

LEARNED MIXED MATERIAL MODELS FOR EFFICIENT CLUSTERING BASED DUAL-ENERGY CT IMAGE DECOMPOSITION

Zhipeng Li[†], Saiprasad Ravishankar^{*}, Yong Long[†], Jeffrey A. Fessler^{*}

[†] University of Michigan - Shanghai Jiao Tong University Joint Institute,
Shanghai Jiao Tong University, Shanghai, China

^{*}Department of Electrical Engineering and Computer Science, University of Michigan, MI, USA

ABSTRACT

Penalized weight-least squares (PWLS) with basis material priors is a promising way to achieve high quality material decompositions for Dual-energy CT (DECT). This paper proposes a new method dubbed DECT-MULTRA for image domain DECT material decomposition that combines conventional PWLS estimation with regularization based on a mixed union of learned transforms (MULTRA) model. Our approach pre-learns from training data a common union of unitary transforms for all the basis materials' patches, as well as a cross-material union of unitary transforms that captures relationships between the different basis material images. The proposed DECT-MULTRA algorithm efficiently obtains material decompositions by alternating between updating the material images and performing clustering of patches in the MULTRA model. Both these steps of the alternating algorithm have closed-form updates. Numerical experiments with the XCAT phantom show that the proposed method significantly improves image quality compared to the recent DECT-ST method that learns different sparsifying transforms for different basis materials and the DECT-EP approach that uses a non-adaptive edge-preserving hyperbola regularizer.

Index Terms— Image domain decomposition, Machine learning, Non-convex optimization, Transform learning, Cross-material models.

1. INTRODUCTION

Dual-energy CT (DECT) has been increasingly used in many clinical applications and industrial applications, such as kidney stone characterization, iodine quantification, and security inspection. It has been a valuable tool because of its ability for tissue characterization and material differentiation.

Methods for material decomposition for DECT can be characterized into direct decomposition [1], projection domain [2], and image domain methods. Compared to direct decomposition methods, image domain methods are more computationally efficient as they do not need to perform expensive forward and back-projection operations. Projection-domain and direct decomposition methods require sinograms or pre-log measurements that usually are not available on commercial DECT scanners. Image-domain methods produce basis

material images directly from attenuation images at low and high energies, but their efficacy may be limited due to sensitivity to noise and artifacts. Extracting prior information from big datasets is a promising way to strengthen image domain DECT decomposition methods.

Various priors such as dictionary models, sparsifying transform models, and manifold models could be learned from datasets to characterize CT images and basis material images. Dictionary learning methods have recently shown promise for low-dose CT (LDCT) image reconstruction [3] and spectral CT reconstruction [4, 5, 6, 7]. But the sparse coding step in the dictionary model and its learning is typically computationally expensive. Recently, Ravishankar et al. [8] proposed learning sparsifying transform (ST) models (generalized analysis models). The ST learning algorithms are computationally efficient with simple thresholding-based sparse coding for the transform model. Zheng et al. [9] demonstrated that pre-learned square sparsifying transforms lead to reduced noise for LDCT reconstruction compared to nonadaptive methods. More recently, Zheng et al. [10] extended the single learned ST approach to a union of learned transforms (ULTRA) model for LDCT reconstruction. That approach jointly pre-learns a collection of transforms such that each image patch is assigned to a corresponding best-matched transform (or cluster) during training and later during image reconstruction.

This paper proposes a basis material decomposition approach for DECT with regularization based on a mixed union of learned transforms (MULTRA) model that captures both the common properties among basis materials and the cross-dependency between materials. Numerical experiments with the XCAT phantom show that the proposed method, dubbed DECT-MULTRA, significantly improves the material images' quality compared to the very recent DECT-ST method that uses different pre-learned sparsifying transforms for different basis materials and the DECT-EP approach that uses a non-adaptive edge-preserving hyperbola regularizer.

2. PROBLEM FORMULATION

For dual energy CT, we denote the stacked two channel high and low energy maps as $\mathbf{y} = (\mathbf{y}_H^T, \mathbf{y}_L^T)^T \in \mathbb{R}^{2N_p}$. Vector $\mathbf{x} = (\mathbf{x}_1^T, \mathbf{x}_2^T)^T \in \mathbb{R}^{2N_p}$ denotes the stacked material density images (unknown), where $\mathbf{x}_l = (x_{l1}, x_{l2}, \dots, x_{ln}, \dots, x_{lN_p})^T \in \mathbb{R}^{N_p}$ represents the l th material for $l = 1, 2$. We model the underlying basis materials using a common-material and a cross-material image model. In the common-material model, patches extracted independently from all materials are assumed sparse in a common union (or collection) of sparsifying transforms. Every extracted patch (from some materi-

S. Ravishankar was a researcher visiting SJTU at the time of submission of this work. This work was supported in part by the SJTU-UM Collaborative Research Program, NSFC (61501292), the Interdisciplinary Program of Shanghai Jiao Tong University (YG2015QN05) and NIH grant U01 E-B018753. *Yong Long (email: yong.long@sjtu.edu.cn).

al image) is assumed to be best sparsified by a particular transform in the collection. The common-material transforms capture features that are common across materials. On the other hand, in the cross-material model, the patches at the same spatial location in different basis material images are stacked together to form larger¹ 3D patches that are assumed to be sparsified using a union of cross-material sparsifying transforms. These transforms help sparsify the materials jointly by exploiting cross-material redundancy.

We formulate the image domain DECT decomposition problem using a pre-learned union of common-material unitary transform matrices $\{\tilde{\Omega}_{k_1}^1\}_{k_1=1}^{K_1}$, and a union of cross-material unitary transform matrices $\{\tilde{\Omega}_{k_2}^2\}_{k_2=1}^{K_2}$, with each $\Omega_{k_r} \in \mathbb{R}^{2m \times 2m}$, as follows:

$$\min_{\mathbf{x} \in \mathbb{R}^{2N_p}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_{\mathbf{W}}^2 + \mathbf{R}(\mathbf{x}), \quad (1)$$

with the regularizer $\mathbf{R}(\mathbf{x})$ defined as

$$\min_{\{\mathbf{z}_j, C_{k_r}^r\}} \sum_{r=1}^2 \sum_{k_r=1}^{K_r} \sum_{j \in C_{k_r}^r} \beta \left\{ \|\tilde{\Omega}_{k_r}^r \mathbf{P}_j \mathbf{x} - \mathbf{z}_j\|_2^2 + \gamma_r^2 \|\mathbf{z}_j\|_0 \right\} \quad (2)$$

where the common-material and cross-material models are denoted using $r = 1, 2$, respectively, and K_r denotes the number of clusters in the r th model. Operator $\mathbf{P}_j \in \mathbb{R}^{2m \times 2N_p}$ extracts the j th patch (overlapping patches assumed) of materials as a vector $\mathbf{P}_j \mathbf{x}$. The patch is constructed by stacking together the vectorized 2D patches extracted from the same spatial location of the multiple material images. Vector $\mathbf{z}_j \in \mathbb{R}^{2m}$ denotes the sparse coefficients for $\mathbf{P}_j \mathbf{x}$, and the ℓ_0 “norm” counts the number of non-zeros in a vector. The set $C_{k_r}^r$ includes the indices of all the patches belonging to the k_r th cluster in the r th model. Each patch $\mathbf{P}_j \mathbf{x}$ belongs to only one cluster in one of the models and the minimization in (2) is over all the patches’ sparse codes and cluster memberships. The parameter β controls the trade-off between noise and image resolution in the decomposition and γ_r for $r = 1, 2$, controls the sparsity in the models.

For $r = 1$, each transform $\tilde{\Omega}_{k_1}^1$ is a block diagonal matrix that sparsifies the individual material images’ patches separately (without mixing them), whereas for $r = 2$, the transforms sparsify the entire 3D patches. Section 3.1 describes the exact structure and learning of these transforms. The transforms are unitary, which simplifies the proposed algorithm for (1).

The matrix $\mathbf{A} \in \mathbb{R}^{2N_p \times 2N_p}$ in (1) is composed of the mass attenuation coefficients and is defined as $\mathbf{A} = \mathbf{A}_0 \otimes \mathbf{I}_{N_p}$, where the operator “ \otimes ” denotes the Kronecker product, and the matrix \mathbf{A}_0 is a 2×2 material decomposition matrix defined as

$$\mathbf{A}_0 = \begin{pmatrix} \varphi_{1H} & \varphi_{2H} \\ \varphi_{1L} & \varphi_{2L} \end{pmatrix}, \quad (3)$$

where φ_{1H} and φ_{1L} denote the mass attenuation coefficients of the l th material at high and low energy, respectively.

The weight matrix $\mathbf{W} \in \mathbb{R}^{2N_p \times 2N_p}$ is designed under the following assumptions: a) the additive noise degrading the reconstructed attenuation image \mathbf{y} follows the Gaussian distribution; b) the high and low energy noise are uncorrelated [11]; and c) the pixels in each attenuation image have the same noise variance [12]. The weight matrix we obtain is $\mathbf{W} = \mathbf{W}_j \otimes \mathbf{I}_{N_p}$, where \mathbf{W}_j represents the weight matrix for the j th pixel, i.e., $\mathbf{W}_j = \text{diag}(\sigma_H^2, \sigma_L^2)^{-1}$, where σ_H^2 and σ_L^2 denote the noise variances for pixels in \mathbf{y}_H and \mathbf{y}_L , respectively.

¹We are focusing on 2D images here so that 3D means 2D with one more channel direction.

We very recently proposed an image domain DECT decomposition technique called DECT-ST [13], where the regularizer $\mathbf{R}(\mathbf{x}) \triangleq \min_{\{\mathbf{z}_l\}} \sum_{l=1}^2 \sum_{j=1}^N \beta_l \left\{ \|\Omega_l \mathbf{P}_{lj} \mathbf{x} - \mathbf{z}_{lj}\|_2^2 + \gamma_l^2 \|\mathbf{z}_{lj}\|_0 \right\}$,

and the transforms Ω_1 and Ω_2 are single (square) transforms learned separately for different material (water and bone) images, and $\mathbf{P}_{lj} \mathbf{x}$ is the j th individual patch extracted from the l th material. Unlike DECT-ST, the proposed model more effectively captures the similarities across materials, while also using a richer union of sparsifying transforms.

3. ALGORITHM

3.1. Learning a Mixed Union of Sparsifying Transforms

We pre-learn a union of common-material unitary transform matrices $\{\tilde{\Omega}_{k_1}^1\}_{k_1=1}^{\sqrt{K_1}}$, and a union of cross-material unitary transform matrices $\{\tilde{\Omega}_{k_2}^2\}_{k_2=1}^{K_2}$, from patches extracted from a set of training material images. For the common-material transforms, we extract patches from all individual basis material images for training and for the cross-material transforms; patches extracted at the same spatial locations from the basis material images of the same object are stacked together to generate the training patches. During the decomposition step (or solving (1)), $\{\tilde{\Omega}_{k_1}^1\}$ are used to form the two blocks (for two materials) of the block diagonal matrices $\{\Omega_{k_1}^1\}$, while $\{\tilde{\Omega}_{k_2}^2\}$ are identical to $\{\Omega_{k_2}^2\}$. We solve the following transform learning problem for each $r = 1, 2$, where \mathbf{I}_{m_r} denotes the $m \times m$ identity matrix for $r = 1$ and the $2m \times 2m$ identity matrix for $r = 2$ [14]:

$$\min_{\{\tilde{\Omega}_{k_r}^r, C_{k_r}^r, \mathbf{z}_{i_r}^r\}} \sum_{k_r=1}^{K_r} \sum_{i_r \in C_{k_r}^r} \left\{ \|\tilde{\Omega}_{k_r}^r \mathbf{Y}_{i_r}^r - \mathbf{z}_{i_r}^r\|_2^2 + \eta^2 \|\mathbf{z}_{i_r}^r\|_0 \right\} \quad (4)$$

s.t. $\tilde{\Omega}_{k_r}^{rT} \tilde{\Omega}_{k_r}^r = \mathbf{I}_{m_r}, \quad 1 \leq k_r \leq K_r,$

where $\eta > 0$ is a scalar parameter controlling sparsity during training, and $\{\mathbf{Y}_{i_1}^1\}$ and $\{\mathbf{Y}_{i_2}^2\}$ denote the vectorized 2D and 3D training patch sets for the two models. The formulation (4) automatically groups the training patches $\{\mathbf{Y}_{i_r}^r\}$ into K_r classes (according to the best matching transforms) for $r = 1, 2$. The transform sparse codes for the training signals are denoted as $\{\mathbf{z}_{i_r}^r\}$ for each r . We use a variant of the algorithm in [10, 15] for training that enforces the unitary constraint [8, 14]. We optimize (4) by alternating between updating $\{C_{k_r}^r, \mathbf{z}_{i_r}^r\}$ (sparse coding and clustering) and $\{\tilde{\Omega}_{k_r}^r\}$ (transform update) with efficient solutions in each step [10, 14, 15]. The cost of the learning algorithm scales in general as $O(m^2 J N')$, where J is the number of iterations of the alternating algorithm and N' is the number of training patches.

3.2. Optimization algorithm

We propose an algorithm for the regularized material separation problem (1) that alternates between updating \mathbf{x} (*image update step*) and updating $\{\mathbf{z}_j, C_{k_r}^r\}$ (*sparse coding and clustering step*).

3.2.1. Image update

With $\{\mathbf{z}_j, C_{k_r}^r\}$ fixed, (1) reduces to the following penalized weighted least squares (PWLS) image decomposition problem:

$$\min_{\mathbf{x} \in \mathbb{R}^{2N_p}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_{\mathbf{W}}^2 + \sum_{r=1}^2 \sum_{k_r=1}^{K_r} \sum_{j \in C_{k_r}^r} \beta \|\Omega_{k_r}^r \mathbf{P}_j \mathbf{x} - \mathbf{z}_j\|_2^2. \quad (5)$$

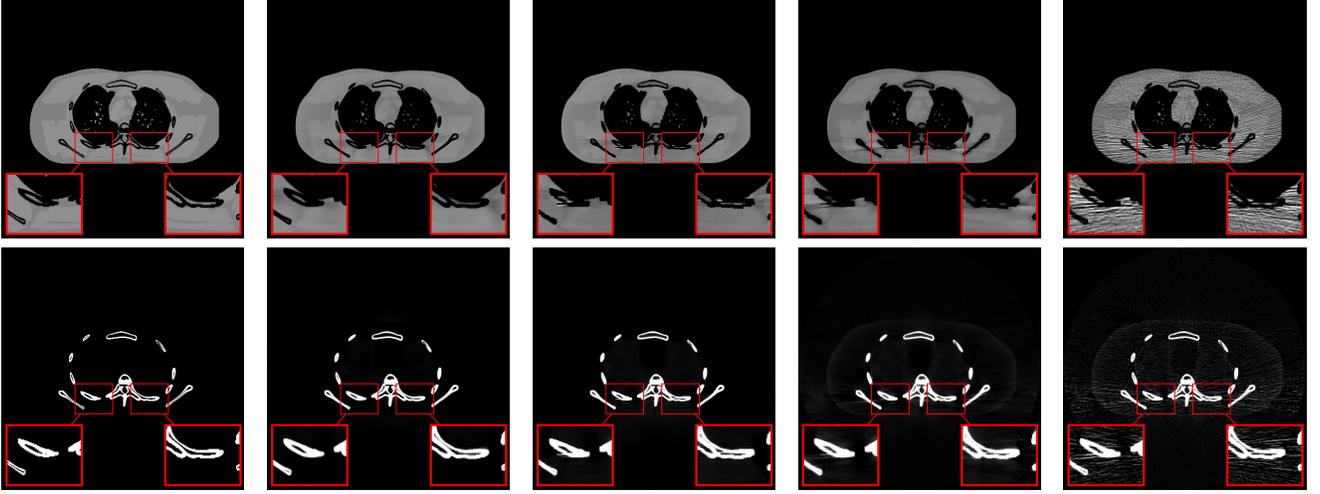


Fig. 1: First to fifth column: material images (from the central slice) of the ground truth, DECT-MULTRA, DECT-ST, DECT-EP, and the direct matrix inversion decomposition, respectively. Water and bone images are in the top and bottom rows with display window [0.6 1.4] g/cm³ and [0 0.8] g/cm³, respectively.

The gradient of $R_2(\mathbf{x})$, the second term in (5), is given as follows:

$$\nabla R_2(\mathbf{x}) = 2\beta \sum_{r=1}^2 \sum_{k_r=1}^{K_r} \sum_{j \in C_{k_r}^r} \mathbf{P}_j^T \boldsymbol{\Omega}_{k_r}^{rT} (\boldsymbol{\Omega}_{k_r}^r \mathbf{P}_j \mathbf{x} - \mathbf{z}_j). \quad (6)$$

Using periodically positioned overlapping image patches with a patch stride of 1 pixel that wrap around at material image (2D) boundaries, the term $\sum_{r=1}^2 \sum_{k_r=1}^{K_r} \sum_{j \in C_{k_r}^r} \mathbf{P}_j^T \mathbf{P}_j$ is a diagonal matrix equal to $m \mathbf{I}_{2N_p}$. Then, because each $\boldsymbol{\Omega}_{k_r}^r$ is a unitary matrix, the gradient in (6) is further simplified as follows:

$$\nabla R_2(\mathbf{x}) = 2\beta \left(m\mathbf{x} - \sum_{r=1}^2 \sum_{k_r=1}^{K_r} \sum_{j \in C_{k_r}^r} \mathbf{P}_j^T \boldsymbol{\Omega}_{k_r}^{rT} \mathbf{z}_j \right). \quad (7)$$

Since $2\beta m \mathbf{I}_2$ and $\mathbf{A}_0^T \mathbf{W}_j \mathbf{A}_0$ are independent of \mathbf{x} , \mathbf{z}_j and $C_{k_r}^r$, we precompute the 2×2 blocks of the Hessian matrix $\mathbf{B}_j = \mathbf{A}_0^T \mathbf{W}_j \mathbf{A}_0 + 2\beta m \mathbf{I}_2$ for all j . Then clearly the image update problem is solved in (5) in closed-form at each pixel j as follows:

$$\mathbf{x}_j = \mathbf{B}_j^{-1} (\mathbf{A}_0^T \mathbf{W}_j \mathbf{y}_j + 2\beta \mathbf{M}_j \sum_{r=1}^2 \sum_{k_r=1}^{K_r} \sum_{j \in C_{k_r}^r} \mathbf{P}_j^T \boldsymbol{\Omega}_{k_r}^{rT} \mathbf{z}_j), \quad (8)$$

where \mathbf{M}_j is a matrix that extracts vector entries corresponding to the j th pixel. In (8), we use the block diagonal structure of \mathbf{A} and (7) to separate the \mathbf{x} -update into N_p pixel-wise updates to avoid explicitly computing the large \mathbf{A} matrix. Obviously (8) provides the global minimizer in the image update step.

3.2.2. Sparse Coding and Clustering Step

With \mathbf{x} fixed, we solve the following sub-problem to obtain the sparse codes and cluster assignments for each patch:

$$\min_{\{\mathbf{z}_j, C_{k_r}^r\}} \sum_{r=1}^2 \sum_{k_r=1}^{K_r} \sum_{j \in C_{k_r}^r} \|\boldsymbol{\Omega}_{k_r}^r \mathbf{P}_j \mathbf{x} - \mathbf{z}_j\|_2^2 + \gamma_r^2 \|\mathbf{z}_j\|_0. \quad (9)$$

For any patch $\mathbf{P}_j \mathbf{x}$, the solution with respect to the sparse code above is obtained by hard-thresholding as $\mathbf{z}_j = H_{\gamma_r}(\boldsymbol{\Omega}_{k_r}^r \mathbf{P}_j \mathbf{x})$, where the operator $H_\gamma(b)$ returns 0 if $|b| < \gamma$ and otherwise returns b . Replacing the variables \mathbf{z}_j in (9) with these optimal values, the problem reduces to a clustering problem, where the optimal cluster assignment for each patch is as follows:

$$(\hat{r}_j, \hat{k}_j) = \arg \min_{\substack{1 \leq k_r \leq K_r \\ 1 \leq r \leq 2}} \left\{ \|\boldsymbol{\Omega}_{k_r}^r \mathbf{P}_j \mathbf{x} - H_{\gamma_r}(\boldsymbol{\Omega}_{k_r}^r \mathbf{P}_j \mathbf{x})\|_2^2 + \gamma_r^2 \|H_{\gamma_r}(\boldsymbol{\Omega}_{k_r}^r \mathbf{P}_j \mathbf{x})\|_0 \right\} \quad (10)$$

Minimizing over k_r and r above finds the best matched transform and model for each patch, i.e., we compute the cost in (10) with respect to each transform in the two models to determine the best match. In particular, for $r = 1$, each of the two blocks (along the diagonal) of $\boldsymbol{\Omega}_{k_1}^1$ can take one of $\sqrt{K_1}$ matrix values (chosen from $\tilde{\boldsymbol{\Omega}}_{k_1}^1$). Since the cost of the 3D patch in (10) for $r = 1$ decomposes into the sum of the clustering costs for the individual material (2D) patches, we find the best matching transform $\boldsymbol{\Omega}_{k_1}^1$ for the two material patches independently and then add these costs, which would efficiently give the smallest cost value in model 1. Comparing this to the smallest cost in model 2 yields the best matched model (and corresponding transform). Then, the optimal sparse codes are updated using hard-thresholding as $\hat{\mathbf{z}}_j = H_{\gamma_{\hat{r}_j}}(\boldsymbol{\Omega}_{\hat{k}_j}^{\hat{r}_j} \mathbf{P}_j \mathbf{x}) \forall j$.

Since the proposed algorithm is an exact alternating minimization method, the objective in (1) is monotone decreasing and converges over the algorithm iterations.

4. EXPERIMENTAL RESULTS

We evaluate the proposed DECT-MULTRA and compare its performance to the direct matrix inversion method (solving (1) without regularization), DECT-EP, and the recent DECT-ST that learns different square transforms for different materials. For DECT-MULTRA, we pre-learned the common-material union of transforms ($\sqrt{K_1} = 15$) from 8×8 patches extracted from five slices of water images and five slices of bone images of the XCAT phantom [16]. The cross-material union of transforms ($K_2 = 10$) was learned from patches

extracted from five slices of cross-material images, with patch size $8 \times 8 \times 2$. For the common-material and cross-material transforms, the training parameter η was set as 0.21 and 0.17, respectively. We ran 2000 iterations of the alternating minimization algorithm for (4) to ensure convergence. For DECT-ST, similar parameters were used during training except that the λ_0 and η values for the water and bone (square) transforms were empirically set as $\{31, 0.12\}$ and $\{31, 0.15\}$, respectively.

For the test data, we first generated three 1024×1024 reference images using the central (77th) slice and the 61st and 150th slices of the XCAT phantom, with a pixel size of $0.49 \times 0.49 \text{ mm}^2$. We generated noisy (Poisson noise) sinograms of size 888 (radial samples) \times 984 (angular views) using GE LightSpeed X-ray CT fan-beam system geometry corresponding to a poly-energetic source at 80kVp and 140kVp with 1.86×10^5 and 1×10^6 incident photons per ray, respectively. We used filtered back projection (FBP) to reconstruct the 2D 512×512 high and low attenuation images, where the pixel size was $0.98 \times 0.98 \text{ mm}^2$.

We obtained the basis material images from attenuation images via the direct matrix inversion method, and used these to initialize the DECT-EP method. Basis material images obtained with DECT-EP were used as initialization for DECT-ST and the proposed DECT-MULTRA to accelerate their convergence. We empirically chose the regularization parameters for all methods in each experiment to achieve the best image quality and decomposition accuracy.

Fig. 1 shows the decompositions by different methods. The proposed DECT-MULTRA suppresses the high noise observed in the direct matrix inversion method. In addition, compared to DECT-EP and DECT-ST, the proposed DECT-MULTRA significantly improves image quality by reducing artifacts and improving the edge details.

Fig. 2 presents an example of the pixel-level clustering in the DECT-MULTRA ($K_1 = 225, K_2 = 10$) decomposition. Two cross-material transforms and the bone and water pixels in the respective clusters are shown. For each image pixel, we perform a majority vote among the 3D patches overlapping the pixel to determine the model (r) to which it belongs and the corresponding cluster. Fig. 2 only shows the pixels (using their estimated densities) from $r = 2$ for specific classes. The clustering depicts the shared edges and other features between materials that are sparsified jointly by the pre-learned cross-material transforms.

To compare the performance of these methods quantitatively, we computed the Root Mean Square Error (RMSE) for the decomposed material images. For a decomposed image \hat{x}_l , RMSE is defined as $\sqrt{\sum_{j=1}^{N_{p,PS}} (\hat{x}_{lj} - x_{lj}^*)^2 / N_{p,PS}}$, where x_{lj}^* denotes the down-sampled true density of the l th material at the j th pixel location and $N_{p,PS}$ is the number of pixels in the phantom support (a circle removes all the black background area outside the image that is not interesting). Table 1 lists the RMSE for the four compared methods. DECT-MULTRA provides significant improvements in RMSE for the basis materials compared to DECT-ST, DECT-EP, and the direct matrix inversion method.

5. CONCLUSIONS

We proposed a new method for image domain DECT decomposition that combines conventional PMLS estimation with regularization based on a mixed union of learned unitary transforms model that exploits both the common properties among material images and their cross-dependencies. The proposed DECT-MULTRA algorithm provided superior material image quality and decomposition accuracy compared to the recent DECT-ST approach and the non-adaptive DECT-EP method. DECT-MULTRA successfully reduces

Method		Direct	DECT	DECT	DECT
		Inversion	EP	ST	MULTRA
Slice	Water	72.8	60.9	51.3	42.8
	Bone	68.4	60.2	51.6	43.9
Slice	Water	92.4	65.9	55.6	38.7
	Bone	89.0	72.2	61.8	49.8
Slice	Water	116.7	69.1	61.7	38.6
	Bone	110.8	76.7	67.0	50.8

Table 1: RMSE of decomposed images of basis materials by Direct Matrix Inversion, DECT-EP, DECT-ST and DECT-MULTRA ($K_1 = 225, K_2 = 10$) for three different slices of the XCAT phantom. The unit for RMSE is 10^{-3} g/cm^3 .

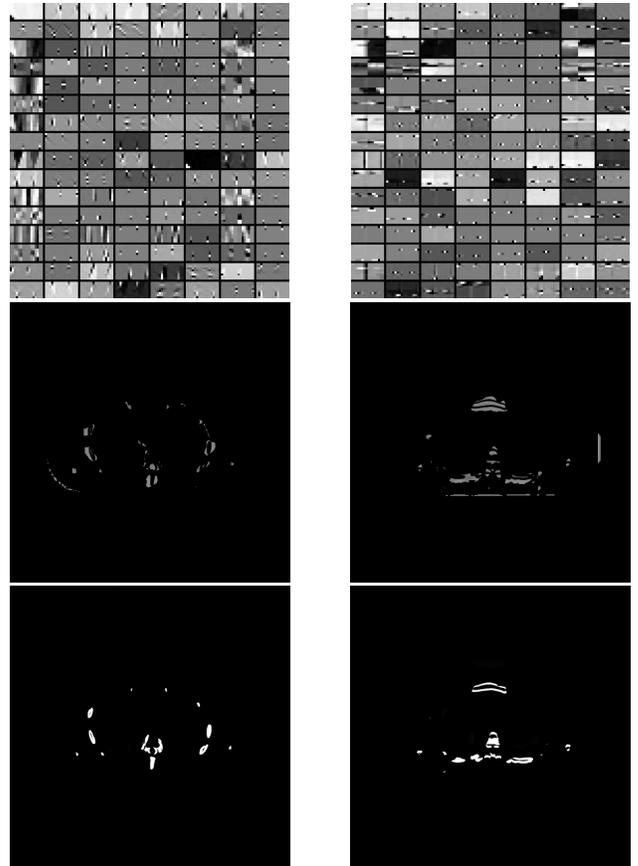


Fig. 2: The cross material transforms (atoms shown as 8×16 patches) for classes 4 (left) and 8 (right) are shown in the top row. The middle and bottom rows show the water and bone pixels (using estimated intensities) belonging to classes 4 (left) and 8 (right) with display windows $[0.6 \ 1.4] \text{ g/cm}^3$ and $[0 \ 0.8] \text{ g/cm}^3$, respectively.

the artifacts at the boundaries of different materials and also provides improved sharpness of edges in the soft tissue. In future work, we plan to investigate experiments with more general multi-material (with several materials) decompositions with DECT-MULTRA, and plan to apply the proposed technique for clinical or industrial CT data.

6. REFERENCES

- [1] Y. Long and J. A. Fessler. Multi-material decomposition using statistical image reconstruction for spectral CT. *IEEE Trans. Med. Imag.*, 33(8):1614–26, August 2014.
- [2] J. Noh, J. A. Fessler, and P. E. Kinahan. Statistical sinogram restoration in dual-energy CT for PET attenuation correction. *IEEE Trans. Med. Imag.*, 28(11):1688–702, November 2009.
- [3] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang. Low-dose X-ray CT reconstruction via dictionary learning. *IEEE Trans. Med. Imag.*, 31(9):1682–97, September 2012.
- [4] B. Zhao, H. Ding, Y. Lu, G. Wang, J. Zhao, and S. Molloy. Dual-dictionary learning-based iterative image reconstruction for spectral computed tomography application. *Phys. Med. Biol.*, 57(24):8217–30, December 2012.
- [5] L. Li, Z. Chen, and P. Jiao. Dual-energy CT reconstruction based on dictionary learning and total variation constraint. In *Proc. IEEE Intl. Symp. Biomed. Imag.*, pages 2358–61, 2012.
- [6] Y. Zhang, X. Mou, H. Yu, G. Wang, and Q. Xu. Tensor based dictionary learning for spectral CT reconstruction. *IEEE Trans. Med. Imag.*, 36(1):142–54, 2017.
- [7] W. Wu, Y. Zhang, Q. Wang, F. Liu, P. Chen, and H. Yu. Low-dose spectral CT reconstruction using image gradient ℓ_0 -norm and tensor dictionary. *Applied Mathematical Modelling*, 63:538 – 57, 2018.
- [8] S. Ravishankar and Y. Bresler. ℓ_0 sparsifying transform learning with efficient optimal updates and convergence guarantees. *IEEE Trans. Sig. Proc.*, 63(9):2389–404, May 2015.
- [9] X. Zheng, Z. Lu, S. Ravishankar, Y. Long, and J. A. Fessler. Low dose CT image reconstruction with learned sparsifying transform. In *Proc. IEEE Wkshp. on Image, Video, Multidim. Signal Proc.*, pages 1–5, July 2016.
- [10] X. Zheng, S. Ravishankar, Y. Long, and J. A. Fessler. PWLS-ULTRA: An efficient clustering and learning-based approach for low-dose 3D CT image reconstruction. *IEEE Trans. Med. Imag.*, 37(6):1498–510, June 2018.
- [11] R. Zhang, J. B. Thibault, C. A. Bouman, K. D. Sauer, and J. Hsieh. Model-based iterative reconstruction for dual-energy X-ray CT using a joint quadratic likelihood model. *IEEE Trans. Med. Imag.*, 33(1):117–34, January 2014.
- [12] T. Niu, X. Dong, M. Petrongolo, and L. Zhu. Iterative image-domain decomposition for dual-energy CT. *Med. Phys.*, 41(4):041901, April 2014.
- [13] Z. Li, S. Ravishankar, Y. Long, and J. A. Fessler. Image-domain material decomposition using data-driven sparsity models for dual-energy CT. In *Proc. IEEE Intl. Symp. Biomed. Imag.*, pages 52–6, April 2018.
- [14] S. Ravishankar and Y. Bresler. Data-driven learning of a union of sparsifying transforms model for blind compressed sensing. *IEEE Trans. Computational Imaging*, 2(3):294–309, September 2016.
- [15] B. Wen, S. Ravishankar, and Y. Bresler. Structured overcomplete sparsifying transform learning with convergence guarantees and applications. *Intl. J. Comp. Vision*, 114(2-3):137–66, September 2015.
- [16] W. P. Segars, M. Mahesh, T. J. Beck, E. C. Frey, and B. M. W. Tsui. Realistic CT simulation using the 4D XCAT phantom. *Med. Phys.*, 35(8):3800–8, August 2008.