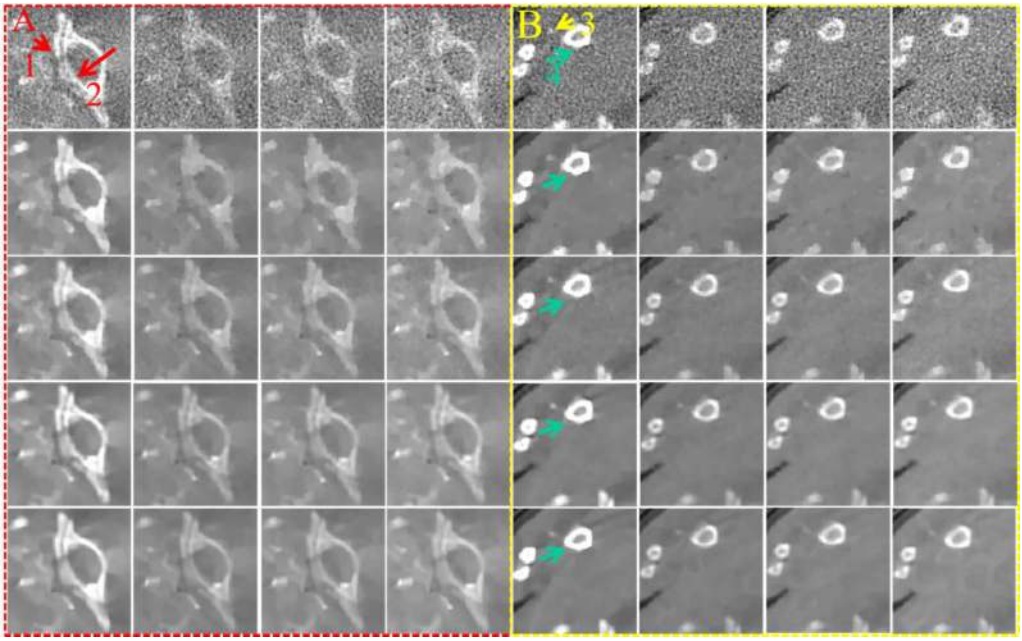

Fig. 14.
The magnified ROIs A and B in Fig.13.

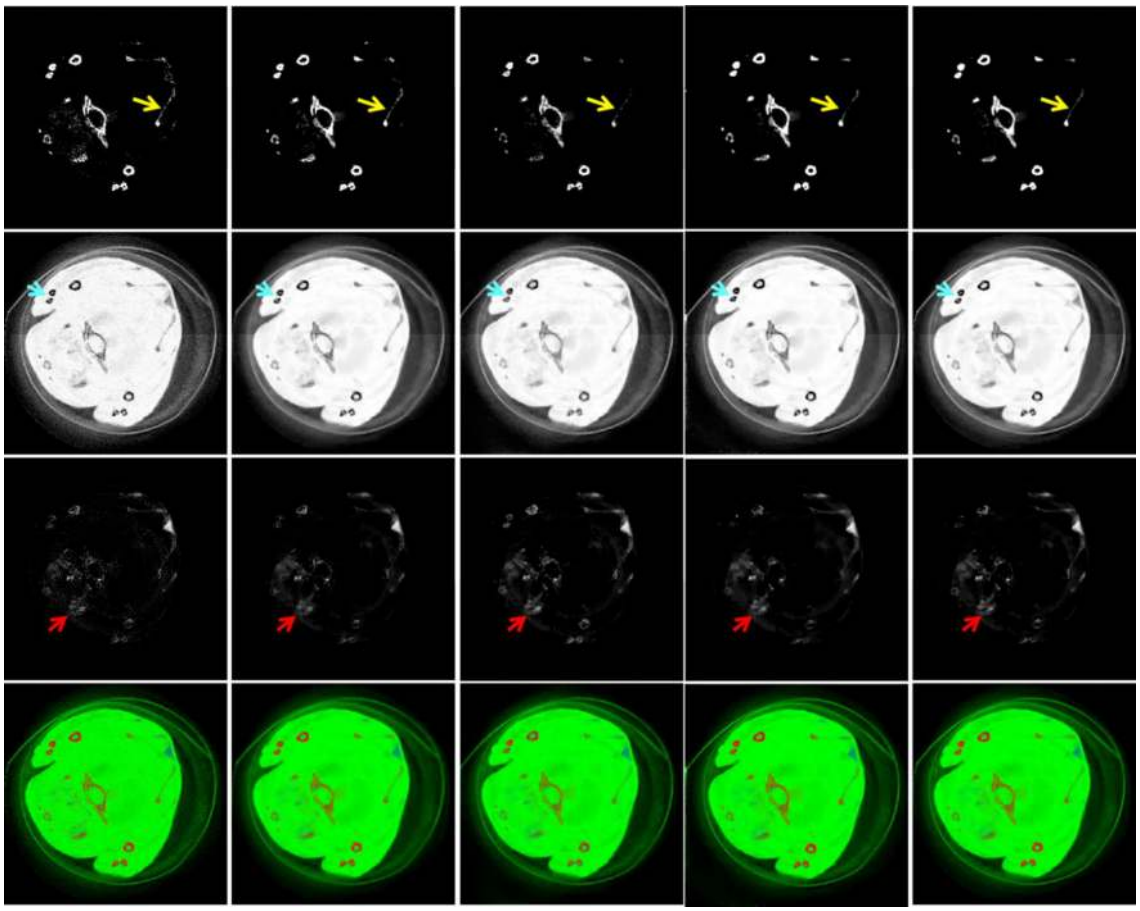

Fig. 15.

Three decomposed basic materials of the Fig. 13. From the 1st to 3rd rows, the images are for bone, soft tissue and GNP, respectively. The 4th row images are merged color images, where red, green and blue represent bone, soft tissue and GNP, respectively. For the 1st to 3rd rows, the display windows are [0.1, 0.5], [0, 1] and [0, 0.5] cm^{-1}, respectively.

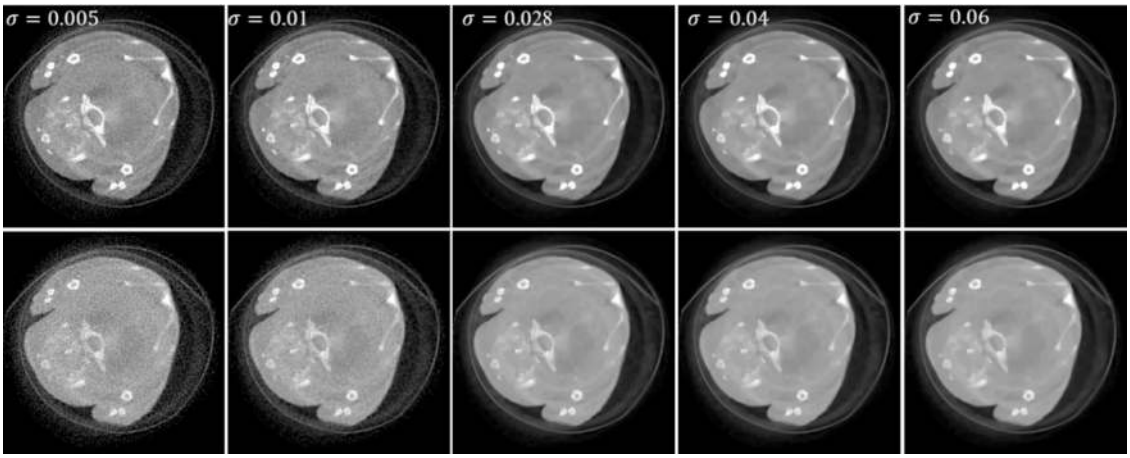

Fig. 16.

Reconstructed results with different σ. The 1st and 2nd rows represent different channels (1st and 13th). From the 1st to 5th columns, the images are reconstructed with different σ and a fixed $\tau = 0.35$. The display window is [0, 0.8] cm^{-1}.

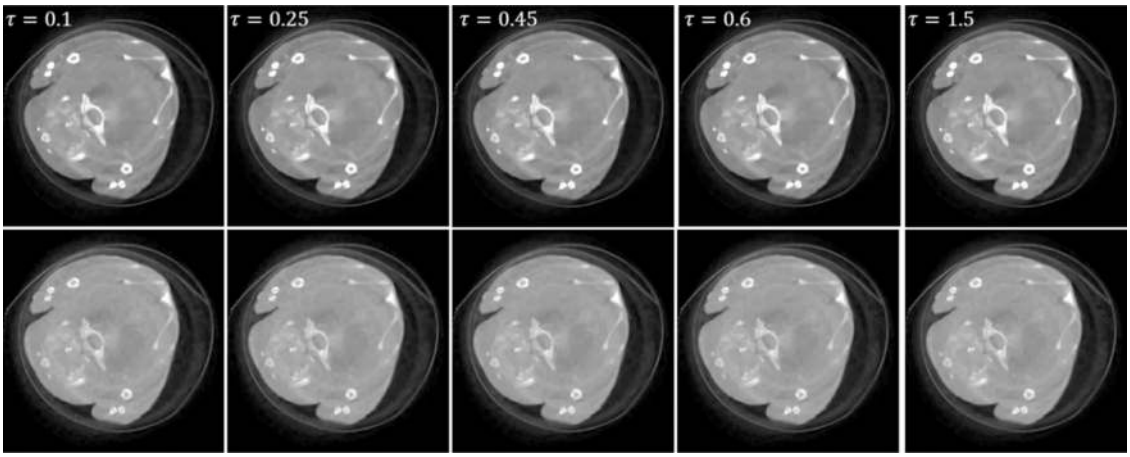

Fig. 17.
Same as Fig. 16 but with different τ and a fixed σ = 0.016.

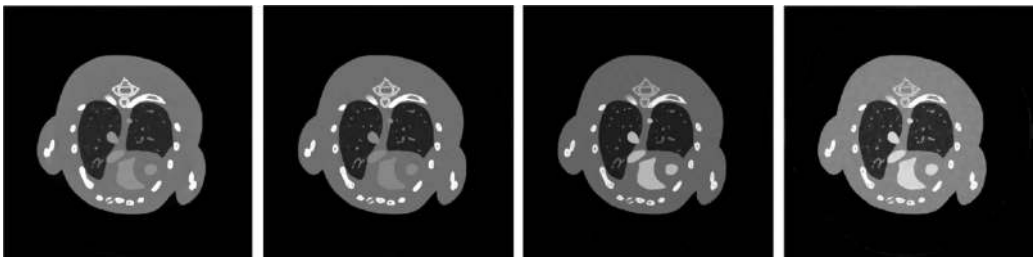

Fig. 18.

Reconstructed results from 4 energy channels. From left to right, the images are for 1–4 energy channels and their display windows are [0, 2] cm^{-1}, [0, 1.2] cm^{-1}, [0, 1.0] cm^{-1} and [0, 0.6] cm^{-1}, respectively.

Table 1.

Reconstruction parameters for both numerical simulations and preclinical datasets.

	Photon number	**σ**	**τ**
Mouse thorax phantom	2×10^4	0.016	12
	5×10^3	0.05	10
Preclinical dataset		0.02	0.35

Table 2.

Quantitative image quality evaluation results.

Index	Method	1st	2nd	3rd	4th	5th	6th	7th	8th
RMSE (10^{-2})	SART	11.88	8.50	6.85	6.23	6.07	6.35	6.00	5.72
	TVM	3.80	2.57	1.98	1.73	1.55	1.53	1.37	1.21
	TV+LR	3.69	2.34	1.73	1.40	1.22	1.34	1.14	0.98
	TDL	3.40	2.14	1.54	1.22	1.05	1.07	0.96	0.87
	SSCMF	**2.68**	**1.62**	**1.18**	**1.03**	**0.88**	**0.88**	**0.78**	**0.73**
PSNR	SART	18.50	21.40	23.28	24.11	24.33	23.94	24.44	24.86
	TVM	28.40	31.81	34.08	35.26	36.17	36.31	37.23	38.34
	TV+LR	28.67	32.60	35.25	37.09	38.29	37.43	38.86	40.13
	TDL	29.37	33.38	36.23	38.24	39.57	39.39	40.32	41.25
	SSCMF	**31.44**	**35.83**	**38.53**	**39.74**	**41.10**	**41.14**	**42.11**	**42.76**
SSIM	SART	0.9125	0.8838	0.8647	0.8273	0.7737	0.7251	0.6801	0.6291
	TVM	0.9901	0.9876	0.9835	0.9746	0.9595	0.9407	0.9220	0.8918
	TV+LR	0.9909	0.9877	0.9843	0.9759	0.9684	0.9583	0.9318	0.9152
	TDL	0.9909	0.9924	0.9943	0.9901	0.9896	0.9874	0.9724	0.9506
	SSCMF	**0.9915**	**0.9956**	**0.9962**	**0.9949**	**0.9930**	**0.9911**	**0.9892**	**0.9865**
FSIM	SART	0.8822	0.8345	0.8223	0.8037	0.7797	0.7567	0.7384	0.7177
	TVM	0.9780	0.9730	0.9713	0.9647	0.9497	0.9326	0.9246	0.9133
	TV+LR	0.9870	0.9836	0.9815	0.9722	0.9564	0.9365	0.9347	0.9238
	TDL	0.9822	0.9860	0.9876	0.9844	0.9811	0.9774	0.9597	0.9308
	SSCMF	**0.9871**	**0.9909**	**0.9885**	**0.9855**	**0.9824**	**0.9794**	**0.9761**	**0.9711**

Table 3.

Quantitative evaluation results of three decomposed basis materials.

Index	Method	Bone	Soft tissue	Iodine contrast agent
RMSE (10^{-2})	SART	1.27	6.77	4.96
	TVM	0.70	2.61	1.95
	TV+LR	0.85	2.67	1.87
	TDL	0.77	2.10	1.20
	SSCMF	**0.50**	**1.90**	**1.19**
PSNR	SART	37.92	23.39	26.08
	TVM	43.11	31.68	34.20
	TV+LR	41.40	31.47	34.59
	TDL	42.25	33.55	38.12
	SSCMF	**46.06**	**34.44**	**38.50**
SSIM	SART	0.9118	0.6176	0.7596
	TVM	0.9976	0.8923	0.9470
	TV+LR	0.9949	0.8654	0.9544
	TDL	0.9982	0.8424	0.9628
	SSCMF	**0.9984**	**0.9687**	**0.9661**
FSIM	SART	0.9899	0.8098	0.9084
	TVM	0.9985	0.9699	0.9578
	TV+LR	0.9981	**0.9760**	0.9581
	TDL	0.9986	0.9753	**0.9604**
	SSCMF	**0.9988**	0.9755	**0.9604**