Dictionary learning based image-domain material decomposition for spectral CT

Weiwen Wu, Haijun Yu, Peijun Chen, Fulin Luo, Fenglin Liu, Qian Wang, Yining Zhu, Yanbo Zhang, Jian Feng and Hengyong Yu

1 Key Lab of Optoelectronic Technology and Systems, Ministry of Education, Chongqing University, Chongqing 400044, People's Republic of China
2 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, People's Republic of China
3 Engineering Research Center of Industrial Computed Tomography Nondestructive Testing, Ministry of Education, Chongqing University, Chongqing 400044, People's Republic of China
4 Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell, MA 01854, United States of America
5 School of Mathematical Sciences, Capital Normal University, Beijing 100048, People's Republic of China
6 Ping An Technology, U.S. Research Laboratory, Palo Alto, CA 94306, United States of America

E-mail: liuf0@cqu.edu.cn and hengyong-yu@ieee.org

Keywords: spectral computed tomography (CT), material decomposition, image domain, dictionary learning

Abstract
The potential huge advantage of spectral computed tomography (CT) is that it can provide accurate material identification and quantitative tissue information by material decomposition. However, material decomposition is a typical inverse problem, where the noise can be magnified. To address this issue, we develop a dictionary learning based image-domain material decomposition (DLIMD) method for spectral CT to achieve accurate material components with better image quality. Specifically, a set of image patches are extracted from the mode-1 unfolding of normalized material images decomposed by direct inversion to train a unified dictionary using the K-SVD technique. Then, the DLIMD model is established to explore the redundant similarities of the material images, where the split-Bregman is employed to optimize the model. Finally, more constraints (i.e. volume conservation and the bounds of each pixel within material maps) are integrated into the DLIMD model. Numerical phantom, physical phantom and preclinical experiments are performed to evaluate the performance of the proposed DLIMD in material decomposition accuracy, material image edge preservation and feature recovery.

1. Introduction
The conventional computed tomography (CT) treats x-ray spectrum as single-energy, and an energy integration detector (EID) is employed to collect photons from the spectrum (Clackdoyle et al 2019, Macdonald et al 2020). As a result, different material components may have the same or similar linear attenuation coefficients in the reconstructed CT images. This implies that it fails to discriminate and decompose material components within the imaging objects (Shikhaliev and Fritz 2011, Lin et al 2012), and it can miss the functional information. Fortunately, due to its advantages in dose reduction, tissue contrast improvement, quantitative tissue analysis, beam hardening artifacts reduction, material discrimination and decomposition (Kim et al 2015), the spectral CT has been developed as an alternative solution. Compared with the traditional CT schemes, spectral CT systems can obtain two or more energy bin projections for the scanned object. The dual energy CT (DECT), as a simple commercialized architecture, has achieved a great success in practical applications (Saito et al 2018). However, the projections from DECT only contain two energy-spectrum measurements. This limits the material decomposition capability of DECT (Johnson et al 2007).

The development of photon-counting detectors (PCDs) makes spectral CT a hot topic in recent years (Taguchi and Iwanczyk 2013). A typical spectral CT system equipped with PCD can collect multiple
projections from different energy bins within one scan. However, different materials may have similar attenuation coefficients in some of the energy bins. Besides, how to perform a quantitative analysis for tissue distribution is still a difficult problem. To realize the aims of material discrimination and tissue quantitative analysis, it is necessary to develop advanced material decomposition methods for spectral CT.

Regarding the material decomposition methods for spectral CT, they can be divided into two categories: direct and indirect methods. Specifically, a direct material decomposition method can directly obtain material components from projections (Long and Fessler 2014, Zhao et al 2015, Barber et al 2016, Liu et al 2016) with known x-ray spectrum. However, due to x-ray beam hardening and scattering during the process of x-ray transmission, it is difficult to achieve x-ray transmission spectral model in practice. Besides, although some regularization priors (for examples, total variation (TV) (Barber et al 2016, Gao et al 2019), nonlocal TV (Liu et al 2016) and so on) have been considered in such material decomposition models, the results from direct material decomposition are still sensitive to noise.

Indirect material decomposition methods can be further divided into projection-based and image-based methods (Wang et al 2017). For the projection-based methods, the projections are first decomposed to material specific sinograms. Then, an image reconstruction algorithm (for example, back-projection (FBP) (Wu et al 2017) and regularization prior based iteration methods (Xiao et al 2018, 2020, Zhang et al 2019)) is followed. For such a material decomposition method, the errors from the specific basis material sinograms can be magnified in the material maps, resulting in a compromised decomposition accuracy. For the image-based methods, the first step is to reconstruct spectral CT images from projections, and the second step is to decompose them into material maps. In terms of the first step, there are a lot of iterative image reconstruction methods available, such as channel-independent TV (Xu et al 2012a) and advanced non-local low-rank cube-based tensor factorization (NLCTF) (Wu et al 2019), etc. Regarding the second step, numerous methods were proposed for DECT (Maaß et al 2009, Niu et al 2014, Zhao et al 2016, Xue et al 2017). Recently, to further obtain better quality of material decomposition maps with higher accuracy, Li et al proposed learned mixed material models (Li et al 2018, 2019b) and a cross-material model (Li et al 2019a). However, the methods for multiple-bin spectral CT are relative scarce. Tao et al (Tao et al 2018) developed a prior knowledge aware iterative denoising material decomposition (MD-PKAD) model by considering the prior image constrained compressed sensing (PICCS) (Chen et al 2008). Xie et al proposed a multiple constraint model for image domain material decomposition and validated it only by a numerical phantom (Xie et al 2019). Besides, the deep learning based multi-material decomposition models (Clark et al 2018, Chen and Li 2019) are also studied. In this work, we will focus on the second step.

The learning-based methods have obtained a great achievement in the medical imaging field, such as low-dose reconstruction(Xu et al 2012b), sparse-view reconstruction (Li et al 2014, Gao et al 2019, Zhang et al 2020, Tao et al 2020, Xu et al 2020), medical image denoising (Li et al 2012), spectral CT reconstruction (Bo et al 2012), positron emission tomography (PET) imaging (da Costa-luis and Reader 2020, Fuller et al 2020, Wang et al 2020), magnetic resonance (MR) imaging (Ritzer et al 2020, Tao et al 2020, Mader et al 2020, Gravel et al 2020), etc. Especially, the dictionary learning-based methods can recover image details and features by training a large number of image patches. Similar to the hyperspectral image (Luo et al 2018), considering the correlation among different energy channels (i.e. image structure similarities of the same object), the tensor based dictionary learning was developed for spectral CT reconstruction, and it obtained excellent performance in our previous studies (Zhang et al 2017, Wu et al 2018). As for the multiple material decomposition of spectral CT, multiple material images can be considered as a tensor, leading to a natural idea to establish tensor based dictionary learning for image-domain material decomposition. However, since it is difficult to ensure a rank-1 tensor within a local region for multiple materials, it is inappropriate to adopt the tensor dictionary learning for material decomposition. Another idea is to develop a multiple-dictionary based decomposition method, i.e. each material corresponds to one dictionary. However, the correlation of material images will be ignored. Besides, because some materials only have simple structures (see iodine contrast agent in figure 1(c) in section 5), the trained dictionary cannot encode material information in the decomposition process. To overcome the aforementioned limitations, a unified dictionary is formulated by training image patches from mode-1 unfolding of normalized material image tensors. Then, a dictionary learning based image-domain material decomposition (DLIMD) model is constructed to fully encode the sparsity and similarities of material images simultaneously. Finally, the volume conservation and bound of each pixel are introduced into the DLIMD model to further improve material decomposition accuracy. The advantages of DLIMD method include good capabilities of noise suppression and edge preservation.

The rest of this paper is organized as follows. In section 2, the material decomposition model for spectral CT will be given. In section 3, we elaborate the mathematical model and solution for the DLIMD. In section 7, numerical mouse and two real datasets are employed to evaluate the proposed decomposition method. In section 5, we discuss some related issues and make a conclusion.
2. Material decomposition model for spectral CT

2.1. Spectral CT imaging

Assuming the detected photon number for the $\ell$th x-ray path within the $n$th energy window $E_n$ is $y_{\ell,n}$ ($1 \leq n \leq N, 1 \leq \ell \leq L$). It can be expressed as

$$y_{\ell,n} = \int_{E_n} I_{\ell,n}(E) e^{-\vartheta_{nm}(E,r) r} dE,$$

(1)

where $\int_{E_n} dE$ integrates over the range of $n$th energy channel and $\int_{E_n} dr$ indicates integral along the $\ell$th x-ray path. In this study, $M$ basis materials within the object is assumed. $x(N, E, r)$ represents linear attenuation coefficient for energy $E$ at position $r$. $I_{\ell,n}(E)$ represents the original x-ray photon intensity emitting from the x-ray source for energy $E$.

Considering the basis material expansion, the attenuation coefficient $x_n(E, r)$ in equation (1) can be further rewritten as a low-dimensional expansion. That is,

$$x_n(E, r) = \sum_{m=1}^{M} \vartheta_{nm}(E) f_m(r),$$

(2)

where $\vartheta_{nm}(E)$ represents the mass-attenuation coefficient for the $m$th material at $n$th energy window $E_n$ and $f_m(r)$ is fraction of $m$th material at location $r$. The task of material decomposition is to recover all the material component maps $f_m(r)$ ($1 \leq m \leq M$) from equation (1). If we directly reconstruct material maps from equation (1), it can be considered as direct material decomposition. In this study, we only consider the indirect material decomposition methods, i.e. image domain material decomposition, where the first step is to reconstruct $x_n(E, r)$. Let $I_{\ell,n}^{(0)}$ be the original x-ray photon flux which can be expressed as

$$I_{\ell,n}^{(0)} = \int_{E_n} I_{\ell,n}(E) dE.$$

(3)

Substituting equations (3) into (1), we have

$$y_{\ell,n} \approx I_{\ell,n}^{(0)} e^{-\int_{E_n} -x_n(E,r) dr}.$$

(4)

Considering the spectral CT imaging model, equation (4) can be discretized as

$$y_{\ell,n}/I_{\ell,n}^{(0)} \approx e^{-A_{\ell,n} x(E_n)}$$

(5)

where $A \in R^{L \times J}$ ($J = J_1 \times J_2$) is the projection matrix, $J_1$ and $J_2$ represent the width and height of the averaged attenuation coefficients image, $L$ is the number of total x-ray paths, and $A_{\ell,n}$ presents the $\ell$th row of $A$. $x(E_n) \in R^J$ is the vectorization of averaged attenuation coefficient image at $n$th energy window. Then, a logarithm operation is operated on both sides of equation (5). By ignoring the index $\ell$, equation (5) can be simplified as

$$p_n = -\ln \left( \frac{y_{\ell,n}}{I_{\ell,n}^{(0)}} \right) \approx A x_n.$$  

(6)

For the image-domain material decomposition method, we can reconstruct multi-energy CT images $x_n$ ($n = 1, \ldots, N$) from the corresponding projection $p_n$. There are numerous advanced multi-energy CT image reconstruction techniques, including tensor dictionary learning (Zhang et al 2017), non-local low-rank cube-based tensor factorization (NLCTF) (Wu et al 2019) and so on. Here, we only focus on image-domain material decomposition, and FBP and NLCTF are employed to obtain $\{x_n\}_n=1^N$.

2.2. Image-domain material decomposition model

To obtain the material component maps $f_m(r)$ ($1 \leq m \leq M$) from $\{x_n\}_n=1^N$, equation (2) can be further rewritten as a matrix form

$$\begin{bmatrix} \vartheta_{11} & \cdots & \vartheta_{1M} \\ \vdots & \ddots & \vdots \\ \vartheta_{NM} & \cdots & \vartheta_{NM} \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_M \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix},$$

(7)

where $\vartheta_{nm}$ represents averaged attenuation coefficients of the $m$th material at the $n$th energy window, which can be determined by calculating the averaged mass coefficients (Badea et al 2011). Equation (7) is equivalent to

$$\vartheta F(\lambda) = X(\lambda),$$

(8)
where $\vartheta = \begin{bmatrix} \vartheta_{11} & \cdots & \vartheta_{1M} \\ \vdots & \ddots & \vdots \\ \vartheta_{N1} & \cdots & \vartheta_{NM} \end{bmatrix} \in \mathbb{R}^{N \times M}$, $\mathbf{F} \in \mathbb{R}^{h \times h \times h}$ and $\mathbf{X} \in \mathbb{R}^{h \times h \times N}$ are two tensors which respectively represent the reconstructed images and material images. The mode-$h$ ($h = 1, 2, 3$) unfolding of tensor $\mathcal{M} \in \mathbb{R}^{h \times h \times h}$ is denoted by $\mathcal{M}_{(h)} \in \mathbb{R}^{h \times (\prod_{i \neq h} h)}$ (Zhang et al 2017, Wu et al 2018). Again, $\mathcal{F}_{(3)} \in \mathbb{R}^{M \times J}$ and $\mathcal{X}_{(3)} \in \mathbb{R}^{N \times J}$ are the mode-3 unfolding of $\mathbf{F}$ and $\mathbf{X}$. However, the reconstructed spectral CT images usually contain noise, which can compromise the decomposed material image quality. Considering the noise $\mathcal{Q}$ in the reconstructed images, equation (8) can be modified as

$$\vartheta \mathcal{F}_{(3)} = \mathcal{X}_{(3)} + \mathcal{Q}. \quad (9)$$

To recover material images from equation (9), we have the following least square linear program problem:

$$\mathcal{F}^* = \arg\min_{\mathcal{F}} \frac{1}{2} \| \mathbf{X}_{(3)} - \vartheta \mathcal{F}_{(3)} \|^2_F, \quad (10)$$

where $\| \cdot \|^2_F$ is Frobenius norm. Equation (10) can be viewed as an image domain basis material decomposition model. Without additional constraints, the simple direct inversion (DI) method (Niu et al 2014) can be employed to solve equation (10) and obtain:

$$\mathcal{F}_{(3)} = (\vartheta^T \vartheta)^{-1} \vartheta^T \mathbf{X}_{(3)}. \quad (11)$$

To improve the decomposed image quality, regularization is a feasible strategy to constrain the solution during iteration process, and equation (10) can be reformulated as

$$\mathcal{F}^* = \arg\min_{\mathcal{F}} \left\{ \frac{1}{2} \| \mathbf{X}_{(3)} - \vartheta \mathcal{F}_{(3)} \|^2_F + \lambda R(\mathcal{F}) \right\}, \quad (12)$$

where $\lambda$ is a regularization factor to balance the data fidelity term $\frac{1}{2} \| \mathbf{X}_{(3)} - \vartheta \mathcal{F}_{(3)} \|^2_F$ and regularization term $R(\mathcal{F})$.

3. Dictionary learning based image-domain material decomposition (DLIMD)

3.1. Dictionary learning

A set of image patches $\mathbf{x}_i \in \mathbb{R}^{s \times t}$, $i = 1, \ldots, I$, are extracted from the training datasets and unfolded as a vector to train the dictionary $\mathbf{D} \in \mathbb{R}^{S \times T}$, where $S = s \times s$ and $T$ represents the number of atoms. The goal of dictionary learning is to find sparse representation coefficients $\alpha \in \mathbb{R}^{T \times I}$ using the dictionary $\mathbf{D}$. This can be established by solving the following optimization problem:

$$\{\mathbf{D}^*, \alpha^*\} = \arg\min_{\mathbf{D}, \alpha} \sum_{i = 1}^{I} \| \mathbf{x}_i - \mathbf{D} \alpha_i \|^2_F \text{s.t.} \| \alpha_i \|_0 \leq L_0, \quad (13)$$

where $L_0$ is the sparsity level, $\| \cdot \|_0$ is the quasi-$l_0$ norm, $\alpha_i \in \mathbb{R}^{T \times 1}$ is sparse representation coefficients for $i$th image patch. Equation (13) is a constrained problem, which can be converted into the following problem:

$$\{\mathbf{D}^*, \alpha^*\} = \arg\min_{\mathbf{D}, \alpha} \left\{ \frac{1}{2} \sum_{i = 1}^{I} \left( \| \mathbf{x}_i - \mathbf{D} \alpha_i \|^2_F + v_i \| \alpha_i \|_0 \right) \right\}, \quad (14)$$

where $v_i$ is a Lagrange multiplier. Equation (14) can be solved by utilizing an alternating minimization scheme. The dictionary $\mathbf{D}$ is first fixed and the best coefficient matrix $\alpha$ can be found. As finding the truly optimal $\alpha$ is impossible, we use an approximation pursuit method. Any algorithm such as the matching pursuit (MP) (Mallat and Zhang 1993) or orthogonal matching pursuit (OMP) algorithm (Chen et al 1989) can be used for the calculation of the coefficients. The algorithm will not stop until the coefficient matrix and all the atoms have been updated and the stop criteria are satisfied. The first step is to update $\alpha_i$ by fixing the dictionary $\mathbf{D}$, $\alpha$ can be updated as

$$\alpha^* = \arg\min_{\alpha} \frac{1}{2} \sum_{i = 1}^{I} \left( \| \mathbf{x}_i - \mathbf{D} \alpha_i \|^2_F + v_i \| \alpha_i \|_0 \right). \quad (15)$$
The second step is to update the dictionary with a fixed sparse representation coefficients \( \alpha \). There are many methods available to train the dictionary, such as K-SVD (Aharon et al 2006) and online robust learning (Lu et al 2013). In this study, the K-SVD method is employed to train dictionary.

### 3.2. DLIMD method

Considering the general image domain material decomposition model equation (12), similar to the dictionary learning for low-dose CT reconstruction (Xu et al 2012b), the proposed dictionary learning based image-domain material decomposition (DLIMD) can be constructed as follow:

\[
\begin{align*}
\{ F^*, \{ \beta^*_m \}_{m=1}^M \} = \text{argmin}_{F, \{ \beta_m \}_{m=1}^M} \left( \frac{1}{2} \| X_{(j)} - \theta F_{(j)} \|_F^2 + \sum_{m=1}^M \lambda_m \frac{1}{2} \sum_{i=1}^J \left( \| H_{(i)} (F_m) - D \beta_{mi} \|_F^2 + \nu m \| \beta_{mi} \|_0 \right) \right),
\end{align*}
\]

where \( F_m \) is the \( m \)th material image \( F \), \( \beta_{mi} \in \mathcal{R}^{T \times I} \) are sparse representation coefficients for \( i \)th image patch from \( m \)th material image and \( \beta_m = \{ \beta_{mi} \}_{i=1}^J \), \( H_{(i)} (F_m) \) is the \( i \)th image patch extraction operator from \( F_m \). \( D \) is the trained dictionary. High-quality dictionary is beneficial for sparse representation and accuracy of material decomposition. A natural idea is to train different \( D \) for different materials. However, this strategy will be limited by three reasons. First, a specific material map may only contain a few image features, \( i.e \) bone and iodine contrast in figure 1. It is difficult for the trained \( D \) to encode the image features and reduce sparse representation level during the process of material decomposition, and this compromises the material decomposition accuracy. Second, training different \( D \) is time consuming, which implies higher computational cost is needed with various materials in practice. Third, if we train the dictionary \( D \) for each material, the correlation between different material maps will be lost. To overcome these limitations, we formulate a unified dictionary \( D \in \mathcal{R}^{K \times T} \) in this study. For a given spectral CT dataset, we can obtain the FBP or other reconstruction result and apply the DI method to decompose material image \( F \). Then, before training the dictionary \( D \), we should normalize material images to avoid data inconsistency of the pixel values from different materials by

\[
F_m' = \frac{F_m - \min (F_m)}{\max (F_m) - \min (F_m)}, \quad m = 1, \ldots, M.
\]

After that, a set of image patches are extracted from \( F_m' \in \mathcal{R}^{I \times (J_1 \times J_2)} \), the mode-\( l \) unfolding of normalized image \( F \), to train a global dictionary \( D \). The normalization is performed on each material image independently rather than the extracted image patches. Since the concentration of contrast agent usually is low, the whole image-based normalization can benefit to extract image patches from contrast agent component and then result in improved accuracy of contrast agent. A higher accuracy of contrast agent is good for other materials accuracy. However, if a patch-based normalization is employed, the image features from iodine contrast would be lost. It may compromise material decomposition accuracy.

Although the DI material maps contain noise, our following experiments demonstrate that the dictionary training can preserve image information and against noise well. Here, the dictionary \( D \) can be learned using the K-SVD algorithm (Aharon et al 2006). The advantages of the formulated unified dictionary \( D \) are threefold. First, it can fully encode image similarities within different materials images (for examples, the image structures within blue boxes from bone and soft tissue, the image structures within red boxes from soft tissue and iodine contrast in figure 1) during the material decomposition process. Second, it can enhance the redundancy with the trained dictionary \( D \). This is good to encode current material image structures by employing the atoms from other material images. Third, we can also save training time to some extent in practice.

To further improve the accuracy of material decomposition, more constraints should be considered. In clinical applications, the contrast agents are usually propagated into soft tissue rather than bone structures. Again, as a given pixel of material images, it is reasonable to assume that bone and contrast agent do not co-exist. Specifically, we first determine the locations belong to bone and contrast agents simultaneously. Then, we need to discriminate what component for these locations by comparing the bone component with a given threshold (for example, 0.5). If the value of referred location is greater than the given threshold, we will determine this location as bone rather than contrast agents. Because the contrast agents cannot be injected into such bone structures. This assumption is employed in our experiments. If air is also treated as one basis material, the summation of pixel values at the same location in different material images should be equal to one (Liu et al 2009), \( i.e \)

\[
\sum_{m=1}^M F_{jh,m} + B_{jh} = 1 (1 \leq j_1 \leq J_1, 1 \leq j_2 \leq J_2),
\]

where \( B_{jh} \) is the background pixel values for the \( j_1 \)th column of the \( j_2 \)th row in the image domain.
Figure 1. Numerically simulated mouse phantom used in (Zhang et al 2017) consists of three materials, i.e. bone, soft tissue and iodine. (a), (b) and (c) represent bone, soft tissue and iodine, respectively.

Figure 2. Setups of physical phantom experiments. (a) is the spectral CT system, (b) and (c) represents the physical phantom.

Figure 3. Physical phantom FBP reconstruction results. From left to right, the images are for 1st—4th energy bins with a display window [0 1.3] cm\(^{-1}\).

where \( \mathcal{F}_{j_1,j_2,m} \) represents the pixel value at the \((j_1,j_2,m)\) th location, \( \mathbf{B} \) is the air map and \( \mathbf{B}_{j_1,j_2} \) represents the pixel value at the \((j_1,j_2)\) th location. Besides, the pixel value within \( \mathcal{F} \) should be in the range of \([0 1]\), that is,

\[ 0 \leq \mathcal{F}_{j_1,j_2,m} \leq 1. \]  

To further constrain the solution for improving the accuracy of material decomposition, equations (18) and (19) can be treated as two conditions and are incorporated into the material decomposition model equation (16). Then, we can obtain equation (20). To solve equation (20), we first introduce a tensor \( \mathcal{U} \) to replace \( \mathcal{F} \), equation (20) can be converted into equation (21), and equation (21) can be further converted into equation(22). When the optimization factor \( \eta > 0 \) is increased to infinity, equation (22) is exactly equal to the constrained optimization problem equation (21). Similar to our previous study in (Wu et al 2018), equation (22) can be solved by the alternative direction minimization method (ADMM).
Figure 4. Material decomposition results from figure 3. From left to right, the columns represent the decomposed results of aluminum, water and iodine, where the display windows are [0.5 1], [0.8 1] and [0 0.003]. From top to bottom, the rows represent the results decomposed by DI, TVMD and DLIMD methods, respectively.

Figure 5. Preclinical experiment. (a) is the preclinical specimen fixed on the spectral CT system. (b)–(e) are FBP reconstruction results from 4 energy bins, where the display window is [0 0.5] cm$^{-3}$. 
Figure 6. Material decomposition results from figure 5. From left to right, the columns represent the DI, TVMD and DLIMD methods. The 1st–3rd rows represent the bone, soft tissue and iodine and 4th – 7th rows represent the magnified ROIs with ‘A’, ‘B’, ‘C’ and ‘D’. From 1st – 7th rows, the display windows are [0.25 0.5], [0.85 0.95], [0.000365 0.001], [0.29 0.33], [0.85 0.95], [0.85 0.95] and [0.85 0.95] respectively.
However, we adopt a simple strategy reported in (Xu et al 2019) for this study.

\[
\{F^*, \{\beta_m^*\}_{m=1}^M \} = \arg \min_{F, \{\beta_m\}_{m=1}^M} \left( \frac{1}{2} \left\| X(3) - \vartheta F(3) \right\|_F^2 + \sum_{m=1}^M \lambda_m \sum_{i=1}^I \left( \left\| \mathcal{H}_i(F_m) - D\beta_{mi} \right\|_F^2 + \nu_{mi} \|\beta_{mi}\|_0 \right) \right),
\]

s.t. \( \sum_{m=1}^M \mathcal{F}_{j_1,j_2,m} = 1, \forall j_1, j_2 \) \hspace{1cm} (20)

\[
\{F^*, \{\beta_m^*\}_{m=1}^M, U^* \} = \arg \min_{F, \{\beta_m\}_{m=1}^M, U} \left( \frac{1}{2} \left\| X(3) - \vartheta F(3) \right\|_F^2 + \sum_{m=1}^M \lambda_m \left( \sum_{i=1}^I \left( \left\| \mathcal{H}_i(U_m) - D\beta_{mi} \right\|_F^2 + \nu_{mi} \|\beta_{mi}\|_0 \right) \right) \right),
\]

s.t. \( \sum_{m=1}^M \mathcal{F}_{j_1,j_2,m} = 1, U = F, F \in [0, 1] \) \hspace{1cm} (21)
Figure 8. Material decomposition results of physical phantom assuming bone and contrast agent can co-exist in a given pixel in material decomposition models. 1st − 2nd rows represent the TVMD and DLIMD methods, respectively.

\[
\{\mathcal{F}^*, \{\beta_m^*\}_{m=1}^M, \mathcal{U}^*\} = \arg\min_{\mathcal{F}, \{\beta_m\}_{m=1}^M, \mathcal{U}} \left( \frac{1}{2} \left\| \mathcal{X}(3) - \partial_\mathcal{F}(3) \right\|^2_F + \frac{\eta}{2} \left\| \mathcal{U} - \mathcal{F} \right\|^2_F + \sum_{m=1}^M \left( \frac{j}{2} \left\| \mathcal{H}_i(\mathcal{U}_m) - \hat{D}_{\beta mi} \right\|^2_F + \gamma_m \left\| \beta_{mi} \right\|_0 \right) \right),
\]

s.t. \( \sum_{m=1}^M \mathcal{F}_{j_1j_2m} = 1, \beta_{mi} = 1, \mathcal{U}_m \in [0, 1] \).

The objective function equation (22) can be divided into the following two sub-problem:

\[
\mathcal{F}^* = \arg\min_{\mathcal{F}} \left( \left\| \mathcal{X}(3) - \partial_\mathcal{F}(3) \right\|^2_F + \frac{\eta}{2} \left\| \mathcal{U} - \mathcal{F} \right\|^2_F \right) \text{s.t. } \left( \sum_{m=1}^M \mathcal{F}_{j_1j_2m} \right) + B_{j_1j_2} = 1, 0 \leq \mathcal{F} \leq 1, \quad (23a)
\]

\[
\{\{\beta_m^*\}_{m=1}^M, \mathcal{U}^*\} = \arg\min_{\{\beta_m\}_{m=1}^M, \mathcal{U}} \left( \frac{\eta}{2} \left\| \mathcal{U} - \mathcal{F}^{(k+1)} \right\|^2_F + \frac{\lambda_m}{2} \sum_{i=1}^M \left( \left\| \mathcal{H}_i(\mathcal{U}_m) - \hat{D}_{\beta mi} \right\|^2_F + \gamma_m \left\| \beta_{mi} \right\|_0 \right) \right),
\]

(23b)

(i) \( \mathcal{F} \) sub-problem: Regarding the optimization of equation (23a), there are two strategies. More details can refer to appendix A. Here, we just employ the strategy 2 to update

\[
\mathcal{F}_{j_1j_2}^* = \arg\min_{\mathcal{F}_{j_1j_2}} \left( \frac{1}{2} \left\| \partial^T_{\mathcal{X}_j} - \eta \mathcal{F}_{j_1j_2} \right\|^2_F + \frac{\tau_{j_1j_2}}{2} \sum_{m=1}^M \left( \left\| \mathcal{H}_i(\mathcal{U}_m) - \hat{D}_{\beta mi} \right\|^2_F + \gamma_m \left\| \beta_{mi} \right\|_0 \right) \right) \text{s.t. } \left( \sum_{m=1}^M \mathcal{F}_{j_1j_2m} \right) = 1, 0 \leq \mathcal{F}_{j_1j_2}^* \leq 1. \quad (24)
\]

In fact, equation (24) is a general DI material decomposition mathematical model.

(ii) \( \mathcal{U} \) sub-problem: As for the optimization problem of equation (23b), it can be rewritten as

\[
\{\{\beta_m^*\}_{m=1}^M, \mathcal{U}^*\} = \arg\min_{\{\beta_m\}_{m=1}^M, \mathcal{U}} \left( \frac{1}{2} \left\| \mathcal{U}_m - \mathcal{F}^{(k+1)}_m \right\|^2_F + \frac{\tau_m}{2} \sum_{i=1}^M \left( \left\| \mathcal{H}_i(\mathcal{U}_m) - \hat{D}_{\beta mi} \right\|^2_F + \gamma_m \left\| \beta_{mi} \right\|_0 \right) \right).
\]

(25)
Algorithm 1. DLIMD.

**Input:** $\eta_1, \varepsilon, L, T, K$ and other parameters; Initialization of $\mathcal{F}^{(0)} = 0, \mathcal{U}^{(0)} = 0, k = 0$.

**Output:** Material decomposition tensor $\mathcal{F}$.

**Part I: Dictionary training**

1: Reconstructing spectral CT images using certain image reconstruction method;
2: Formulating the material attenuation matrix using reconstruction results;
3: Decomposing the reconstructed images using DI method;
4: Normalizing the DI results using equation (17);
5: Extracting image patches to form a dictionary training dataset;
6: Training a dictionary using the K-SVD technique.

**Part II: Material decomposition**

7: While not convergence do
8: Updating $\mathcal{F}^{(k+1)}$ using equation (24);
9: Finding and processing the pixels in $\mathcal{F}^{(k+1)}$ using DI technique;
10: Updating $\mathcal{U}$ and $\{\beta_m^k\}_{m=1}^M$ using equation (26);
11: $k = k + 1$;
12: End While

where $\tau_m = \lambda_m / \eta_1$, equation (25) can be divided into $M$ sub-problems. Again, $U_m (m = 1, \ldots, M)$ can be updated independently. Here, equation (25) can be read as

$$\begin{align*}
\left\{ U_m^*, \{\beta_m^k\}_{i=1}^I \right\} & = \argmin_{\{\beta_m^k\}_{i=1}^I, U_m} \left( \frac{1}{2} \left\| U_m - \mathcal{F}_m^{(k+1)} \right\|_F^2 + \frac{\tau_m}{2} \sum_{i=1}^I \left( \left\| \mathcal{H}_i(U_m) - D\beta_{mi}^{(k+1)} \right\|_F^2 + \nu_{mi}\|\beta_{mi}\|_0 \right) \right), \\
1 \leq m \leq M.
\end{align*}$$

Equation (26) can be solved by the method in (Wu et al. 2018). The parameters of sparsity level $L$ and tolerance of representation error $\varepsilon$ play important roles in controlling the dictionary quality and material decomposition accuracy. In this study, the sparsity level $L$ is set as the same value for different materials. However, the tolerance of representation error $\varepsilon$ is $\varepsilon_m (m = 1, \ldots, M)$ should be carefully chosen. Again, the $m$th basis material corresponds to $\varepsilon_m$. Here, the selection of $\varepsilon_m$ depends on specific material image. A smaller $\varepsilon_m$ can induce noise/false structures, while a greater one can lose some detailed structures (Tan et al. 2015, Wu et al. 2018). The overall pseudo-codes of the DLIMD algorithm is summarized in algorithm 1.

### 3.3. Implementation details

The DLIMD algorithm can be divided into two main parts: dictionary learning and material decomposition. To implement the dictionary learning, we first reconstruct spectral CT images from projections using analytic/iterative reconstruction methods. In this study, both FBP and NLCTF (Wu et al. 2019) methods are employed to implement image reconstruction. By carefully design the system matrix for iterative algorithms, both the FBP and regularization based iteration methods can obtain the same gray value values in terms of linear attenuation coefficient upon a constant scale. In fact, all of them cannot obtain linear attenuation coefficients directly without extra calibration operations (e.g. normalization with respect to Water). However, it does not affect the material decomposition results.

Calculating the basis material attenuation matrix by averaging the selected uniform regions is a key step for material decomposition. This is a common strategy to obtain basis material attenuation matrix in practice (Niu et al. 2014, Zhang et al. 2017). To demonstrate the procedure, the physical phantom (more details can refer to the following section 4.1) is taken as an example. Because the physical phantom contains iodine solution rather than pure iodine, we can only directly obtain attenuation coefficients of water and aluminum by averaging uniform regions, it is difficult to obtain the attenuation coefficient of iodine. Note that we can obtain the attenuation coefficient of water easily. According to figure 2(a), the physical phantom contains three different concentrations (i.e. $5 \text{ mg ml}^{-1}$, $10 \text{ mg ml}^{-1}$ and $15 \text{ mg ml}^{-1}$) of iodine. Three ROIs can be extracted from three different cylinders with iodine. From the extracted ROIs, we can obtain three attenuation coefficients of different iodine concentrations. Finally, we can compute the material attenuation matrix of physical phantom $\hat{\theta}_1$ by weighting. A natural question is that how to compute the matrix if we do not know which materials present in the object. In fact, the patient usually contains three materials, i.e. bone, soft tissue and contrast agent. We can decompose the contrast agent in practice even though we do not know the specific concentration (Wu et al. 2019). In addition, the calibrated phantom is also a common method for determining material attenuation coefficient matrix (Bateman et al. 2018).

Then, we adopt the DI method to decompose the reconstructed images into basis material maps using the formulated material attenuation matrix. Finally, $10^4$ image patches are extracted from normalized
material images to train the dictionary $\mathbf{D}$ using the K-SVD algorithm. Here, the sizes of extracted image patches are set as $10 \times 10, 8 \times 8$ and $8 \times 8$ for numerical phantom, physical phantom and preclinical experiments. The dictionary $\mathbf{D}$ is overcomplete to enforce the sparsity level. Again, the number of atoms in a dictionary should be much greater than the size of an atom, i.e. $T \gg s \times s$. Usually, $T$ should be greater than $4 \times s \times s$. Here, the number of atoms $T$ within the dictionary $\mathbf{D}$ is set as 512. The sparsity level $L$ in the dictionary training can be set empirically from 5 to 10, and it is uniformly set as 6. The iteration number to train dictionary is uniformly set as 200. Three dictionaries used in numerical mouse, physical phantom and preclinical datasets experiments are trained from their corresponding DI results. In other words, the dictionary is trained using its DI results merely. In fact, it is feasible to train dictionaries using its DI results in practice because we always can obtain its DI results. Besides, most of image structures of different material images are included in their DI results.

Particularly, equation (24) is solved by adopting the Isqlin function in Matlab 2014(b). The parameter selections of sparsity representation level $L$ and tolerance of representation error $\varepsilon = \varepsilon_m (m = 1, \ldots, M)$ depend on specific applications. All the methods are stopped after 30 iterations.

### 4. Experiments and results

To evaluate the performance of the proposed DLIMD method in material decomposition, a numerical mouse and two real datasets are employed. The advantages are demonstrated by comparing with the DI and total variation-regularized material decomposition (TVMD) methods. To quantitatively evaluate the performance of all the material decomposition techniques, the root means square error (RMSE), peak-signal-to-noise ratio (PSNR) and structural similarity (SSIM) are employed. The optimized parameters used in all experiments are summarized in table 1. The material number is set as 3 ($M = 3$) for all experiments. Here, the number of energy bins $N$ for numerical mouse, physical phantom and preclinical experiment is 4. The experiment results of numerical mouse can refer to appendix B.

### 4.1. Physical phantom

A physical phantom consisting of three basis materials, i.e. water, iodine and aluminum, is employed to evaluate our method. As shown in figure 2, the phantom contains five cylinders, and each of them represents different basis material or different concentration of iodine solution. Aluminum is used to mimic bone. The spectral CT system is equipped with a micro-focus x-ray source (YXLON, 225kV) and a flat-panel PCD (Xcounter, XC-Hydra FX20). For the central slice, the PCD contains 2048 detector cells and each of them covers a length of 0.1 mm. Here, every four cells are combined to reduce noise. Because the PCD only has two energy channels, the phantom is scanned twice to obtain four energy bin projections. The experiments is performed on an industrial CT system with high precision. If the parameters of scan configuration is fixed, the collection locations of photon counting detector is also fixed with high precision. Although there may be a very small variations between two scans, actually it is not necessary to calibrate the misregistration for two scans. The full-scan projections are collected from 1080 uniformly distributed views with 137kV source. The distances from the x-ray source to rotational center and detector are 182.68 mm and 440.50 mm, respectively. The field of view diameter is calculated as 82.6 mm. The reconstructed material image matrix is $512 \times 512$, and each pixel covers an area of $0.162 \times 0.162 \text{mm}^2$.

To evaluate the proposed material decomposition methods in the case of analytic reconstruction results, figure 3 shows the reconstructed spectral CT images from four different energy bins using FBP. This can comprise the material decomposition accuracy to some extent. The material decomposition results with three basis materials (aluminum, water and iodine) are given in figure 4. From figure 4, it can be observed that the material decomposition with regularization terms can also provide higher image quality than that obtained by the DI method. In terms of aluminum results, the proposed DLIMD method can better protect image edges and avoid the blocky artifacts compared with the TVMD method. As for the decomposed water results, the proposed DLIMD can provide much clearer image edge with higher accuracy, which is confirmed by the magnified regions of interest (ROI). Regarding the iodine results, three cylinders from the DLIMD are more complete and uniform than those obtained by other decomposition methods.
To further evaluate the performance of DLIMD algorithm in material decomposition improvement, five ROI indicated with 1–5 are extracted from different material maps and their quantitative evaluation results are listed in table 2. Since the concentrations of iodine contrast agent are designed in advance, the ground truth of different concentration can be estimated. The proposed method always performs the best followed by the DI and TVMD methods.

In this work, all the involved methods are programmed in Matlab (version 0020) on a PC (i7-6700, 8.0 GB memory). The proposed DLIMD algorithm can be divided into two subroutines: dictionary training and material decomposition. As for the computational cost of material decomposition, the higher the number of materials is, the greater the computational cost is. Particularly, the DI, TVMD and DLIMD consume 22.79, 24.16 and 44.17 s, respectively. The dictionary learning needs more time to implement sparse representation than others because its process refer to sparse encoding during the process of iteration. Compared with the DI, the DLIMD needs more time (387.22 s) to train the dictionary. Obviously, the DLIMD algorithm requires more computational cost than other algorithms. However, once the dictionary has been well trained, it can be used for material decomposition of similar objects.

4.2. Preclinical experiments

In this study, the PILATUS3 PCD with four energy-channels by DECTRIS is employed to acquire multi-energy projections in preclinical application. Such a PCD consists of 515 cells and each has a length of 0.15 mm. Here, a full-scan is performed to obtain projections with four energy bins from 720 views. Figure 5(a) shows the imaged object consisting of chicken foot and 5 mg ml$^{-1}$ iodine solution cylinder. The distances from x-ray source to rotational center and detector are 35.27 cm and 43.58 cm, respectively. The FBP reconstruction results (see figures 5(b)–(e)) have 512 × 512 pixels and each pixel covers an area of 0.122 × 0.122 mm$^2$.

To demonstrate the advantages of the DLIMD method, the material decomposition results from FBP images with different algorithms are shown in figure 6. Since the FBP reconstructed images may contains noise, noise suppression sometimes is a common strategy to improve the accuracy of material decomposition, including noise propagation multiple materials decomposition (Jiang et al 2019). Here, for bony components, the image edges of bony structure are blurred in the DI results. Although the TV based material decomposition method can provide clear image edges, the blocky artifacts appear in the component map. This point can be confirmed by the bony structure indicated with red arrow within the magnified bony ROI 'A' in figure 6. As for the soft tissue decomposition results, the DLIMD can obtain clear image structure and features. To clarify this point, three soft tissue ROIs marked with 'B', 'C' and 'D' are also extracted and magnified in figure 6. From figure 6, it can be easily seen that image edges and structures are remarkably clear than those obtained by the TVMD and DI methods. Similar to the soft tissue components, as for iodine contrast, the material images from all decomposition methods contain small ring artifacts while the ring artifacts removal method is used (Chang et al 2020). DI can wrongly classify the pixels of bone component into iodine contrast followed by the TVMD results. Obviously, the DLIMD method can further improve the accuracy of decomposed materials. This conclusion can be observed from the magnified ROI 'E' in the 3rd row of figure 6. According to figure 5, the iodine concentration is 5.0 mg ml$^{-1}$ and it can be considered as gold standard to evaluate the decomposition accuracy of different algorithms.

---

Table 2. Quantitative evaluation results of ROIs 1–5.

| ROI | RMSE(10$^{-4}$) | PSNR | SSIM |
|-----|----------------|------|------|
| DI  | 1006           | 19.952 | 0.9348 |
| TVMD| 944.6          | 20.496 | 0.9842 |
| DLIMD| 902.5       | 20.891 | 0.9923 |
| DI  | 355.4          | 28.983 | 0.9524 |
| TVMD| 321.8          | 29.848 | 0.9794 |
| DLIMD| 298.4        | 30.504 | 0.9953 |
| DI  | 5.537          | 65.134 | 0.4336 |
| TVMD| 2.979          | 70.520 | 0.6872 |
| DLIMD| 2.348        | 72.585 | 0.8271 |
| DI  | 7.445          | 62.563 | 0.6213 |
| TVMD| 3.093          | 70.193 | 0.9156 |
| DLIMD| 8.596        | 61.314 | 0.8220 |
Table 3. Quantitative evaluation results of ROI-6 (unit: mg/mL).

|         | DI   | TVMD   | DLIMD   |
|---------|------|--------|---------|
| FBP     | Mean | 1.864  | 2.282   | 2.283   |
|         | RMSE | 3.4088 | 2.7839  | 2.7621  |
| NLCTF   | Mean | 1.830  | 2.353   | 2.353   |
|         | RMSE | 3.2187 | 2.6917  | 2.6896  |

Advanced spectral CT image reconstruction methods can improve the accuracy of material decomposition by providing high quality reconstructed images (Wu et al. 2019). To further demonstrate the advantages of the proposed DLIMD algorithm in advanced reconstruction case, the NLCTF method is employed to implement image reconstruction and the corresponding material decomposition results are given in figure 7. Because the TVMD can obtain higher material decomposition accuracy than the DI method, only the decomposition results from TVMD and NLCTF methods are presented to save space. Although the TVMD method can provide higher quality for bone component map, the blocky artifacts still appear. This point can be confirmed by the magnified ROIs ‘F’ and ‘G’. From the 3rd and 4th columns in figure 7, it can be observed that the image structures indicated by ‘3’ and ‘4’ from the TVMD contain blocky artifacts. Fortunately, these structures are clear in the DLIMD results. Besides, the structures indicated by arrows ‘5’ and ‘6’ from the proposed DLIMD method are clearer than those obtained by the TVMD. As for the soft tissue decomposition results, the DLIMD can also obtain better image quality with clear image edges. However, these image structures are blurred in the TVMD results. Especially, two ROIs indicated by ‘H’ and ‘K’ are extracted and magnified in figure 7. The image feature indicated by arrow ‘2’ disappears in the TVMD. It is still available in the DLIMD result. The image structures labeled by arrows ‘7’ and ‘8’ further confirm these points. Finally, the RMSEs and mean value from location of ROI-6 are also calculated and listed in table 3. From table 3, one can see the DLIMD can obtain the smallest RMSE. In terms of mean value, the TVMD and DLIMD can obtain the same accuracy.

5. Discussions and conclusions

To improve the accuracy of material decomposition for spectral CT, we propose a DLIMD technique. Compared with the previous image-domain material decomposition method for spectral CT (Zhao et al. 2015), the innovations of DLIMD are threefold. First, considering the similarities of different material images, we construct a unified dictionary to encode material image sparsity by training a set of image patches. Here, those image patches are extracted from normalized material images. Second, we formulate a DLIMD mathematical model by enhancing sparsity of material maps with the dictionary and analyzing the image-domain material decomposition. Third, additional constraints (volume conservation (Liu et al. 2009) and the bound of each material pixel value) are incorporated into basis material decomposition model to further improve the decomposition accuracy. The experiments demonstrate the advantages of DLIMD in improving image quality and material decomposition accuracy.
The beam hardening correction plays an important role in spectral CT imaging. Because the x-ray photons within lower energy bins imply low energy, they can be easily attenuated and result in beam hardening effects. However, our focus is to improve image-domain material decomposition rather than remove beam hardening artifacts. How to reduce the beam hardening artifacts is an interesting topic and it should be given a great attention in our future work. As a material decomposition method in image domain, beam hardening artifact cannot be avoided. Figure 8 shows the decomposed material results with no constraint over the model. It can be seen that the effect of x-ray beam hardening, many pixels of aluminum can be wrongly classified into iodine. Again, considering the referred constraints into the TVMD and DLIMD methods, the artifacts from x-ray beam hardening can be reduced and the material accuracy can be improved.

In practice, the material decomposition matrix \( \vartheta \) needs to be predetermined from known phantom before scan an unknown object. In our phantom experiments, the imaging object and the corresponding basis materials are known. We determined the matrix \( \vartheta \) from several uniform small ROIs that can serve as the known information of basis materials. Then, this matrix is used for material decomposition for the whole image object where the regions out of the ROIs can serve as the test object. In other words, in the same scanning setting, our phantom can be divided into two parts: one is for the determination of the matrix \( \vartheta \) with known basis materials, the other is treated as unknown object for the evaluation of material decomposition, and there is no overlap between the two parts. Because the reconstructed spectral CT images usually contain noise and artifacts, they can cause the instabilities of image domain material decomposition. Particularly, the main goal of material decomposition is to decompose more accurate material component. If the number of materials \( M \) is smaller than the number of energy-bin \( N \), the image-domain material decomposition model would search an over-determined solution for equation (8). Otherwise, equation (8) would be an ill-posed inverse problem. An appropriate energy-bin number with good threshold setting can benefit material decomposition accuracy improvement, radiation dose reduction, and image quality improvement. In fact, the number and threshold of energy bins are designed by specific imaging applications and the associated material decomposition algorithms. Again, we need to adjust the number and the corresponding energy threshold by material components within imaged object. This is a continuous procedure for specific applications. In this study, we mainly focus on three material decomposition of four energy bins without consideration of energy thresholds optimization. In addition, the attenuation coefficient matrix is manually computed by averaging uniform regions of specific basis materials. Due to the effects of charge sharing, K-edge, fluorescence x-ray emission and pulse pileups during projection collection, the linear fidelity term of equation (7) may be inaccuracy in practice. In these cases, the regularization prior in the material decomposition basic model can benefit to improve material decomposition accuracy. The proposed DLIMD can encode the image features and reduce sparse representation level during the course of material decomposition using a unified dictionary rather than split dictionary to improve material decomposition accuracy in practice. Meanwhile, training different dictionaries for different materials is time consuming, which implies higher computational cost for dictionary training can be reduced in practice. Compared with other regularization prior based decomposition methods, the correlation between different material maps can be explored, which benefit image edge preservation and detail recovery in practice. Although the DLIMD method can obtain satisfied decomposition results in image-domain, there are still some remaining issues that should be addressed. First, the parameters are only empirically optimized by comparing extensive experiment results. Regarding the selection of \( \varepsilon \), it is necessary to select an \( \varepsilon_m \) for each material component independently. However, if \( m \) is great, it is difficult to select an appropriate \( \varepsilon \). Again, when the image structure is simple, the function of \( \varepsilon \) mainly focus on noise suppression and a relative larger value should be given to reduce noise. When we focus on the structure and details, a smaller \( \varepsilon \) value can be chosen to recover
image edges and profiles. In fact, the distribution of contrast agents and bone are usually relative simple in terms of clinical application, a small $\varepsilon$ can be selected to recover image profile and preserve image edge. Regarding as soft tissue component, $\varepsilon$ should be optimized carefully depending on clinical applications due to its structure and details are complicated. In this case, it is feasible to select a reasonable $\varepsilon$ by trying some automatic strategies to obtain optimization. Second, two real datasets only contain three different basis materials. However, the imaging objects may contain multiple (greater than 3) materials. It is necessary to validate and evaluate the performance of DLIMD in such cases. Third, as a material decomposition method in image domain, the DLIMD cannot avoid the beam hardening artifacts. Indeed, we can still observe small

Figure B2. Material decomposition results of the numerical mouse phantom. The 1st–3rd columns represent bone, soft tissue and iodine contrast agent, where the display windows are [0.01 0.2], [0.05 0.55] and [0.0007 0.003]. The 1st row represent the ground truth, and 2nd–4th rows are material decomposition results using DI, TVMD and DLIMD methods, respectively.
beam hardening artifacts in the physical phantom study. To further improve the accuracy of materials decomposition with beam hardening artifacts reduction, advanced robust beam hardening reduction methods can be incorporated by modeling the x-ray spectrum (Zhao et al 2018). Fourth, material cross-talks can be observed within the material images by according to the material decomposition results, the block-matching technique by clustering pixel patches into groups may be employed to reduce the inaccuracy in our future plan (Xue et al 2019). Fifth, since the manufacture skill of photon counting detector is immature, ring artifacts usually appear. Although the ring artifacts has been reduced by utilizing a Fourier filtering method (Münch et al 2009), ring artifacts still appear. How to remove ring artifacts by employing deep learning is listed on our schedule (Chang et al 2020).

In conclusion, based on the dictionary learning theory, we proposed a DLIMD method for image-domain material decomposition of spectral CT. Real dataset experiments demonstrate the outperformance of the DLIMD technique. This will be useful for image-domain material decomposition of spectral CT.

**Strategy 1:** The air can be treated as a basis material belongs to $\mathcal{F}$. In this case, the material attenuation matrix $\vartheta$ can be modified as $\vartheta' = \begin{bmatrix} \vartheta_{11} & \cdots & \vartheta_{1M} \\ \vdots & \ddots & \vdots \\ \vartheta_{N1} & \cdots & \vartheta_{NM} \end{bmatrix} \in \mathcal{R}^{N \times (M+1)}$, where $\vartheta_{ij}$ is a small positive value to represent air attenuation coefficient. The air can increase the number of basis materials in this strategy, and it can result in instability of material decomposition and compromise the decomposition accuracy, especially in the case of $M > N$.

**Strategy 2:** In this method, equation (23a) can be divided into two steps. The first step is to solve the objection function equation (A.1):

$$
\mathcal{F}^{(k+1)} = \arg\min_{\mathcal{F}} \left( \frac{1}{2} \left\| \mathcal{X}_3 - \vartheta \mathcal{F}_3 \right\|_F^2 + \frac{\eta}{2} \left\| \mathcal{F} - \mathcal{U}^{(k)} \right\|_F^2 \right) s.t. \left( \sum_{m=1}^{M} \mathcal{F}_{j; i; m} = 1, \forall j, i, 0 \leq \mathcal{F} \leq 1. \right) \tag{A.1}
$$

The second step is to find the air region of $\mathcal{F}^{(k+1)}$ by operating a threshold method on it. Again, if a pixel value within $\mathcal{F}^{(k+1)}$ is greater than the given threshold, the pixel value can be computed by DI. The given threshold is set as 0.99 in this work.

To compare strategies 1 and 2, figure A1 shows the material decomposition results using DI of FBP images from noise-free projections. From figure A1, it can be seen that soft tissue results using strategy 1 contain severe strip artifacts. However, the artifacts are suppressed by using strategy 2.

Now, let us discuss the details of solving equation (A.1). Considering equation (A.1) with pixel level, it can be rewritten as

$$
\mathcal{F}^* = \arg\min_{\mathcal{F}} \left( \frac{1}{2} \sum_{j=1}^{J} \left\| \mathcal{X}_{(3)}_{j} - \vartheta \mathcal{F}_{(3)}_{j} \right\|_F^2 + \frac{\eta}{2} \sum_{j=1}^{J} \sum_{j=1}^{J} \left\| \mathcal{F}_{j;i;#} - \mathcal{U}_{j;i;#}^{(k)} \right\|_F^2 \right) s.t. \left( \sum_{m=1}^{M} \mathcal{F}_{j;i; m} = 1, \forall j, i, 0 \leq \mathcal{F}_{j;i;#} \leq 1. \right) \tag{A.2}
$$

where $\mathcal{X}_{(3)}_{j} = [\mathcal{X}_{j,1}, \cdots, \mathcal{X}_{j,N}]^T$, $\mathcal{F}_{(3)}_{j} = [\mathcal{F}_{j,1}, \cdots, \mathcal{F}_{j,M}]^T$ and $\mathcal{U}_{j;i;#}^{(k)} = [\mathcal{U}_{j;i;1}^{(k)}, \cdots, \mathcal{U}_{j;i;M}^{(k)}]^T$. Noting that $j = j_1 \times j_2$, equation (A.2) is equivalent to

$$
\mathcal{F}^* = \arg\min_{\mathcal{F}} \left( \frac{1}{2} \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} \left\| \mathcal{X}_{j;#} - \vartheta \mathcal{F}_{j;#} \right\|_F^2 + \frac{\eta}{2} \left\| \mathcal{F}_{j;#} - \mathcal{U}_{j;#}^{(k)} \right\|_F^2 \right) s.t. \left( \sum_{m=1}^{M} \mathcal{F}_{j;#} = 1, 0 \leq \mathcal{F}_{j;#} \leq 1 \forall j_1, j_2. \right) \tag{A.3}
$$

We optimize equation (A.3) by minimizing the following problem:

$$
\mathcal{F}^*_{j;#} = \arg\min_{\mathcal{F}_{j;#}} \left( \frac{1}{2} \left\| \mathcal{X}_{j;#} - \vartheta \mathcal{F}_{j;#} \right\|_F^2 + \frac{\eta}{2} \left\| \mathcal{F}_{j;#} - \mathcal{U}_{j;#}^{(k)} \right\|_F^2 \right), \forall j_1, j_2 s.t. \left( \sum_{m=1}^{M} \mathcal{F}_{j;#} = 1, 0 \leq \mathcal{F}_{j;#} \leq 1. \right) \tag{A.4}
$$

Equation (A.4) is a constrained convex programmable optimization problem. It is equal to the following optimization problem
Table B1. Quantitative evaluation results of three basis materials.

|                    | RMSE($10^{-2}$) | PSNR   | SSIM   |
|--------------------|-----------------|--------|--------|
| Bone               |                 |        |        |
| DI                 | 4.200           | 27.542 | 0.9812 |
| TVMD               | 4.070           | 27.812 | 0.9848 |
| DLIMD              | 3.745           | 28.529 | 0.9873 |
| Soft tissue        |                 |        |        |
| DI                 | 10.643          | 19.459 | 0.7363 |
| TVMD               | 8.103           | 21.828 | 0.7688 |
| DLIMD              | 8.051           | 21.883 | 0.7900 |
| Iodine contrast agent |               |        |        |
| DI                 | 0.0918          | 60.740 | 0.9553 |
| TVMD               | 0.0528          | 65.547 | 0.9595 |
| DLIMD              | 0.0521          | 65.671 | 0.9728 |

Equation (A.5) is a constrained least square problem, which can be easily solved.

\[
F_{j_{1}j_{2}#}^{*} = \arg\min_{F_{j_{1}j_{2}#}} \left( \frac{1}{2} \left\| \left( \vartheta^{T} \vartheta + \eta I \right) F_{j_{1}j_{2}#} - \left( \vartheta^{T} X_{j_{1}j_{2}#} + \eta \varphi_{j_{1}j_{2}#}^{(k)} \right) \right\|_{F}^{2} \right), \forall j_{1}, j_{2} s.t. \left( \sum_{m=1}^{M} F_{j_{1}j_{2}m} \right) = 1, 0 \leq F_{j_{1}j_{2}#} \leq 1. \tag{A.5}
\]

A numerical mouse thorax phantom with a 1.2% iodine contrast agent is employed for simulation study. This phantom consists of three basic materials, i.e. bone, soft tissue and iodine contrast agent, as shown in figure B1(a). A polychromatic 50KVp x-ray source is assumed and divided into eight energy bins ([16, 25) keV, [25, 31) keV, [31, 37) keV, [37, 50) keV]. A PCD consisting of 512 cells (each element covers 0.1 mm) is used to collect 640 × 4 projections over a full scan. The distances starting from source to PCD and object are 180 mm and 132 mm, respectively. In this study, the photon number of each emitting x-ray path is $10^{4}$ and Poisson noise is imposed. The reconstructed spectral CT images by FBP consist of $512 \times 512 \times 4$ pixels and three of them are shown in figures B1(b)(d).

To evaluate the proposed material decomposition method in the case of numerical mouse, the material decomposition results using DI, TVMD and DLIMD with three basis materials (bone, soft tissue and iodine contrast agent) are given in figure B2. In terms of bone results, many pixels of iodine contrast agent are wrongly classified into bone in the DI results. Compared with the TVMD method, the DLIMD method can better preserve image edges. As for the decomposed soft tissue results, the proposed DLIMD can provide much clearer image edge with higher accuracy, which can be observed in 2nd column of figure B2. Compared with the DI and TVMD methods, our proposed method can provide higher accuracy of iodine contrast agent and it can be observed in the 3rd column of figure B2.

To further evaluate the performance of DLIMD algorithm in improving material decomposition, the quantitative evaluation results from three basis material are listed in table B1. From table B1, we can see the proposed DLIMD can always obtain the smallest RMSE value with the highest PSNR and SSIM values.

ORCID iDs

Weiwen Wu  https://orcid.org/0000-0002-8295-5104
Fenglin Liu  https://orcid.org/0000-0002-8952-311X
Qian Wang  https://orcid.org/0000-0001-8988-126X
Yanbo Zhang  https://orcid.org/0000-0002-7627-488X
Hengyong Yu  https://orcid.org/0000-0002-5852-0813

References

Aharon M, Elad M and Bruckstein A 2006 K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation IEEE Trans. Signal Proc. 54 6311–22
Badea C, Johnston S, Qi Y, Ghaghada K and Johnson G 2011 Dual-energy micro-CT imaging for differentiation of iodine- and gold-based nanoparticles Proc. SPIE 7961 79611X
Barber R F, Sidky E Y, Schmidt T G and Pan X 2016 An algorithm for constrained one-step inversion of spectral CT data Phys. Med. Biol. 61 3784–818
Bateman C, Knight D, Brandwacht B, Mc Mahon J, Healy J, Panta R, Aamir R, Rajendran K, Moghiseh M and Ramyar M 2018 MARS-MD: rejection based image domain material decomposition J. Instrum. 13 P05020
Bo Z, Huanjun D, Yang L, Ge W, Jun Z and Sabee M 2012 Dual-dictionary learning-based iterative image reconstruction for spectral computed tomography application Phys. Med. Biol. 57 8217
Chang S, Chen X, Duan J and Mou X 2020 A CNN based hybrid ring artifact reduction algorithm for CT images IEEE Trans. Radiat. Plasma Med. Sci. (https://doi.org/10.1109/TRPMS.2020.2983391)
Chen G H, Tang J and Leng S 2008 Prior image constrained compressed sensing (PICCS) Med. Phys. 35 660
Chen S, Billings S A and Luo W 1989 Orthogonal least squares methods and their application to non-linear system identification Int. J. Control 50 1873–96
Chen Z and Li L 2019 Robust multitemporal decomposition of spectral CT using convolutional neural networks Opt. Eng. 58 013104
Clackdoyle R and Neof 2019 Quantification of tomographic incompleteness in cone-beam reconstruction IEEE Trans. Radiat. Plasma Med. Sci. 4 63–80
Clark D P, Hollbrook M and Badea C T 2018 Multi-energy CT decomposition using convolutional neural networks Proc. SPIE 10573 105731O
da Costa-Izquierdo C O and Reader A J 2020 Micro-networks for robust MR-guided low count PET imaging IEEE Trans. Radiat. Plasma Med. Sci. (https://doi.org/10.1109/TRPM2020.2986414)
Fuller O K, Angelis G I and Meikle S R 2020 Classification of neurotransmitter response in dynamic PET data using machine learning approaches IEEE Trans. Radiat. Plasma Med. Sci. 4 708–19
Gao Y, Liang Z, Zhang H, Yang J, Ferretti J, Billinger T, Yaddanapudi K, Schweiter M, Bhattacharji P and Moore W 2019 A task-dependent investigation on dose and texture in CT image reconstruction IEEE Trans. Radiat. Plasma Med. Sci. 4 441–9
Gravel P, Li Y and Matej S 2020 Effects of TOF resolution models on edge artifacts in PET reconstruction from limited-angle data IEEE Trans. Radiat. Plasma Med. Sci. 4 603–12
Jiang Y, Xue Y, Liu X, Luo C, Yang P, Yang C, Wang J, Hu X and Zhang X 2020 Noise suppression in image-domain multi-material decomposition for dual-energy CTs IEEE Trans. Biomed. Eng. 67 323–35
Johnson T R, Krauss B, Sedlmair M, Graasnick M, Bruder H, Morhard D, Fink C, Weckbach S, Lenhard M and Schmidt B 2007 Material differentiation by dual energy CT: initial experience Eur. Radiol. 17 1510–7
Kim K, Ye J C, Worstell W, Ouyang Y, Rakvongthai Y, El Fakhri G and Li Q 2015 Sparse-view spectral CT reconstruction using spectral patch-based low-rank penalty IEEE Trans. Med. Imaging 34 748–60
Li S, Cao Q, Chen Y, Hu Y, Luo L and Toumoulion C 2014 Dictionary learning based sinogram inpainting for CT sparse reconstruction Optik-Int J. Light Electron Opt. 125 2862–7
Li S, Yin H and Fang L 2012 Group-sparse representation with dictionary learning for medical image denoising and fission IEEE Trans. Biomed. Eng. 59 3450–9
Li Z, Ravishankar S, Long Y and Fessler J A 2019B DECT-MULTA: dual-energy CT image decomposition with learned mixed material models and efficient clustering (arXiv: 1901.00106)
Lin X Z, Wu Z Y, Tao R, Guo Y, Li J Y, Zhang J and Chen K M 2012 Dual energy spectral CT imaging of insulinoma—value in preoperative diagnosis compared with conventional multi-detector CT Eur. J. Radiol. 81 2487–94
Liu J, Ding H, Moll S, Zhang X and Gao H 2016 TCMR: total image constrained material reconstruction via nonlocal total variation regularization for spectral CT IEEE Trans. Med. Imaging 35 2578–86
Liu X, Yu L, Primak A N and McCollough C H 2009 Quantitative imaging of element composition and mass fraction using dual-energy CT: three-material decomposition Med. Phys. 36 1662–9
Long Y and Fessler J A 2014A Multi-material decomposition using statistical image reconstruction for spectral CT IEEE Trans. Med. Imaging 33 1614–26
Lu C, Shi J and Jia J 2013 Proc. IEEE Conf. on Computer Vision and Pattern Recognition pp 415–22
Luo F, Du B, Zhang L, Zhang I and Tao D 2018 Feature learning using spatial-spectral hyperspectral discrimination analysis for hyperspectral image IEEE Trans. Cybern. 49 2406–19
Maaß C, Baer M and Kachelrieß M 2009 Image-based dual energy CT using optimized precorrection functions: A practical new approach of material decomposition in image domain Med. Phys. 36 3818–29
Macdonald L R, Lo J Y, Sturgeon G M, Zeng C, Harrison R L, Kinahan P E and Segars W P 2020 Impact of using uniform attenuation models and efficient clustering (arXiv: 1901.00106)
Mather C E, Fuchs T, Ferraro D A and Burger I A 2020 Potential clinical applications of PET/MR IEEE Trans. Radiat. Plasma Med. Sci. 4 293–9
Mallat S G and Zhang Z 1993 Matching pursuits with time-frequency dictionaries IEEE Trans. Signal Process. 41 3397–415
Münch B, Trittik P, Marone F and Stampari M 2009 Stripe and ring artifact removal with combined wavelet—Fourier filtering Opt. Express 17 8567–91
Niu T, Dong X, Petrongolo M and Zhu L 2014 Iterative image-domain decomposition for dual-energy CT Med. Phys. 41 041901
Ritter C, Becker R, Buck A, Commichau V, Debus J, Djambazov L, Eleftheriou A, Fischer J, Fischer P and Ito M 2020 Initial characterisation of the SAFIR prototype PET-MR scanner IEEE Trans. Radiat. Plasma Med. Sci. 4 613–21
Saito H et al 2018 Reduced iodinated contrast media for abdominal imaging by dual-layer spectral detector computed tomography for patients with kidney disease Radiol. Case Rep. 13 437–43
Shikhaliyev P M and Fritz S G 2011 Photon counting spectral CT versus conventional CT: comparative evaluation for breast imaging reconstruction Phys. Med. Biol. 56 1905
Taguchi K and Iwanczyk J S 2013 Vision 20/20: single photon counting x-ray detectors in medical imaging Opt. Eng. 52 100901
Tan S, Zhang Y, Wang G, Mou X, Cao G, Wu Z and Yu H 2015 Tensor-based dictionary learning for dynamic tomographic reconstruction Phys. Med. Biol. 60 2803
Tao L, Fisher J, Anaya E, Li X and Levin C S 2020 Pseudo CT image synthesis and bone segmentation from MR images using adversarial networks with residual blocks for MR-based attenuation correction of brain PET data IEEE Trans. Radiat. Plasma Med. Sci. (https://doi.org/10.1109/TRPM2020.2989073)
Tao S, Rajendran K, Mccollough C H and Leng S 2018 Material decomposition with prior knowledge aware iterative denoising (MD-PKAID) Phys. Med. Biol. 63 195003
Wang B, Ruan D and Liu H 2020 Noninvasive estimation of macro-parameters by deep learning IEEE Trans. Radiat. Plasma Med. Sci. 4 684–95
Wang Z, Wang Y and Yu H 2017 Locally linear constraint based optimization model for material decomposition Phys. Med. Biol. 62 8314
Wu W, Liu E, Zhang Y, Wang Q and Yu H 2019 Non-local low-rank cube-based tensor factorization for spectral CT reconstruction IEEE Trans. Med. Imaging 38 1079–93
Wu W, Yu H, Wang S and Liu F 2017 BFP-type region-of-interest reconstruction for parallel translational computed tomography J. X-Ray Sci. Technol. 25 487–504
Wu W, Zhang Y, Wang Q, Liu E, Chen P and Yu H 2018 Low-dose spectral CT reconstruction using image gradient ℓ0-norm and tensor dictionary Appl. Math. Modell. 63 338–57
Xiao D, Zhang X, Li J, Bao N and Kang Y 2020 Low-Dose X-ray CT reconstruction algorithm using shearlet sparse regularization J. Med. Imaging Health. Inf. 10 620–7

Xiao D, Zhang X, Yang Y, Guo Y, Bao N and Kang Y 2018 Low-dose x-ray computed tomography reconstruction using curvelet sparse regularization J. Med. Imaging Health. Inf. 8 1665–72

Xie B, Su T, Kafandjian V, Niu P, Yang F, Robini M, Zhu Y and Duvauchelle P 2019 Material decomposition in x-ray spectral CT using multiple constraints in image domain J. Nondestr. Eval. 38 16

Xu J, Zhao Y, Li H and Zhang P 2019 An image reconstruction model regularized by edge-preserving diffusion and smoothing for limited-angle computed tomography Inverse Probl.

Xu M, Hu D, Luo F, Liu F, Wang S and Wu W 2020 Limited angle X-ray CT reconstruction using image gradient $\ell^0$ norm with dictionary learning IEEE Trans. Radiat. Plasma Med. Sci. (https://doi.org/10.1109/TRPMS.2020.2991887)

Xu Q, Yu H, Bennett J, He P, Zainon R, Doesburg R, Opie A, Walsh M, Shen H and Butler A 2012a Image reconstruction for hybrid true-color micro-CT IEEE Trans. Biomed. Eng. 59 1711–9

Xu Q, Yu H, Mou X, Zhang L, Hsieh J and Wang G 2012b Low-dose X-ray CT reconstruction via dictionary learning IEEE Trans. Med. Imaging 31 1682–97

Xue Y, Jiang Y, Yang C, Lyu Q, Wang J, Luo C, Zhang L, Desrosiers C, Feng K and Sun X 2019 Accurate multi-material decomposition in dual-energy CT: a phantom study IEEE Trans. Comput. Imaging 5 515–29

Xue Y, Ruan R, Hu X, Kuang Y, Wang J, Long Y and Niu T 2017 Statistical image-domain multimaterial decomposition for dual-energy CT Med. Phys. 44 886–901

Zhang F, Zhang M, Qin B, Zhang Y, Xu Z, Liang D and Liu Q 2020 REDAEP: robust and enhanced denoising autoencoding prior for sparse-view CT reconstruction IEEE Trans. Radiat. Plasma Med. Sci. (https://doi.org/10.1109/TRPMS.2020.2989634)

Zhang X, Xiao D and Kang Y 2019 Quantitative reconstruction of dual K-edge contrast agent distribution for photon-counting computed tomography IEEE Access 7 124441–51

Zhang Y, Mou X, Wang G and Yu H 2017 Tensor-based dictionary learning for spectral CT reconstruction IEEE Trans. Med. Imaging 36 142–54

Zhao W, Li D, Niu K, Qin W, Peng H and Niu T 2018 Robust beam hardening artifacts reduction for computed tomography using spectrum modeling IEEE Trans. Comput. Imaging 5 333–42

Zhao W, Niu T, Xing L, Xie Y, Xiong G, Elmore K, Zhu J, Wang L and Min J K 2016 Using edge-preserving algorithm with non-local mean for significantly improved image-domain material decomposition in dual-energy CT Phys. Med. Biol. 61 1332

Zhao Y, Zhao X and Zhang P 2015 An extended algebraic reconstruction technique (E-ART) for dual spectral CT IEEE Trans. Med. Imaging 34 761–8