Super resolution dual-energy cone-beam CT imaging with dual-layer flat-panel detector

Ting Su,¹ Jiongtao Zhu,² Xin Zhang,¹ Dong Zeng,³ Yuhang Tan,¹ Han Cui,¹ Hairong Zheng,⁴ Jianhua Ma,³ Dong Liang,^{1,4} and Yongshuai Ge^{1,4,a)}

¹⁾Research Center for Medical Artificial Intelligence, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518055, China.

²⁾College of Physics and Optoelectronic Engineering, Ministry of Education and Guangdong Province, Key Laboratory of Optoelectronic Equipment and Systems, Shenzhen University, Shenzhen 518060, China.

³⁾School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China.

⁴⁾Paul C Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518055, China.

(Dated: 18 October 2022)

For medical cone-beam computed tomography (CBCT) imaging, the native receptor array of the flat-panel detector (FPD) is usually binned into a reduced matrix size. By doing so, the signal readout speed can be increased by over 4-9 times at the expense of sacrificing the spatial resolution by at least 50%-67%. Clearly, such tradition poses a main bottleneck in generating high spatial resolution and high temporal resolution CBCT images at the same time. In addition, the conventional FPD is also difficult in generating dual-energy CBCT images. In this paper, we propose an innovative super resolution dual-energy CBCT imaging method, named as suRi, based on dual-layer FPD (DL-FPD) to overcome these aforementioned difficulties at once. With suRi, specifically, an 1D or 2D sub-pixel (half pixel in this study) shifted binning is applied to replace the conventionally aligned binning to double the spatial sampling rate during the dual-energy data acquisition. As a result, the suRi approach provides a new strategy to enable high signal readout speed and high spatial resolution CBCT imaging with FPD. Moreover, a penalized likelihood material decomposition algorithm is developed to directly reconstruct the high resolution bases from the dual-energy CBCT projections containing spatial sub-pixel shifts. Experiments based on the single-layer FPD and DL-FPD are performed with physical phantoms and biological specimen to validate this newly developed suRi method. The synthesized monochromatic CT imaging results demonstrate that suRi can significantly improve the spatial image resolution by 46.15%. We believe the developed suRi method would be capable to greatly enhance the imaging performance of the DL-FPD based dual-energy CBCT systems in future.

Keywords: System modeling, dual-energy imaging, sub-pixel shift, CT image reconstruction.

^{a)}Electronic mail: ys.ge@siat.ac.cn.

I. INTRODUCTION

Over the past two decades, X-ray flat-panel detector (FPD) made from a single layer of CsI:TI scintillator and amorphous silicon (α -Si) based thin film transistor (TFT) backplate has been widely used in medical imaging applications. Due to its small native pixel dimension, e.g., 0.07-0.2 mm, FPD shows superior performance in detecting the ultra-fine details of the anatomical structures such as the micro-calcification, bone marrow and contrast enhanced blood vessels in two dimensional digital radiography (DR) imaging. For three and four dimensional cone-beam computed tomography (CBCT) imaging, flat-panel detector (FPD) is also irreplaceable in applications such as oral imaging¹, image-guided radiation $therapy^2$ and interventional therapy³. Despite of these advancements, however, the dilemma between the intrinsic pixel-size-defined high spatial resolution and the pixel-number-confined slow data acquisition speed in FPD impedes its further developments in advanced CBCT imaging tasks which require both high spatial and high temporal resolution simultaneously. Take the PaxScan 4030CB FPD (Varex, USA) as an example, the native pixel dimension is 0.194 mm, and the full receptor array consists of 2048×1536 detector elements. With the 1×1 binning mode (limiting resolution 2.58 lp/mm), this FPD can only acquire 7.5 frames per second (fps) at most^{4,5}, and at least 40 seconds are needed to complete a CBCT scan. The excessive CBCT scanning time would cause the motion artifacts and the loss of temporal variations of the contrast agent. Therefore, the FPD is usually worked at 2×2 binning mode (0.388 mm effective pixel dimension, corresponding to a reduced limiting spatial resolution of 1.29 lp/mm). By doing so, a quick signal readout speed of 30.0 fps can be achieved to complete the CBCT data acquisition within 6-10 seconds. Compared to the 1×1 binning mode, the 2×2 binning mode saves about 80% CBCT scan time, but at the expense of losing 50% image spatial resolution⁴. Besides, the current FPD is also difficult in performing dual-energy CBCT imaging to generate quantitative material-specific images for accurate disease diagnoses⁶.

As a promising dual-energy CBCT imaging approach, the flat-panel detector made from dual layers of CsI:TI scintillator and α -Si TFT back-plate (DL-FPD) can acquire the dualenergy CBCT data simultaneously: the low-energy (LE) X-ray projection is acquired from the top layer, and the high-energy (HE) X-ray projection is acquired from the bottom layer. The two CsI:TI scintillator layers may have different thicknesses, and additional beam filtration made of 1.0 mm Copper may also be inserted between the two layers to further separate the beam spectra. By far, several investigations based on the DL-FPD have been reported. For example, Shi *et al.* and Ståhl *et al.* demonstrated the feasibility of performing accurate dual-energy CBCT imaging based on the DL-FPD^{7,8}. Wang *et al.* proposed a model-based high resolution material decomposition method for the DL-FPD CBCT⁹. Even though the DL-FPD helps realizing the dual-energy CBCT imaging, however, the conflict between the high spatial resolution imaging and the high speed imaging in single-layer FPD (SL-FPD) is still inherited. In other words, the DL-FPD encounters the same trade-off between the spatial resolution and the temporal resolution as the SL-FPD.

To overcome this long-standing difficulty, we propose an innovative imaging method, named as suRi, for DL-FPD to realize high temporal and spatial resolution CBCT imaging at the same time. In particular, the conventionally aligned binning approach is altered into a sub-pixel (half pixel in this study) shifted binning approach, see the illustrations in Fig. 1 for more details. Depending on the imaging task, such half pixel shift could be one-dimensional along the horizontal direction, or two-dimensional along the diagonal direction. Take the simple one-dimensional half pixel shift as an example, the information recorded on the top and bottom layers are spatially shifted by half-pixel. As shown in Fig. 1, the line integral of the polychromatic X-ray beam that passes through a certain object position is sampled by two different FPD pixel elements: the top one is shifted by half pixel with regard to the bottom one. Consequently, the sub-pixel shift used by suRi doubles the spatial sampling rate assuming parallel incident beam. Therefore, the proposed suRi approach not only enables fast dual-energy CBCT imaging, but also enables super spatial resolution CBCT imaging. With such redefined DL-FPD data acquisition scheme, we believe the dual-energy CBCT imaging with high spatial resolution and high temporal resolution performance can be achieved. To the best of our knowledge, no such investigations have been reported before.

The major contributions of this work are as follows: (1) The sub-pixel shift dual-energy CBCT data acquisition scheme is proposed for DL-FPD for the first time. This method perfectly fits with the stacked structure of the DL-FPD to increase the spatial sampling rate while shortening the signal readout time by introducing additional "geometric mismatch" between the two detector layers. (2) A high performance one-step material decomposition algorithm is developed for the proposed novel DL-FPD based sub-pixel shift dual-energy CBCT imaging. The transmitted X-ray intensity of the sub-ray in each sub-pixel together

with the CBCT imaging geometry, the data noise statistics and the beam spectra are utilized to build a more accurate forward imaging model, and the penalized likelihood function is optimized to generate the basis images with high spatial resolution.

The rest of this paper is organized below: Section II introduces some related works. Section III presents the mathematical model of the sub-pixel shift data acquisition method for DL-FPD, the penalized likelihood material decomposition method, the experimental setup, the implementation details and evaluation metrics. Section IV presents the experimental results of different objects. Section V provides the discussions and a brief conclusion.

II. RELATED WORK

A. Super resolution CT imaging method

There are two main strategies to realize super spatial resolution imaging in medical CBCT applications. First, the flying focal spot (FFS) technique^{10,11} used by some advanced X-ray tube doubles the spatial sampling density in the horizontal and vertical directions, and thus can significantly reduce the in-plane and the out-of-plane aliasing artifacts. Since the detector matrix size is fixed, the FFS technique usually augments the total data size and adds burdens to the detector read-out speed. Second, the sub-pixel shifting technique in the detector end, which has been proposed for optical imaging for a long time¹², and can also be utilized to greatly improve the spatial resolution of CT images. For example, Yan et al. proposed to shift the single-layer detector array and make multiple scans to jointly reconstruct the high spatial resolution CT image¹³. Li *et al.* suggested to move the rotation stage in a fixed trajectory during the data acquisition to obtain multiple projection images with sub-pixel displacements¹⁴. Szczykutowicz *et al.* developed a spectral-spatial encoding method that acquires multiple measurements at different beam energies and different spatial positions to decompose the line integrals into two unique bases¹⁵. However, one major limitation of these studies is the prolonged data acquisition period due to the repeated measurements. In this work, the proposed suRi method can efficiently overcome such limitation by incorporating the sub-pixel shifts into the DL-FPD.

B. Material decomposition algorithm

The dual-energy CT material decomposition methods can be mainly divided into three categories. The first one is known as the projection-domain material decomposition method^{16,17}, which discriminates the material-specific density integrals from paired dualenergy projections. This kind of method is difficult to be implemented straightforwardly in this study due to the spatial mismatch between the dual-energy projections. The second one is the image-domain decomposition method^{18–20}, in which the CT images reconstructed using the filtered back-projection (FBP) or iterative algorithm are decomposed into different bases. Despite that the spatial mismatch of projections can be compensated after reconstruction, however, the image-domain method is not able to generate CT bases with the desired high spatial resolution. The third approach is the so called direct (one-step) material decomposition method $^{21-23}$, which generates the material-specific CT maps directly from the dual-energy projections by using a certain imaging model. Since the imaging geometry (e.g., the sub-pixel shifts), the noise fluctuation and the beam spectra can be incorporated into the formulated imaging model, therefore, the direct material decomposition method is selected in this study to reconstruct high spatial resolution CT bases directly from the acquired sub-pixel shifted dual-energy projections.

III. METHOD

A. Imaging model in suRi

To better explain the proposed suRi data acquisition method, we need to define several coordinate systems on the detector planes, see Fig. 2. First, the coordinates of the unbinned detector elements are denoted by (u, v), and they are assumed to be spatially fixed for all of the K ($K \ge 2$) detector layers, see Fig. 2(a). Second, the coordinates of the $B_u \times B_v$ binned detector elements are denoted by (U, V), see Fig. 2(b). As a consequence, the beam intensity recorded by the binned detector element (U, V) in the kth detector layer, denoted as $\hat{y}_{U,V}^k$, can be expressed as a summation of the responses $\hat{y}_{u,v}^{\prime k}$ that are acquired from the unbinned detector elements, namely,

$$\hat{y}_{U,V}^{k} = \sum_{u,v \in \mathcal{N}_{U,V}^{k}} \hat{y}_{u,v}^{'k}, \qquad (3.1)$$



FIG. 1. Illustration of the proposed suRi method in this study. (a) The overall imaging setup of the DL-FPD based DECT system. In DL-FPD, the top layer and bottom layer record the lowenergy and high-energy X-ray photons, respectively. (b) The conventionally used data acquisition scheme, denoted as aligned binning, in which the signal of four neighboring pixels at the same spatial locations of the top and bottom detector layers are summed up. (c) The proposed suRi data acquisition scheme, in which the binned pixels in the top layer and bottom layer use are shifted diagonally by half pixel size. As clearly seen from the difference image, inconsistent spatial information are recorded in the top layer and bottom layer, separately. Due to this reason, the proposed suRi method can be used to restore the high spatial resolution information and thus achieve high spatial resolution DECT imaging. Note that 2×2 binning is assumed in (b) and (c) to demonstrate the aligned binning and sub-pixel shift binning modes.

where $\mathcal{N}_{U,V}^k$ represents the neighboring $B_u \times B_v$ detector elements that belong to the binned detector element (U, V) on the kth detector layer.

To determine $\mathcal{N}_{U,V}^k$, we assume that the binned detector element (U, V) is spatially displaced between two adjacent detector layers:

$$u_i^{(k+1)}(U,V) - u_u^{(k)}(U,V) = \Delta_u^k \times \delta_{del},$$
(3.2)

$$u_v^{(k+1)}(U,V) - u_v^{(k)}(U,V) = \Delta_v^k \times \delta_{del},$$
(3.3)



FIG. 2. The discrete coordinates for (a) the high resolution unbinned projection data $\hat{y}_{u,v}^{'k}$, and (b) the low resolution binned projection data $\hat{y}_{U,V}^{k}$. The binning mode $B_u \times B_v$ is applied to generate $\hat{y}_{U,V}^{k}$ from $\hat{y}_{u,v}^{'k}$. Quantitatively, $N_u = B_u \times N_U$, $N_v = B_v \times N_V$.

where $u_u^{(k)}(U, V)$ and $u_v^{(k)}(U, V)$ are the vertical and horizontal physical locations of the binned detector element (U, V) in the *k*th layer, δ_{del} denotes the physical dimension of one native detector element, Δ_u^k and Δ_v^k denote the relative displacements along the vertical and horizontal directions, respectively. Without loss of generality, herein we assume $\Delta_u^k \ge 0$ and $\Delta_v^k \ge 0$. It is easy to derive that the coordinates of the neighboring pixels in $\mathcal{N}_{U,V}^k$ satisfies:

$$(U-1) \times B_u + 1 + \sum_{\kappa=1}^k \Delta_u^{\kappa} \le u \le U \times B_u + \sum_{\kappa=1}^k \Delta_u^{\kappa}$$
(3.4)

$$(V-1) \times B_v + 1 + \sum_{\kappa=1}^k \Delta_v^{\kappa} \le v \le V \times B_v + \sum_{\kappa=1}^k \Delta_v^{\kappa}$$
(3.5)

The above two relationships define the spatial correspondence between the native pixel (u, v)and the corresponding binned element (U, V). In particular, if $\Delta_u^k = \Delta_v^k = 0$, the equality $\mathcal{N}_{U,V}^1 = \dots = \mathcal{N}_{U,V}^K$ holds for the K number of detector layers. Under such circumstance, the imaging model provided in (3.1) corresponds to the aligned binning data acquisition, as schemed in Fig. 1(b). However, if the displacements satisfies $0 < \Delta_u^{k\geq 2} < B_u$ or $0 < \Delta_v^{k\geq 2} < B_v$ (herein, we assume $\Delta_u^{k=1} = \Delta_v^{k=1} = 0$, meaning the element binning on the first detector layer is not displaced.), then $\mathcal{N}_{U,V}^k$ varies from $\mathcal{N}_{U,V}^{k+1}$, indicating that some detector elements are overlapped, as schemed in Fig. 1(c). Mathematically, the indices of $u, v \in \mathcal{N}_{u,v}^k \cap \mathcal{N}_{U,V}^{k+1}$, are found to be:

$$(U-1) \times B_u + 1 + \sum_{\kappa=1}^{k+1} \Delta_u^{\kappa} \le u \le U \times B_u + \sum_{\kappa=1}^k \Delta_u^{\kappa}$$
(3.6)

$$(V-1) \times B_v + 1 + \sum_{\kappa=1}^{k+1} \Delta_v^{\kappa} \le v \le V \times B_v + \sum_{\kappa=1}^k \Delta_v^{\kappa}$$
(3.7)

The transition from the unbinned projection $\hat{\mathbf{Y}'}^k \in \mathbb{R}^{N_u N_v}$ of the *k*th detector layer,

$$\hat{\boldsymbol{Y}'}^{k} = [\hat{y}_{1,1}^{'k}, \hat{y}_{2,1}^{'k}, ..., \hat{y}_{1,2}^{'k}, \hat{y}_{2,2}^{'k}, ..., \hat{y}_{u,v}^{'k}, ... \hat{y}_{N_u-1,N_v}^{'k}, \hat{y}_{N_u,N_v}^{'k}]^T$$
(3.8)

to the binned projection $\hat{\boldsymbol{Y}}^k \in \mathbb{R}^{N_U N_V}$

$$\hat{\boldsymbol{Y}}^{k} = [\hat{y}_{1,1}^{k}, \hat{y}_{2,1}^{k}, ..., \hat{y}_{1,2}^{k}, \hat{y}_{2,2}^{k}, ..., \hat{y}_{U,V}^{k}, ... \hat{y}_{N_{U}-1,N_{V}}^{k}, \hat{y}_{N_{U},N_{V}}^{k}]^{T}$$
(3.9)

can be derived based upon (3.1):

$$\hat{\boldsymbol{Y}}^{k} = \boldsymbol{D}^{k} \hat{\boldsymbol{Y}}^{'k}, \qquad (3.10)$$

where $\mathbf{D}^k \in \mathbb{R}^{N_U N_V \times N_u N_v}$ denotes the displaced binning matrix and is composed by 0 and 1.

A simple example is discussed below to demonstrate that the unbinned projection $\hat{\mathbf{Y}'}^k \in \mathbb{R}^{N_u N_v}$ may be recovered from the binned projection $\hat{\mathbf{Y}}^k \in \mathbb{R}^{N_U N_V}$. We assume a linear dual-layer (K=2) detector array with 8 native elements $(N_u = 1, N_v = 8)$ in each layer, and the 1 × 2 binning $(B_u = 1, B_v = 2)$ is applied. Thus, the binned detector matrix has a size of 1 × 4 $(N_U = 1, N_V = 4)$ for each detector layer. To avoid the illness of the matrix inversion in (3.10), the displacement of binned detector element is set to $\Delta_u^k = 0$ and $\Delta_v^k = 1$. Hence, the displacement matrix \mathbf{D}^k in each detector layer is written as:

$$\boldsymbol{D}^{1} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix},$$
(3.11)
$$\boldsymbol{D}^{2} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$
(3.12)

Ignoring the difference of X-ray beam spectra recorded by the two detector layers, i.e., $\hat{\boldsymbol{Y}}^{'1} = \hat{\boldsymbol{Y}}^{'2} = \hat{\boldsymbol{Y}}^{'}$ for the monochromatic X-ray beam, consequently, we obtain:

$$\begin{bmatrix} \hat{\boldsymbol{Y}}^{1} \\ \hat{\boldsymbol{Y}}^{2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}^{1} \\ \boldsymbol{D}^{2} \end{bmatrix} \times \hat{\boldsymbol{Y}}^{'}.$$
(3.13)

With a standard matrix inversion, the unbinned $\hat{\boldsymbol{Y}}'$ corresponding to the high resolution projection can be readily obtained from the two binned projections $\hat{\boldsymbol{Y}} = [\hat{\boldsymbol{Y}}^1, \hat{\boldsymbol{Y}}^2]^T$:

$$\hat{\boldsymbol{Y}}' = \boldsymbol{D}^{-1} \hat{\boldsymbol{Y}} \tag{3.14}$$

With the sub-pixel shift technique, the spatial resolution of the projection image obtained from the DL-FPD or ML-FPD can be potentially improved by a factor of K (along the uor v direction). Table I provides more examples using the dual-layer and tripe-layer FPD. Additionally, the potential acceleration rate of signal readout speed is also estimated. Be aware that only the 1D horizontal sub-pixel shifts are considered.

TABLE I. Detector settings in suRi with 1D horizontal sub-pixel shift. The physical shift and resolution are listed along the u and v directions. The potential acceleration rate of signal readout speed is compared with the 1×1 binning mode.

Detector	Binning	Physical shift	Resolution	Acceleration
Dual-layer	1×2	$[0, \delta_{del}]$	$[\delta_{del},\delta_{del}]$	2
	1×4	$[0, 2\delta_{del}]$	$[\delta_{del},2\delta_{del}]$	4
	1×6	$[0, 3\delta_{del}]$	$[\delta_{del}, 3\delta_{del}]$	6
Triple-layer	1×3	$[0, \delta_{del}]$	$[\delta_{del},\delta_{del}]$	3
	1×6	$[0, 2\delta_{del}]$	$[\delta_{del}, 2\delta_{del}]$	6
	1×9	$[0, 3\delta_{del}]$	$[\delta_{del},3\delta_{del}]$	9

In reality, it is difficult to solve Eq. (3.10) due to the different beam spectra recorded by the individual detector layer in DL-FPD. Nevertheless, it is still possible to improve the CT image spatial resolution via suRi since the needed high spatial resolution information have already been encoded in the acquired projections. To reconstruct such CT images with high spatial resolution, the iterative CT reconstruction algorithm has to be employed.

B. Penalized likelihood material decomposition

In CT imaging, the expected intensity of the *i*th ray registered by a binned detector pixel is treated as a sum of those obtained from several discrete sub-rays r (corresponding to different sub-pixels). The linear attenuation coefficient of a compound material can be regarded as the weighted combination of different basis materials. Thus (3.1) can be rewritten as:

$$\hat{y}_{i}^{k}(\boldsymbol{\rho}) = \sum_{r \in \mathcal{N}_{i}^{k}} \sum_{e=1}^{E} S_{ie}^{kr} \exp\left(-\sum_{m=1}^{M} f_{e}^{m} \sum_{j=1}^{J} a_{ij}^{kr} \rho_{j}^{m}\right).$$
(3.15)

where S_{ie}^{kr} denotes the spectral intensity at energy e for the rth sub-ray of ray i, f_e^m denotes the mass attenuation coefficient of material m at energy e, a_{ij}^{kr} denotes the intersection length of the rth sub-ray of ray i with the jth voxel, ρ_j^m denotes the density value of material mat voxel j, $\boldsymbol{\rho} = (\rho_1^1, ..., \rho_J^1; ...; \rho_1^m, ..., \rho_J^m; ...; \rho_1^M, ..., \rho_J^M)^T$ is a matrix of the density values. Assuming that the projection measurements follow Poisson distributions, we can obtain the negative log-likelihood function to solve for the unknown material basis maps:

$$-L(\boldsymbol{\rho}) = -\log\left(\mathcal{P}(y|\hat{y})\right) = \sum_{k=1}^{K} \sum_{i=1}^{I} \hat{y}_{i}^{k} - y_{i}^{k} \ln \hat{y}_{i}^{k}.$$
(3.16)

where y_i^k denotes the measurement of the *i*th ray from layer k. To suppress the image noise, a regularization term $R(\rho)$ is added, resulting in the following penalized likelihood objective function:

$$\Phi(\boldsymbol{\rho}) = -L(\boldsymbol{\rho}) + R(\boldsymbol{\rho}). \tag{3.17}$$

Herein, the edge-preserving Huber regularization is utilized:

$$R(\boldsymbol{\rho}) = \sum_{m=1}^{M} \sum_{j=1}^{J} \sum_{j' \in \mathcal{V}_j} \lambda^m \psi(\rho_j^m - \rho_{j'}^m), \qquad (3.18)$$

with

$$\psi(\Delta) = \begin{cases} \frac{\Delta^2}{2} & \text{if } |\Delta| \le \gamma \\ \gamma |\Delta| - \frac{\gamma^2}{2} & \text{if } |\Delta| > \gamma \end{cases},$$
(3.19)

in which λ^m is a factor that balances the regularization weight between different basis materials, \mathcal{V}_j is a neighborhood of the voxel j, γ is a tuning parameter that decides whether the regions use quadratic regularization (when the neighborhood difference is smaller than γ) or linear regularization (when the neighborhood difference is larger than γ).

Since the objective function in (3.17) is difficult to minimize directly, therefore, the optimization transfer principles^{24,25} is used to optimize a surrogate function, which is much more easier and efficient to minimize. The derivations of the surrogate functions are based on²⁶ and²⁷. In addition, an extra surrogate function that extracts the sub-ray related summation outside the negative log-likelihood objective function is introduced as well. To simplify the derivation, we define

$$h_{i}^{k}(\hat{y}_{i}^{k}) = \hat{y}_{i}^{k} - y_{i}^{k} \ln \hat{y}_{i}^{k}, \quad \hat{y}_{i}^{k} = \sum_{r \in \mathcal{N}_{i}^{k}} \hat{y}_{i}^{'kr},$$

$$l_{ie}^{krm}(\boldsymbol{\rho}^{m}) = f_{e}^{m} \sum_{j=1}^{J} a_{ij}^{kr} \rho_{j}^{m}, \qquad l_{ie}^{kr}(\boldsymbol{\rho}) = \sum_{m=1}^{M} l_{ie}^{krm},$$

$$t_{ie}^{kr}(\boldsymbol{\rho}) = \exp\left(-l_{ie}^{kr}\right),$$

$$\theta_{i}^{kr,(n)} = \frac{\hat{y}_{i}^{k,(n)}}{\hat{y}_{i}^{'kr,(n)}}, \quad \beta_{ie}^{kr,(n)} = \frac{\hat{y}_{i}^{'kr,(n)}}{t_{ie}^{kr,(n)}}.$$
(3.20)

Since it holds that:

$$\sum_{r \in \mathcal{N}_i^k} \frac{1}{\theta_i^{kr,(n)}} = 1, \quad \sum_{e=1}^E \frac{S_{ie}^{kr}}{\beta_{ie}^{kr,(n)}} = 1, \tag{3.21}$$

according to the Jensen's inequality²⁸, the integrals over the sub-ray r and spectrum k can be moved out of the negative log-likelihood function successively. As a consequence, two surrogate functions $Q_1(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})$ and $Q_2(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})$ can be obtained immediately:

$$-L(\boldsymbol{\rho}) = \sum_{k=1}^{K} \sum_{i=1}^{I} h_i^k \left(\sum_{r \in \mathcal{N}_i^k} \hat{y}_i^{\prime k r} \right)$$

$$\leq \sum_{k=1}^{K} \sum_{i=1}^{I} \frac{1}{\theta_i^{k r, (n)}} \sum_{r \in \mathcal{N}_i^k} h_i^{k r} \left(\hat{y}_i^{\prime k r} \theta_i^{k r, (n)} \right) \equiv Q_1(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)}), \qquad (3.22)$$

and

$$Q_{1}(\boldsymbol{\rho};\boldsymbol{\rho}^{(n)}) = \sum_{k=1}^{K} \sum_{i=1}^{I} \frac{1}{\theta_{i}^{kr,(n)}} \sum_{r \in \mathcal{N}_{i}^{k}} h_{i}^{kr} \left(\sum_{e=1}^{E} S_{ie}^{kr} t_{ie}^{kr} \theta_{i}^{kr,(n)}\right)$$

$$\leq \sum_{k=1}^{K} \sum_{i=1}^{I} \sum_{r \in \mathcal{N}_{i}^{k}} \sum_{e=1}^{E} \frac{S_{ie}^{kr}}{\theta_{i}^{kr,(n)} \beta_{ie}^{kr,(n)}} h_{ie}^{kr} \left(t_{ie}^{kr} \beta_{ie}^{kr,(n)} \theta_{i}^{kr,(n)}\right)$$

$$\equiv Q_{2}(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)}). \qquad (3.23)$$

Subsequently, $Q_2(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})$ is approximated into a quadratic surrogate function in terms of the attenuation line integral l_{ie}^{krm} . With the definition of $g_{ie}^{kr}(l_{ie}^{krm}) \equiv h_{ie}^{kr}\left(t_{ie}^{kr}\beta_{ie}^{kr,(n)}\theta_{i}^{kr,(n)}\right)$, g_{ie}^{kr} can be expanded into a Taylor series about the line integral l_{ie}^{krm} up to the second order:

$$g_{ie}^{kr}(l_{ie}^{krm}) \approx q_{ie}^{kr}(l_{ie}^{krm}; l_{ie}^{krm,(n)}) = g_{ie}^{kr}(l_{ie}^{krm,(n)}) + \sum_{m=1}^{M} \left. \frac{\partial g_{ie}^{kr}}{\partial l_{ie}^{krm}} \right|_{l_{ie}^{krm} = l_{ie}^{krm,(n)}} (l_{ie}^{krm} - l_{ie}^{krm,(n)}) + \frac{1}{2} \sum_{c,d=1}^{M} T_{ie}^{krcd,(n)}(l_{ie}^{krc} - l_{ie}^{krc,(n)})(l_{ie}^{krd} - l_{ie}^{krd,(n)}),$$
(3.24)

where $l_{ie}^{krm,(n)} = l_{ie}^{krm}(\boldsymbol{\rho}^{m,(n)})$. Since it is difficult to choose a curvature $T_{ie}^{krcd,(n)}$ to fulfill all the conditions in the optimization transfer principles, therefore, the second derivative of $g_{ie}^{kr}(l_{ie}^{krm})$ is applied to replace the curvature by relaxing the monotonicity according to^{27,29}. Combining (3.23) and (3.24), the third surrogate function $Q_3(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})$ is obtained:

$$Q_{3}(\boldsymbol{\rho};\boldsymbol{\rho}^{(n)}) = \sum_{k=1}^{K} \sum_{i=1}^{I} \sum_{r \in \mathcal{N}_{i}^{k}} \sum_{e=1}^{E} \frac{S_{ie}^{kr}}{\theta_{i}^{kr,(n)} \beta_{ie}^{kr,(n)}} q_{ie}^{kr}(l_{ie}^{krm}; l_{ie}^{krm,(n)}).$$
(3.25)

Next, the line integral l_{ie}^{krm} is made separable over discrete density voxels ρ_j^m . According to²⁶, the surrogate function $Q_4(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})$ can be obtained:

$$Q_4(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)}) = \sum_{k=1}^K \sum_{i=1}^I \sum_{r \in \mathcal{N}_i^k} \sum_{e=1}^E \sum_{j=1}^J \frac{\omega_{ij}^{kr} S_{ie}^{kr}}{\theta_i^{kr,(n)} \beta_{ie}^{kr,(n)}} q_{ie}^{kr} \left(\xi_{ije}^{krm}(\rho_j^m; \rho_j^{m,(n)})\right),$$
(3.26)

where

$$\omega_{ij}^{kr} = \frac{a_{ij}^{kr}}{\sum_{j}^{J} a_{ij}^{kr}},$$

$$\xi_{ije}^{krm}(\rho_{j}^{m};\rho_{j}^{m,(n)}) = \frac{a_{ij}^{kr}f^{m}}{\omega_{ij}^{kr}}\left(\rho_{j}^{m}-\rho_{j}^{m,(n)}\right) + l_{ie}^{krm,(n)}.$$
(3.27)

A separable surrogate function for the regularization term can be derived 29 as follows:

$$\mathcal{R}(\boldsymbol{\rho};\boldsymbol{\rho}^{(n)}) = \frac{1}{2} \sum_{m=1}^{M} \sum_{j=1}^{J} \sum_{j' \in \mathcal{V}_j} \lambda^m \left[\psi(2\rho_j^m - \rho_j^{m,(n)} - \rho_{j'}^{m,(n)}) + \psi(2\rho_{j'}^m - \rho_j^{m,(n)} - \rho_{j'}^{m,(n)}) \right].$$
(3.28)

With the above separable surrogate functions for both $-L(\rho)$ and $R(\rho)$ in (3.17), the final surrogate objective function becomes $Q_4(\rho; \rho^{(n)}) + \mathcal{R}(\rho; \rho^{(n)})$. We use the Newton-Raphson method to iteratively minimize the objective function:

$$\boldsymbol{\rho}^{(n+1)} = \boldsymbol{\rho}^{(n)} - (H_{Q_4}^{(n)} + H_{\mathcal{R}}^{(n)})^{-1} \cdot (\nabla Q_4 + \nabla \mathcal{R})|_{\boldsymbol{\rho} = \boldsymbol{\rho}^{(n)}}, \qquad (3.29)$$

where $H_{Q_4}^{(n)}$ and $H_{\mathcal{R}}^{(n)}$ denote the Hessian matrices of $Q_4(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})$ and $\mathcal{R}(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})$ at point $\boldsymbol{\rho} = \boldsymbol{\rho}^{(n)}$, ∇Q_4 and $\nabla \mathcal{R}$ denote the gradient vectors. Detailed expressions can be found in the Appendix. To this end, the penalized likelihood objective function defined in (3.17) can be minimized and optimized readily. To reduce the computation burden in each iteration and to accelerate the convergence speed, the ordered subset optimization²⁹ and Nesterov's momentum technique²⁶ are applied.

C. Experiments

Two experiments were performed on our in-house CBCT imaging benchtop. The Xray source is a rotating-anode diagnostic grade X-ray tube (Varex G-242, Varex Imaging Corporation, UT, USA). In the first experiment, a SL-FPD (Varex 4343CB, Varex Imaging Corporation, UT, USA) with native pixel dimension of 0.139 mm \times 0.139 mm and matrix dimension of 3072×3072 was used. It was 3×3 binned during the rotation-by-rotation dualenergy scans with 75 kVp (1.5 mm aluminium(Al) plus 0.2 mm Cu filtration) and 125 kVp (1.5 mm Al plus 1.2 mm Cu filtration) beam settings. The source to detector distance (SDD) was 1560.6 mm, and the source to rotation center distance (SOD) was 1156.3 mm. Projections were collected from one full 360° rotation with an angular interval of 0.8°. In this experiment, a pig leg sample and the Catphan-700 phantom (CTP714 high resolution module) were scanned.



FIG. 3. The experimental dual-energy CBCT imaging benchtop with a dual-layer FPD.

The second experiment was performed with a DL-FPD prototype (CareView 560RF-DE, CareRay Digital Medical Technology Co., Ltd., China), as shown in Fig 3. This detector has 1536×1536 pixels with native pixel dimension of 0.154 mm \times 0.154 mm. The 2×2 detector binning mode was used during the data acquisition. The source to rotation center distance was 1475.0 mm, and the source to the top and bottom detector layers were 1612.8 mm and 1619.4 mm, respectively. An additional 1.0 mm Copper filtration was inserted between the top layer and the bottom layer to harden the X-ray beam. More detailed parameter settings are listed in Table II. The pig leg sample and an iodine phantom were scanned. For the iodine phantom, four iodine inserts with varied concentrations (2.5 mg/cm³, 5 mg/cm³, 10 mg/cm³, 20 mg/cm³) were emerged in a cylinder water tank with diameter of 10 cm. Be aware that this DL-FPD was not worked under the newly proposed suRi method with sub-pixel shifting since such innovative data acquisition protocol is still under development by the manufacturer.

Detector type	SL-FPD	DL-FPD	
Tube voltage (kVp)	75; 125	125	
Tube Current (mA)	8; 12.5	12.5	
Deem filtration (nom)	75 kVp: 1.5 Al, 0.2 Cu	0.4.0	
Deam intration (inin)	125 kVp: 1.5 Al, 1.2 Cu	0.4 Uu	
Total views	450	450	
SOD (mm)	1156.3	1475.0	
	1500 0	1612.8 (top)	
SDD (mm)	1000.0	1619.4 (bottom)	
Ino mo motivi	Pig leg: 512×512	384×384	
Image matrix	Catphan: 660×660		
	Pig leg: 0.25×0.25	0.28×0.28	
voxei size (mm)	Catphan: 0.31×0.31		
Detector array	3072×3072	1536×1536	
Pixel dimension (mm)	0.139×0.139	0.154×0.154	
Detector binning	3×3	2×2	
Manual binning	1×2	1×2	
Sub-pixel shift (mm)	0.417	0.308	
CaLTI thickness (ram)	0.75	0.26 (top)	
USI: 11 UNICKNESS (IIIIII)	0.70	0.55 (bottom)	

TABLE II. Experimental data acquisition parameters.

For both experiments, the X-ray beam was collimated into a narrow width (<20 mm on the detector surface) to minimize the Compton scatters. Moreover, additional 1×2 binning was implemented manually on the projections. To mimic the 1D sub-pixel shifting, the

high-energy projection (bottom layer) is shifted by half binned pixel dimension relative to the low-energy projection (top layer) along the horizontal direction. Unless specified, the binning mode mentioned in the rest of the manuscript refers to the manual binning.

D. Comparison method and parameter selection

The conventional image domain decomposition (IDD) method was implemented for comparison. In brief, the dual-energy CT images were reconstructed with the FBP algorithm, then the density maps of two basis materials were generated via pixel-wise matrix inversion of the equation below:

$$\begin{bmatrix} f^{1,\text{LE}} & f^{2,\text{LE}} \\ f^{1,\text{HE}} & f^{2,\text{HE}} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \begin{bmatrix} \mu^{\text{LE}} \\ \mu^{\text{HE}} \end{bmatrix}, \qquad (3.30)$$

where $f^{m,k}$ (m = 1, 2; k = LE, HE) denotes the effective mass attenuation coefficient of material b at spectrum k, which is estimated by:

$$f^{m,k} = \frac{\sum_{e=1}^{E} S_e^k f_e^{m,k}}{\sum_{e=1}^{E} S_e^k}.$$
(3.31)

Moreover, the same material decomposition algorithm used for suRi was also implemented on the aligned binning data. The obtained results are referred to as DD-align.

Some key parameters need to be determined for the proposed suRi algorithm, including the number of sub-rays in \mathcal{N}_i^k , the size of neighboring pixels \mathcal{V}_j in Huber regularization, the tuning parameter γ , the regularization weight λ^m and the number of iterations N_{iter} . Empirically, we set the size of \mathcal{V}_j equal to 3×3 . Increasing the number of sub-rays would result in more accurate imaging model, however, this would also cause intense computation burden. To balance the trade-off, 8 sub-rays was selected. The regularization weight λ^m is critical to the image spatial resolution: the basis images would become blurred if λ^m gets too large, and they would become very noisy if λ^m gets too small. The specific values of λ^m , γ and N_{iter} are listed in Table III. Notice that the same parameters were used for the DD-align and suRi methods.

E. Characterization of image spatial resolution

To characterize the spatial resolution, the modulation transfer function (MTF) was calculated upon the reconstructed Catphan CT images. The method proposed by Droege et

Parameters	SL-FF	D CT	DL-FPD CT	
	Pig leg	Catphan	Pig leg	Iodine
λ^m	[3, 3]	[1, 1]	[1, 1]	$[10^2, 10^5]$
γ	[0.1, 0.1]	[0.1,0.1]	[0.1,0.1]	[0.1, 0.001]
N_{iter}	40	70	40	100

TABLE III. Iterative parameters used by suRi.

 $al.^{30,31}$ was used to determine the MTF from the line-pair patterns:

$$MTF(\mathcal{F}) = \frac{\pi}{\sqrt{2}M_0} \sqrt{\sum_n C_n \frac{V(n\mathcal{F})}{n^2}}, \qquad n = 1, 3, 5, \dots$$

with
$$V(n\mathcal{F}) = V'(n\mathcal{F}) - V_0,$$

$$C_n = \begin{cases} -1 & \text{if } p' = 1 \text{ and } n > 1\\ 0 & \text{if } p > p' \qquad , \\ 1 & \text{otherwise} \end{cases}$$
(3.32)

where \mathcal{F} represents the spatial frequency (unit: lp/mm), $M_0 = |\mu_1 - \mu_2|$ is the pixel intensity difference between the largest bar and its adjacent background regions, $V'(n\mathcal{F})$ denotes the variance measured within a region-of-interest (ROI) of the bar pattern with spatial frequency $n\mathcal{F}$, V_0 denotes the variance of the CT image values within an uniform ROI, C_n depends on the total number of prime factors p and the number of different prime factors p' of the number n. The spatial frequency $\mathcal{F}_{20\%}$ where the MTF decreases to 0.2 was calculated.

IV. RESULTS

A. Results obtained from SL-FPD

The decomposition results of the pig leg specimen are shown in Fig. 4. Images in the first and second columns are reconstructed from the projections with aligned 1×1 binning and sub-pixel shifted 1×2 binning using the IDD method, images in the third and fourth columns are reconstructed from the projections with aligned 1×2 binning and sub-pixel shifted 1×2 binning using the proposed material decomposition algorithm for suRi. As seen, the water and bone basis images obtained from the IDD method contain higher image noise than those obtained from the iterative image reconstruction method. Besides,



FIG. 4. The decomposed dual-energy imaging results of the pig leg specimen data acquired from the SL-FPD. Images in the first column are the reference with high spatial resolution. As seen, the images in the fourth column also show very high spatial resolution as the reference images. Images in the first and second rows are the decomposed water and bone density maps, respectively. The display windows are [0.5, 1.5] g/cm³ and [1, 1.5] g/cm³. Images in the third and fourth rows correspond to the monochromatic images synthesized from the generated basis material images. The display window is [0, 0.45] cm⁻¹. The display windows for the zoomed-in ROIs in the third and fourth rows are [0.18, 0.42] cm⁻¹ and [0.16, 0.30] cm⁻¹, respectively. The scale bar denotes 20 mm.

some streaking artifacts may appear in Fig 4(b) and (f), as highlighted by the red arrows. These streaking artifacts are partially mitigated in the images reconstructed by the DD-align method, and most of them are efficiently suppressed by the suRi method. By comparing the monochromatic images in Fig 4(i) to (p), one can clearly observe that the proposed suRi method is able to recover the fine bony structures (shown in the zoomed-in ROIs) from the 1×2 binned data to a similar level as those reconstructed from the 1×1 binned data. Without the sub-pixel shifting, however, CT images of such high spatial resolution can not be generated, see those blurry CT images obtained from the DD-align method with the aligned 1×2 binning data.

The imaging results of the Catphan phantom are shown in Fig. 5. Similar to the pig leg results, the suRi method is capable to generate basis images and monochromatic images with higher spatial resolution than the IDD and DD-align method. The zoomed-in ROIs with line pairs in Fig. 5(1) and (p) show similar visual performance compared to those in (i) and (m). The MTF curve of the dual-energy FBP images and the synthesized 50 keV monochromatic images are plotted in Fig. 6. Quantitatively, the image reconstructed from the 1 × 2 binned data using the FBP method has the lowest MTF value. The penalized likelihood algorithm helps to improve the image spatial resolution: the MTF values of the DD-align and suRi method at frequency less than 0.6 lp/mm are similar, and both are higher than the FBP method (1 × 2 binned data). In regard to the high spatial frequency region, the suRi method outperforms the DD-align method due to the sub-pixel shift in the 1 × 2 binned projection data . The $\mathcal{F}_{20\%}$ of the FBP (1 × 1), FBP (1 × 2), DD-align and suRi correspond to 1.13 lp/mm, 0.78 lp/mm, 0.98 lp/mm and 1.14 lp/mm, respectively. As a results,the suRi method improves the spatial resolution by 46.15% compared with the FBP reconstruction (1 × 2 binned) method, and 16.33% compared with the DD-align method.

B. Results obtained from DL-FPD

Figure 7 shows the dual-energy imaging results of the pig leg specimen obtained from the DL-FPD. Images in the first and second columns are dual-energy CT images reconstructed by the FBP algorithm. Images in the rest columns are monochromatic CT images synthesized from the IDD, DD-align and suRi method. Since the CsI:TI scintillator in the top and bottom layers have different thicknesses, slight spatial resolution difference in the



FIG. 5. The dual-energy imaging results of the Catphan phantom acquired from the SL-FPD. Images in the first column are the reference with high spatial resolution. As seen, the images in the fourth column also show very high spatial resolution as the reference images. Images in the first and second rows are the decomposed water and aluminum density maps, respectively. The display windows are [0, 1.6] g/cm³ and [0, 1] g/cm³. Images in the third and fourth rows correspond to the monochromatic images synthesized from the generated basis images. The display window is [0, 0.5] cm⁻¹. The display windows for the zoomed-in ROIs in the third and last row are [0.20, 0.60] cm⁻¹ and [0.17, 0.39] cm⁻¹, respectively. The scale bar denotes 30 mm.

FBP reconstructed low-energy and high-energy CT images are observed, see Fig. 7(a) - (d). The synthesized monochromatic CT images from the IDD method look similar to the FBP reconstructed LE and HE CT images. Specifically, the IDD reconstructed CT images from 1×2 binned projections show worse spatial resolution than the results obtained from the



FIG. 6. The MTF curves of the FBP reconstructed low-energy CT images and the monochromatic (50 keV) CT images synthesized from the DD-align and suRi method.



FIG. 7. Imaging results of the pig leg data obtained from the DL-FPD. The FBP reconstructed low-energy and high-energy CT images are shown on the left two columns, and the synthesized monochromatic CT images using the IDD and suRi algorithms are shown on right four columns. The display window is [0.15, 0.37] cm⁻¹ for the low-energy CT images and the monochromatic 65 keV images, and is [0.15, 0.33] cm⁻¹ for the high-energy CT images and the monochromatic 85 keV images. The scale bar denotes 20 mm.

 1×1 binned projections. The DD-align method (1×2 binned data) slightly enhances some of the fine bony structures, see the highlighted red circles, but it is still difficult to recover the full image resolution. On the contrary, the proposed suRi method generates the best CT images with the highest spatial resolution, which are visually comparable to ones



reconstructed from the 1×1 binned projections by the FBP algorithm.

FIG. 8. The dual-energy imaging results of the iodine phantom obtained from the DL-FPD. Images in the first and second rows are the decomposed water and iodine density maps, respectively. The display windows are [0, 1.2] g/cm³ and [0, 25] mg/cm³. Images in the third and fourth rows are the synthesized monochromatic images from the generated basis material images. The display window is [0, 0.3] cm⁻¹. The insert with 20 mg/cm³ iodine is zoomed-in. The scale bar denotes 20 mm.

The imaging results of the iodine phantom are shown in Fig. 8. For all methods, the iodine material can be well separated from water. Compared with the IDD algorithm, the DD-align and suRi method could efficiently suppress the noise of the basis images. The mean values and standard deviations (STDs) of the estimated iodine concentrations are plotted in Fig. 9. It is found that the IDD method slightly underestimate the iodine densities, and the decomposed results of the DD-align and suRi methods are more consistent with the ground

truth. Quantitatively, the average relative errors of the four methods are -14.77%, -14.71%, 11.26% and 6.26%, respectively. The small decomposition error demonstrates the viability of the proposed reconstruction algorithm for suRi in generating accurate dual-energy basis images.



FIG. 9. The mean values and standard deviations of the iodine concentrations $(2.5 - 20 \text{ mg/cm}^3)$ measured by the IDD, DD-align and suRi method.

V. DISCUSSIONS AND CONCLUSION

In this paper, a sub-pixel shifting technique, which is referred to as suRi, is firstly proposed for the DL-FPD based dual-energy CBCT to generate high spatial resolution and potentially high temporal resolution material-specific basis images. In suRi, the low-energy and highenergy projection data acquired from the top and the bottom detector layers contain varied spatial information. A dedicated penalized likelihood algorithm is developed to reconstruct high spatial resolution basis images directly from these projections containing sub-pixel shifts. Physical experiments from a SL-FPD and a DL-FPD were conducted to evaluate the performance of suRi. Results demonstrate that the proposed suRi method is capable to generate dual-energy CBCT images with high spatial resolution. Compared with the conventional FBP reconstruction $(1 \times 2 \text{ binning})$, the suRi method with sub-pixel shifted binning technique could improve the spatial resolution by 46.15% (at 20% MTF).

The small gap (approximately 6.6 mm) between the top and bottom layers in the DL-FPD would cause intrinsic geometrical mismatches to the acquired dual-energy CBCT projections. Such geometrical mismatches correspond to the minor pixel shifts when X-rays are incident

with large cone angles, see the illustration in Figure 10(a). As shown in Figure 10(b) and (c), such relative pixel shifts are proportional to the layer gap, but inversely proportional to the binned pixel size. As a result, the proposed suRi method would generate the highest spatial resolution improvement only in the following cases: (1) Small object size, in which the field-of-view (FOV) is mostly captured by the centeral area of the detector. (2) Small gap distance between layers. It is observed from Fig 10(b) that the intrinsic pixel shift can be reduced by a factor of 3.3 if the gap distance is decreased from 6.6 mm down to 2 mm³². (3) Large pixel binning size. As shown in Fig 10(c), this would result in reduced intrinsic pixel shift. Consequently, the suRi method would become especially useful when high acquisition speed is required.



FIG. 10. (a) Illustration of the intrinsic pixel shift in dual-layer detector due to the oblique incident X-ray beam. The intrinsic pixel shift between two layers with varied (b) gap distance w_{gap} and (c) pixel binning size B_{bin} . Notice that the DL-FPD used in this study has $w_{gap} = 6.6$ mm and $B_{bin} = 4$ (combination of both the detector binning and manual binning).

The present work has several limitations. First, only small sized objects with 10 cm diameter were imaged by our DL-FPD to demonstrate the super resolution imaging feasibility of the proposed suRi method. This corresponds to small FOV and thus negligible intrinsic pixel shifts. Second, only the half-pixel shift along the horizontal direction is manipulated, and the two-dimensional half-pixel shifts along both the horizontal and vertical directions have not investigated yet. It would be studied as one of the future works. Third, the low-energy and high-energy CT images acquired with the DL-FPD have inconsistent spatial resolutions due to the varied thickness of the scintillator in each layer. This may lead to degraded image quality of the decomposed basis materials. To overcome this problem, model-based reconstruction algorithm that considers the projection data blur⁹, or novel DL-FPD structures with equal scintillator thickness³² could be investigated. Fourth, the proposed one-step material decomposition algorithm is sensitive to the value of the regularization weight λ^m of the basis materials. For example, the water basis image may get blurred when λ^m of the bone basis decreases, even if λ^m of the water basis keeps unchanged. Careful selections of the regularization weights and development of a new cross-material regularizer^{33,34} may be viable solutions. Fifth, we assumed the Poisson distribution of the measured data in the likelihood model. Although this is a good approximation³⁵, the projection data may not faithfully follow the Poisson statistics because of the energy weighted signal integration in the FPD. Thus, a more realistic statistical model needs be designed in future to further improve the accuracy of the decomposed dual-energy CT images. Last, the potential advancements of suRi in increasing the temporal resolution for CBCT imaging is not investigated in this study. We believe this should be a very interesting topic and may bring changes to the current four-dimensional (4D) CBCT imaging applications such as the 4D neurovascular or cardiac imaging.

In future, we believe the proposed suRi method in this work would bring many new opportunities to innovate the current CBCT imaging systems. First, detector binning sizes that are larger than 1×2 (e.g., 6×6 or 8×8) could be utilized in DL-FPD to achieve even faster signal readout speed while obtaining CT images with moderate spatial resolution. If the signal readout speed is doubled, the entire CBCT scan time could be reduced by 50% assuming the same total number of projections, or alternatively, the total number of projections could be doubled (i.e., much finer angular sampling interval) when the entire CBCT scan time is fixed. For the former case, the temporal resolution would be greatly improved³⁶; for the latter case, the CBCT image quality would be significantly enhanced due to the mitigation of streaking artifacts induced by the insufficient angular sampling rate³⁷. Second, the suRi method can also be applied on the multi-layer FPD (ML-FPD), e.g., triple-layer FPD³⁸, based spectral CBCT imaging system. To do that, more complex subpixel shifting strategy needs to be investigated and designed, as well as the reconstruction algorithm to obtain multiple sets of high resolution spatial-spectral (-temporal) CT images. Third, the suRi method is inherently compatible with the 1×1 pixel binning mode in either SL-FPD or ML-FPD. For this special scenario, sub-pixel shifts among these layers have to be physically determined before laterally moving the SL-FPD¹³ or assembling the DL/ML-FPD hardwares. Fourth, the deep learning techniques, especially the convolution neural networks (CNN), can be investigated to learn the high resolution LE and HE projections in the sinogram domain³⁹, or to reconstruct the high resolution basis material CT images directly from the sub-pixel shifted projections⁴⁰. The utilization of CNN could significantly reduce the image reconstruction time compared with the iterative algorithms, and thus make the suRi method more viable in practice. Fifth, the flying focal spot technique can be combined with the suRi method to further improve the spatial (temporal) resolution of the reconstructed CT images. Finally, the proposed suRi method would also benefit the DL/ML-FPDs made from the latest IGZO (indium gallium zinc oxide) based X-ray receptors.

In conclusion, a super resolution dual-energy CBCT imaging method, suRi, is developed based on the DL-FPD. It is capable to overcome the long running difficulty of achieving high spatial resolution and high signal readout speed at the same time in CBCT imaging. By using of the penalized likelihood material decomposition algorithm, experimental results demonstrate that high spatial resolution basis images can be directly reconstructed from the dual-energy projections containing sub-pixel shifts. As a result, we believe this developed suRi method would greatly enhance the medical imaging performance of the DL-FPD (or ML-FPD) based dual-energy (spectral) CBCT systems in future.

APPENDIX

The gradient of $Q_4(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})$ is

$$\frac{\partial Q_4(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})}{\partial \rho_j^m} \bigg|_{\boldsymbol{\rho} = \boldsymbol{\rho}^{(n)}} = \sum_{k=1}^K \sum_{r \in \mathcal{N}_i^k} \sum_{i=1}^I a_{ij}^{kr} \left(\frac{y_i^k}{\hat{y}_i^{k,(n)}} - 1 \right) \sum_{e=1}^E S_{ie}^{kr} f_e^m t_{ie}^{kr}(\boldsymbol{\rho}^{(n)}).$$
(A.1)

The second derivative of the Hessian matrix of $Q_4(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})$ is

$$\frac{\partial^2 Q_4(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})}{\partial \rho_j^c \partial \rho_j^d} \bigg|_{\boldsymbol{\rho} = \boldsymbol{\rho}^{(n)}} = \sum_{k=1}^K \sum_{r \in \mathcal{N}_i^k} \sum_{i=1}^I a_{ij}^{kr} \left(\sum_{i=1}^J a_{ij}^{kr} \right) \sum_{e=1}^E \frac{S_{ie}^{kr} T_{ie}^{kr,(n)}}{\theta_i^{kr,(n)} \beta_{ie}^{kr,(n)}} f_e^c f_e^d.$$
(A.2)

By substituting the curvature $T_{ie}^{kr,(n)}$ with the second derivative $\partial^2 g_{ie}^{kr}(l_{ie}^{krm})/\partial l_{ie}^{krc}\partial l_{ie}^{krd} = \hat{y}_i^{k,(n)}$, and combining the definition of $\theta_i^{kr,(n)}$ and $\beta_{ie}^{kr,(n)}$ in (3.20), we have:

$$\frac{\partial^2 Q_4(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})}{\partial \rho_j^c \partial \rho_j^d} \bigg|_{\boldsymbol{\rho} = \boldsymbol{\rho}^{(n)}} = \sum_{k=1}^K \sum_{r \in \mathcal{N}_i^k} \sum_{i=1}^I a_{ij}^{kr} \left(\sum_{i=1}^J a_{ij}^{kr} \right) \times$$

$$\sum_{e=1}^{E} S_{ie}^{kr} f_e^c f_e^d \exp\left(-\sum_{m=1}^{M} f_e^m \sum_{j=1}^{J} a_{ij}^{kr} \rho_j^{m,(n)}\right).$$
(A.3)

The gradient of the surrogate function of the regularization term is

$$\frac{\partial \mathcal{R}(\boldsymbol{\rho};\boldsymbol{\rho}^{(n)})}{\partial \rho_j^m}\bigg|_{\boldsymbol{\rho}=\boldsymbol{\rho}^{(n)}} = \sum_j^J \sum_{j'\in\mathcal{V}_j} \lambda^m \psi'(\rho_j^{m,(n)} - \rho_{j'}^{m,(n)}), \tag{A.4}$$

with

$$\psi'(\Delta) = \begin{cases} \Delta & \text{if } |\Delta| \le \gamma \\ |\gamma| & \text{if } |\Delta| > \gamma \end{cases}.$$
 (A.5)

The second derivative of the Hessian matrix of $\mathcal{R}(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})$ is denoted as:

$$\frac{\partial^2 \mathcal{R}(\boldsymbol{\rho}; \boldsymbol{\rho}^{(n)})}{\partial \rho_j^c \partial \rho_j^d} \Big|_{\boldsymbol{\rho} = \boldsymbol{\rho}^{(n)}} = \begin{cases} \sum_j^J \sum_{j' \in \mathcal{V}_j} 2\lambda^m \psi''(\rho_j^{c,(n)} - \rho_{j'}^{d,(n)}) \\ & \text{if } c = d \\ 0 & \text{if } c \neq d \end{cases}$$
(A.6)

with

$$\psi''(\Delta) = \begin{cases} 1 & \text{if } |\Delta| \le \gamma \\ 0 & \text{if } |\Delta| > \gamma \end{cases}$$
(A.7)

REFERENCES

- ¹R. Schulze and N. Drage, "Cone-beam computed tomography and its applications in dental and maxillofacial radiology," Clinical radiology **75**, 647–657 (2020).
- ²M. Oldham, D. Létourneau, L. Watt, G. D. Hugo, D. Yan, D. Lockman, L. H. Kim, P. Y. Chen, A. A. Martinez, and J. W. Wong, "Cone-beam-ct guided radiation therapy: A model for on-line application." Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology **75 3**, 271–8 (2005).
- ³M. J. Wallace, M. D. Kuo, C. B. Glaiberman, C. A. Binkert, R. C. Orth, and G. Soulez, "Three-dimensional c-arm cone-beam ct: applications in the interventional suite." Journal of vascular and interventional radiology : JVIR **19 6**, 799–813 (2008).
- ⁴"Paxscan 4030x product description," https://www.vareximaging.com/wp-content/ uploads/2022/01/4030Xpds_0.pdf, accessed: 2022-09-01.
- ⁵N. M. Sheth, A. Uneri, P. A. Helm, W. Zbijewski, and J. H. Siewerdsen, "Technical assessment of 2d and 3d imaging performance of an igzo-based flat-panel x-ray detector," Medical physics **49**, 3053–3066 (2022).

- ⁶W. Zbijewski, G. Gang, J. Xu, A. Wang, J. Stayman, K. Taguchi, J. Carrino, and J. Siewerdsen, "Dual-energy cone-beam ct with a flat-panel detector: Effect of reconstruction algorithm on material classification," Medical physics **41**, 021908 (2014).
- ⁷L. Shi, M. Lu, N. R. Bennett, E. Shapiro, J. Zhang, R. Colbeth, J. Star-Lack, and A. S. Wang, "Characterization and potential applications of a dual-layer flat-panel detector," Medical physics **47**, 3332–3343 (2020).
- ⁸F. Ståhl, D. Schäfer, A. Omar, P. van de Haar, F. van Nijnatten, P. Withagen, A. Thran, E. Hummel, B. Menser, Å. Holmberg, *et al.*, "Performance characterization of a prototype dual-layer cone-beam computed tomography system," Medical Physics **48**, 6740–6754 (2021).
- ⁹W. Wang, Y. Ma, M. Tivnan, J. Li, G. J. Gang, W. Zbijewski, M. Lu, J. Zhang, J. Star-Lack, R. E. Colbeth, *et al.*, "High-resolution model-based material decomposition in duallayer flat-panel cbct," Medical physics **48**, 6375–6387 (2021).
- ¹⁰M. Kachelriess, M. Knaup, C. Penssel, and W. A. Kalender, "Flying focal spot (ffs) in cone-beam ct," IEEE Transactions on Nuclear Science 53, 1238–1247 (2006).
- ¹¹T. G. Flohr, K. Stierstorfer, S. Ulzheimer, H. Bruder, A. Primak, and C. H. McCollough, "Image reconstruction and image quality evaluation for a 64-slice ct scanner with z-flying focal spot." Medical physics **32** 8, 2536–47 (2005).
- ¹²H. Ur and D. Gross, "Improved resolution from subpixel shifted pictures," CVGIP Graph. Model. Image Process. 54, 181–186 (1992).
- ¹³Z. Yan, J. Li, Y. Lu, H. Yan, and Y. Zhao, "Super resolution in ct," International Journal of Imaging Systems and Technology 25, 92–101 (2015).
- ¹⁴Y. Li, J. Li, T. Wang, M. Lian, and S. Luo, "Micro-ct imaging of super-resolution mbir algorithm based on sub-pixel displacement," in *Medical Imaging 2020: Physics of Medical Imaging*, Vol. 11312 (SPIE, 2020) pp. 916–923.
- ¹⁵T. P. Szczykutowicz, S. D. Rose, and A. Kitt, "Sub pixel resolution using spectral-spatial encoding in x-ray imaging," PloS one **16**, e0258481 (2021).
- ¹⁶R. E. Alvarez and A. Macovski, "Energy-selective reconstructions in x-ray computerised tomography," Physics in Medicine & Biology **21**, 733 – 744 (1976).
- ¹⁷J.-P. Schlomka, E. Roessl, R. Dorscheid, S. V. Dill, G. Martens, T. Istel, C. Bäumer, C. Herrmann, R. Steadman, G. Zeitler, A. Livne, and R. Proksa, "Experimental feasibility of multi-energy photon-counting k-edge imaging in pre-clinical computed tomography,"

Physics in Medicine & Biology 53, 4031 – 4047 (2008).

- ¹⁸T. Niu, X. Dong, M. Petrongolo, and L. Zhu, "Iterative image-domain decomposition for dual-energy ct," Medical physics **41**, 041901 (2014).
- ¹⁹P. R. S. Mendonça, P. Lamb, and D. V. Sahani, "A flexible method for multi-material decomposition of dual-energy ct images," IEEE Transactions on Medical Imaging **33**, 99– 116 (2014).
- ²⁰D. Zeng, L. Yao, Y. Ge, S. Li, Q. Xie, H. Zhang, Z. Bian, Q. Zhao, Y. Li, Z. Xu, D. Meng, and J. Ma, "Full-spectrum-knowledge-aware tensor model for energy-resolved ct iterative reconstruction," IEEE Transactions on Medical Imaging **39**, 2831–2843 (2020).
- ²¹C. Mory, B. Sixou, S. A. Si-Mohamed, L. Boussel, and S. Rit, "Comparison of five one-step reconstruction algorithms for spectral ct," Physics in Medicine & Biology **63** (2018).
- ²²B. Chen, Z. Zhang, D. Xia, E. Y. Sidky, and X. Pan, "Algorithm-enabled partial-angularscan configurations for dual-energy ct." Medical physics **45** 5, 1857–1870 (2018).
- ²³Y. Long and J. A. Fessler, "Multi-material decomposition using statistical image reconstruction for spectral ct," IEEE Transactions on Medical Imaging **33**, 1614–1626 (2014).
- ²⁴K. L. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions," Journal of Computational and Graphical Statistics 9, 1 – 20 (2000).
- ²⁵A. R. D. Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," IEEE Trans. Medical Imaging **14**, 132–137 (1995).
- ²⁶K. Mechlem, S. Ehn, T. Sellerer, E. Braig, D. Münzel, F. Pfeiffer, and P. B. Noël, "Joint statistical iterative material image reconstruction for spectral computed tomography using a semi-empirical forward model," IEEE Transactions on Medical Imaging **37**, 68–80 (2018).
- ²⁷T. Weidinger, T. M. Buzug, T. G. Flohr, S. Kappler, and K. Stierstorfer, "Polychromatic iterative statistical material image reconstruction for photon-counting computed tomography," International Journal of Biomedical Imaging **2016** (2016).
- ²⁸E. J. McShane, "Jensen's inequality," Bulletin of the American Mathematical Society 43, 521–527 (1937).
- ²⁹H. Erdogan and J. A. Fessler, "Ordered subsets algorithms for transmission tomography." Physics in medicine and biology **44 11**, 2835–51 (1999).
- ³⁰R. T. Droege and R. L. Morin, "A practical method to measure the mtf of ct scanners." Medical physics 9, 758–760 (1982).
- ³¹R. T. Droege and M. S. Rzeszotarski, "Modulation transfer function from the variance of

cyclic bar images," Optical Engineering 23, 230168 (1984).

- ³²S. Gu, H. Zhou, Z. Wang, and H. Gao, "Evaluation of x-ray and ct imaging using a duallayer flat-panel detector without interlayer metal filter," Chinese Journal of Stereology and Image Analysis **27**, 81–88 (2022).
- ³³W. S. Huh and J. A. Fessler, "Iterative image reconstruction for dual-energy x-ray ct using regularized material sinogram estimates," 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro , 1512–1515 (2011).
- ³⁴W. Wang, M. Tivnan, G. J. Gang, and J. W. Stayman, "Prospective prediction and control of image properties in model-based material decomposition for spectral ct," in *Medical Imaging 2020: Physics of Medical Imaging*, Vol. 11312 (SPIE, 2020) pp. 475–480.
- ³⁵B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, "An iterative maximumlikelihood polychromatic algorithm for ct," IEEE transactions on medical imaging 20, 999–1008 (2001).
- ³⁶G.-H. Chen and Y. Li, "Synchronized multiartifact reduction with tomographic reconstruction (smart-recon): A statistical model based iterative image reconstruction method to eliminate limited-view artifacts and to mitigate the temporal-average artifacts in timeresolved ct." Medical physics **42** 8, 4698–707 (2015).
- ³⁷J. Bian, J. H. Siewerdsen, X. Han, E. Y. Sidky, J. L. Prince, C. A. Pelizzari, and X. Pan, "Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam ct," Physics in Medicine & Biology 55, 6575 – 6599 (2010).
- ³⁸S. L. Maurino, A. Badano, I. A. Cunningham, and K. S. Karim, "Theoretical and monte carlo optimization of a stacked three-layer flat-panel x-ray imager for applications in multispectral diagnostic medical imaging," in *Medical Imaging 2016: Physics of Medical Imaging*, Vol. 9783 (SPIE, 2016) pp. 1061–1074.
- ³⁹C. Tang, W. Zhang, L. Wang, A. Cai, N. Liang, L. Li, and B. Yan, "Generative adversarial network-based sinogram super-resolution for computed tomography imaging," Physics in Medicine & Biology 65, 235006 (2020).
- ⁴⁰T. Su, X. Sun, J. Yang, D. Mi, Y. Zhang, H. Wu, S. Fang, Y. Chen, H. Zheng, D. Liang, and Y. Ge, "Direct-net: a unified mutual-domain material decomposition network for quantitative dual-energy ct imaging." Medical Physics 49, 917–934 (2022).