

# Statistical Iterative CBCT Reconstruction Based on Neural Network

Binbin Chen, Kai Xiang, Zaiwen Gong, Jing Wang<sup>ID</sup>, and Shan Tan<sup>ID</sup>, *Member, IEEE*

**Abstract**—Cone-beam computed tomography (CBCT) plays an important role in radiation therapy. Statistical iterative reconstruction (SIR) algorithms with specially designed penalty terms provide good performance for low-dose CBCT imaging. Among others, the total variation (TV) penalty is the current state-of-the-art in removing noises and preserving edges, but one of its well-known limitations is its staircase effect. Recently, various penalty terms with higher order differential operators were proposed to replace the TV penalty to avoid the staircase effect, at the cost of slightly blurring object edges. We developed a novel SIR algorithm using a neural network for CBCT reconstruction. We used a data-driven method to learn the “potential regularization term” rather than design a penalty term manually. This approach converts the problem of designing a penalty term in the traditional statistical iterative framework to designing and training a suitable neural network for CBCT reconstruction. We proposed using transfer learning to overcome the data deficiency problem and an iterative deblurring approach specially designed for the CBCT iterative reconstruction process during which the noise level and resolution of the reconstructed images may change. Through experiments conducted on two physical phantoms, two simulation digital phantoms, and patient data, we demonstrated the excellent performance of the proposed network-based SIR for CBCT reconstruction, both visually and quantitatively. Our proposed method can overcome the staircase effect, preserve both edges and regions with smooth intensity transition, and provide reconstruction results at high resolution and low noise level.

**Index Terms**—CBCT, image reconstruction, neural network, regularization term.

## I. INTRODUCTION

CONE-Beam computed tomography (CBCT) is an important technique that provides volumetric information for

patient setup and target localization during radiation therapy. However, repeated exposure to CBCT scanning and extra radiation doses during treatment, may increase the risk of cancer and genetic defects in patients [1], [2]. Decreasing mAs (the product of tube current and exposure time) levels during CBCT acquisition is an optional method that reduces the radiation dose. However, low-dose images reconstructed with classic methods such as FDK may contain a high level of noise which dramatically degrades image quality [3], [4].

Many algorithms have been proposed to improve the quality of CBCT images with lower mAs level projections. Statistical iterative reconstruction (SIR) algorithms have shown better performance in low-dose CBCT imaging than the classic FDK method [5]. Different kinds of penalty terms can be applied to SIR. Of these, the total variation (TV) penalty is one of the most common studied ones. TV has shown a state-of-the-art ability in preserving edges and suppressing noise for CBCT reconstruction [6]–[10]. Although the TV penalty presents many advantages, a major disadvantage is its so-called staircase effect [11], [12]. TV penalizes the image gradient, regardless of the structures, and produces reconstruction results with piecewise-constant artifacts that negatively affect the correct interpretation of images in medical imaging. Recently, various higher-order differential operators were proposed to overcome the staircase effect in various inverse problems. Examples of operators include total generalized variation (TGV) [13], fourth order partial differential equation (PDE) [14], isotropic and anisotropic higher degree TV penalties [15], and the Schatten norms of the Hessian matrix [16]. Higher-order penalties can effectively suppress the staircase effect, but usually tend to slightly over-smooth object edges for CBCT reconstruction [17]–[19].

Besides the TV penalty, other kinds of regularization terms have been proposed to solve various medical imaging problems. In the last several years, regularization methods using sparse representation and dictionary learning have gained popularity for computed tomography imaging [20]–[22]. Compressed sensing has been employed especially for image reconstruction from highly under-sampled measurements to exploit image sparsity [23], [24]. Non-local self-similarity has also been explored widely to improve image quality in SIR [25]–[30]. Other regularization methods such as wavelet transform and low rank matrix decomposition have been used for medical image reconstruction [31], [32].

With the increasingly widespread use of deep learning in computer vision, interest has grown in applying deep learning to solve various inverse problems, such as image denoising and super-resolution. Zhang *et al.* [33] reported

Manuscript received April 8, 2018; revised April 17, 2018; accepted April 17, 2018. Date of publication April 25, 2018; date of current version May 31, 2018. This work was supported by the National Natural Science Foundation of China under Grant 61375018 and Grant 61672253. The work of J. Wang was supported in part by the Cancer Prevention and Research Institute of Texas under Grant RP160661 and in part by the National Institute of Biomedical Imaging and Bioengineering under Grant R01 EB020366. (Binbin Chen and Kai Xiang are co-first authors.) (Corresponding authors: Jing Wang; Shan Tan.)

B. Chen, K. Xiang, Z. Gong, and S. Tan are with the Key Laboratory of Image Processing and Intelligent Control of Ministry of Education of China, Huazhong University of Science and Technology, Wuhan 430074, China, and also with the School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: ace7chan@hust.edu.cn; u201314635@hust.edu.cn; m201472350@hust.edu.cn; shantan@hust.edu.cn).

J. Wang is with the Department of Radiation Oncology, The University of Texas Southwestern Medical Center, Dallas, TX 75390 USA (e-mail: jing.wang@utsouthwestern.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2018.2829896

0278-0062 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See [http://www.ieee.org/publications\\_standards/publications/rights/index.html](http://www.ieee.org/publications_standards/publications/rights/index.html) for more information.

that a simple convolution neural network (CNN) framework could address denoising with state-of-the-art performance. Also, they used the half quadratic splitting method to extend their CNN denoising model to other image restoration problems, such as image deblurring and image super-resolution [34]. Zhang *et al.* [35] applied the CNN denoiser to the gradient domain rather than the image intensity domain, and used the splitting method to solve the non-blind deblurring problem. Lefkimmiatis [36] integrated the non-local method into the network and proposed a non-local network for image denoising. Chakrabarti [37] used a neural network to learn the blur kernel in the frequency domain to solve the blind deblurring problem. Wieschollek *et al.* [38] extended the work of Chakrabarti [37] and coalesced multiple images to solve the deblurring problem. To address the super-resolution problem, new neural network techniques have been applied, including pyramid networks [39] and generative adversarial networks [40]. Many neural networks methods surpassed the classic schemes to solve various inverse problems.

Despite the success of deep learning in computer vision and image processing, only a few studies reported using deep learning for CT reconstruction [41]. Kang *et al.* [42] applied CNN to wavelet transform coefficients to improve the quality of low-dose CT images. Chen *et al.* [43] trained a three-layer network from low-dose images directly, leading to low-dose reconstruction results with similar quality as the normal-dose reconstruction results that use classic methods. Yang *et al.* [44] constructed an objective loss function of the denoising CNN using feature descriptors, which shifted the idea from style-transfer to CT reconstruction [45], [46]. They also used the Wasserstein generative adversarial network (WGAN) [47] for CT reconstruction to enhance image quality [48]. Most of these methods tried to solve CT reconstruction as a denoising problem. The networks involved in these methods were trained in an end-to-end way, which required many low-dose/high-dose CT image pairs, and needed to be retrained when the dose levels changed.

In this study, we proposed a novel neural network-based statistical iterative CBCT reconstruction method. Particularly, we designed a network that learns a “potential powerful regularization term” through data, instead of manually selecting the regularization term in SIR. The idea was to employ a deep neural network that played a similar role to that of the traditional regularization term in CBCT reconstruction. We did not only consider reconstruction as a denoising problem as reported earlier [43], [48], but also a deblurring problem, in contrast to previous methods. To this end, we constructed a new objective function, which considers the blurring effect usually seen in CBCT reconstruction. The half quadratic splitting method [49] was used to divide the reconstruction problem into two sub-problems. One sub-problem was the common reconstruction form with a tradeoff portion that constrains the reconstruction image to be similar to the deblurred image. Another sub-problem that can be regarded as a deblurring problem was solved using the CNN deblurring method. Benefiting from the iterative process, the proposed SIR method can retain information from the projection data and allow the deblurring process do not have to be fully functional in the

starting phase (with a high noise level) during the whole reconstruction process.

Our experimental results showed that the proposed network-based SIR not only suppressed the staircase effect, which was commonly seen in reconstruction results when using the TV penalty, but also preserved edges and removed noise. Our method obtained significantly better results than TV in most regions, and can produce reconstruction images with low noise level without largely sacrificing spatial resolution.

## II. METHOD

### A. PWLS Reconstruction

The X-ray CT projection data represents the line integration along the X-ray path  $l$  of the tissue attenuation, which can be calculated as the logarithm transform of the incident to detected intensities ratio [5]:

$$p = \ln \frac{N_0}{N} = \int u(x, y, z) dl, \quad (1)$$

where  $p$  is the projection,  $N_0$  is the incident intensity,  $N$  is the detected intensity, and  $u$  is the attenuation coefficient. The noise of the X-ray CT projection approximately follows a Gaussian distribution [5], [50]. The noise variance  $\sigma_i^2$  associated with the projection  $p_i$  at the  $i$ th detector bin can be determined as follows [5]:

$$\sigma_i^2 = \exp(\bar{p}_i) / N_{i0}, \quad (2)$$

where  $N_{i0}$  is the incident photon number, and  $\bar{p}_i$  is the mean value of the projection  $p_i$ . The purpose of CBCT reconstruction is to find out the optimal solution by minimizing the objective function [51]:

$$\Phi(u) = \frac{1}{2} (p - Au)^T \Sigma^{-1} (p - Au) + \beta R(u). \quad (3)$$

The first term in Eq. (3) corresponds to the weighted least-squares (WLS) criterion.  $A$  represents the projection matrix, and  $\Sigma$  is a diagonal matrix with its  $i$ th element  $\sigma_i^2$ . The symbol  $T$  denotes the transpose operator. The second term  $R(u)$  is a penalty term. The penalty term is generally a constraint over the solution. The TV and Huber penalties are two typical regularization examples in CBCT reconstruction. Importantly, our penalty term is learned via the neural network and therefore, does not take an explicit form.  $\beta$  is a parameter that balances the WLS and penalty terms.

### B. The Blurring Model in PWLS Reconstruction

The spatial resolution in a CBCT image depends on factors such as size of the focal spot and detector channels, and also the amount of channel-to-channel crosstalk. The in-plane spatial resolution in CT is typically defined as the value at which the modulation transfers function (MTF) reaches a given percentage of its maximum. Current clinical CT scanners have an in-plane spatial resolution between 5 lp/cm and 15 lp/cm (at 10% MTF), depending on the resolution kernel [52].

To take the blurring effect into account, we introduced blur kernel  $G$ . Therefore, the PWLS function that considered the blur effect (we called BPWLS) can be expressed as follows:

$$\Phi(u) = (p - AGu)^T \Sigma^{-1} (p - AGu) + \beta R(u). \quad (4)$$

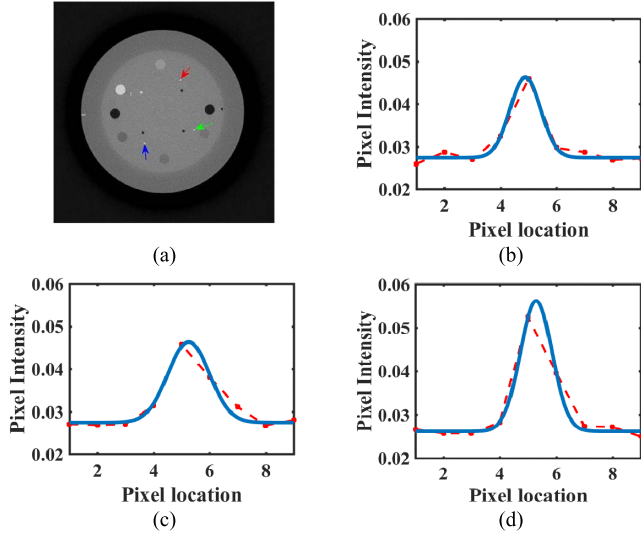


Fig. 1. (a) High-dose CatPhan 600; Gaussian fitting of source points: (b) corresponding to the blue arrow, (c) corresponding to the red arrow, (d) corresponding to the green arrow. The blue curves indicate the fitting curve, and the red dotted curves represent the attenuation coefficients along the line segments centered at the source points.

To integrate CNN into SIR, we used the variable splitting technique (half quadratic splitting). Introducing an auxiliary variable  $x$ , Eq. (4) can be reformulated as follows [49]:

$$\Phi(u, x) = (p - Ax)^T \Sigma^{-1} (p - Ax) + \lambda \|x - Gu\|_2^2 + \beta R(u), \quad (5)$$

where  $\lambda$  is a penalty parameter that varies with the phantoms but was fixed in iterations for each individual phantom in our experiments. Minimizing  $\Phi(u, x)$  in Eq. (5) over the auxiliary variable  $x$  and the attenuation coefficient  $u$ , respectively, yields the joint minimization problem [49]:

$$x_{sol} = \arg \min_x (p - Ax)^T \Sigma^{-1} (p - Ax) + \lambda \|x - Gu\|_2^2, \quad (6)$$

$$u_{sol} = \arg \min_u \lambda \|x - Gu\|_2^2 + \beta R(u). \quad (7)$$

Eq. (6) represents a standard reconstruction problem with a weighting parameter  $\lambda$  that balances the deblurred and reconstruction images, while Eq. (7) can be seen as a deblurring problem. In other words, the BPWLS problem can be solved iteratively between a pure reconstruction sub-problem and a deblurring sub-problem.

### C. Blur Kernel Estimate

Using the Gaussian fitting method [52], we tested the degree of blurring in both high-dose and low-dose FDK (with a Hanning filter kernel) reconstruction results of the CatPhan 600 phantom. The source points we selected are reported in Fig. 1 (a). The corresponding values of the full width at half maximum (FWHM) are listed in Table I, which shows that PSFs varies with spatial location in the reconstruction images. The high-dose reconstruction results had slightly higher spatial resolution.

TABLE I

BLUR DEGREE (FWHM) OF SELECTED LOCATIONS IN CATPHAN 600

Source Point	High-dose	Low-dose
Green	1.32	1.17
Blue	1.28	1.93
Red	1.74	1.67
Mean	1.45	1.59

We used the *mean* PSF of multiple source points in the high-dose FDK reconstruction results to approximate the blurring kernel. The reasons are as follows. First, the low-dose FDK reconstruction contains high noise levels, which might affect PSF estimation accuracy. Second, estimating a PSF map for the whole field of view would be hard and time-consuming. We assumed that the whole image has the same PSF for simplicity. Third, our experiments showed that the blurring kernel that was estimated using this method worked well. The kernel size was chosen to be  $9 \times 9$ .

### D. Pure Reconstruction

The objective function of Eq. (6) can be written as:

$$\Phi_1(x) = (p - Ax)^T \Sigma^{-1} (p - Ax) + \lambda \|x - Gu\|_2^2, \quad (8)$$

and its derivative is expressed as:

$$\frac{\partial \Phi_1(x)}{\partial x} = 2A^T \Sigma^{-1} (Ax - p) + 2\lambda (x - Gu). \quad (9)$$

The objective function  $\Phi_1(x)$  is convex, and its minimizer can be obtained by  $\partial \Phi_1(x)/\partial x = 0$ , which can be solved using the Gauss-Seidel update strategy [51]. The initialization of the Gauss-Seidel iterative algorithms was set as the FDK reconstruction result, and the parameters in the Gauss-Seidel iterative algorithms were set as:

$$\begin{aligned} \hat{x} &= FDK(p), \\ \hat{r} &= p - A\hat{x}, \\ \Sigma &= \text{diag}\{\sigma_i^2(p_i)\} = \text{diag}\{\exp(p_i)/N_{i0}\}, \\ s_j &= a_j^T \Sigma^{-1} a_j, \end{aligned} \quad (10)$$

where  $\hat{r}$  and  $s_j$  are intermediate variables and  $a_j$  is the  $j$ th column of  $A$ . In the calculation of  $\Sigma$  reported above, we used the measured projection  $p_i$  to approximate the mean projection  $\bar{p}_i$  since the mean projection was unknown in practice, as previously reported [19], [51]. For each  $j$ , the iterative process is expressed as:

$$\begin{aligned} \hat{x}_j^{old} &= \hat{x}_j, \\ \alpha_j &= s_j + \lambda, \\ \hat{x}_j^{new} &= \left( a_j^T \Sigma^{-1} \hat{r} + s_j \hat{x}_j^{old} + \lambda (Gu)_j \right) / \alpha_j, \\ \hat{x}_j &= \max \{0, \hat{x}_j^{new}\}, \\ \hat{r} &= \hat{r} + a_j (\hat{x}_j^{old} - \hat{x}_j). \end{aligned} \quad (11)$$

In the reconstruction process, we adopted separable footprint (SF) projectors [53] to determine projection matrix  $A$ .



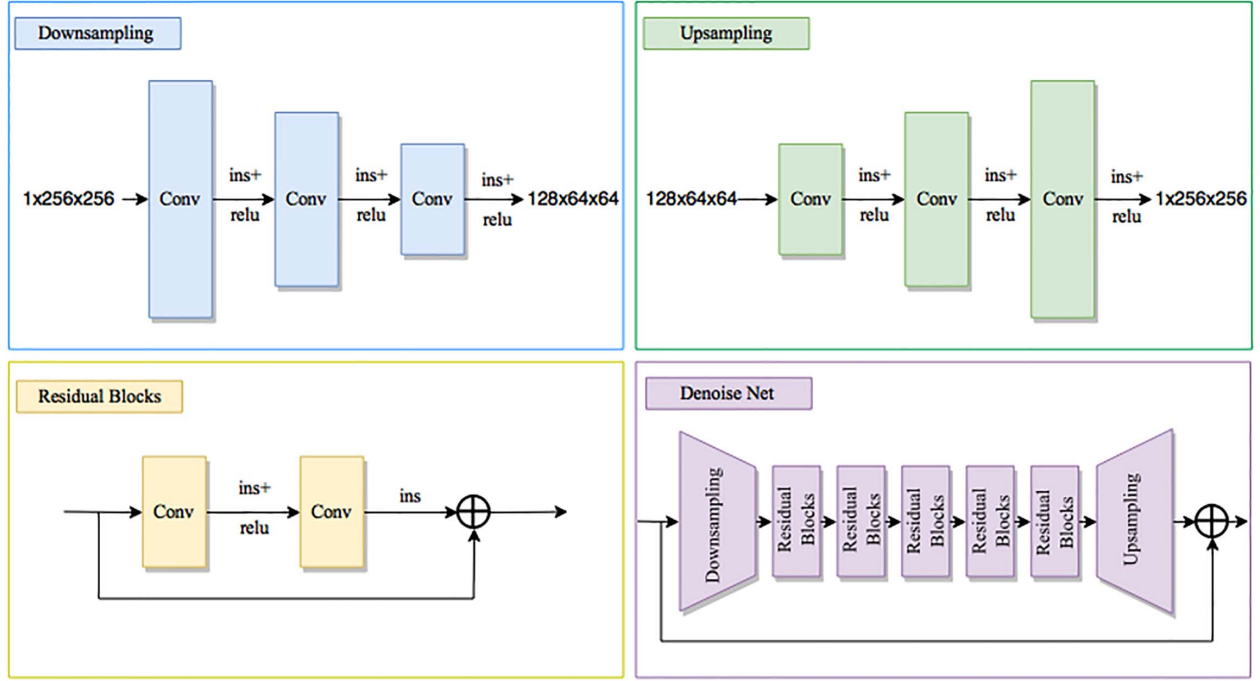


Fig. 2. The architecture of our denoising network. Ins represents the instance normalization, conv represents the convolution operation, and relu represents the rectified linear units.

We specifically used the SF projector with a trapezoid/rectangle function (SF-TR) [53]. SF-TR approximates the voxel footprint functions as 2D separable functions, and uses the trapezoid function in the transaxial direction and the rectangular function in the axial direction.

#### E. Deblurring Based on CNN

The objection function of Eq. (7) can be re-written as:

$$\Phi_2(u) = \|x - Gu\|_2^2 + \eta R(u), \quad (12)$$

with  $\eta = \beta/\lambda$ . This is a standard form of the deblurring problem. To fit the pure reconstruction sub-problem that generates reconstruction results with different blurring kernels and noise levels at different iterations, we used an iterative deblurring method [34] instead of other direct deblurring methods [37], [38].

Using again the half quadratic splitting method [49], Eq.(12) can be rewritten as:

$$\Phi_2(u, z) = \|x - Gu\|_2^2 + \eta R(z) + \xi \|z - u\|_2^2. \quad (13)$$

Eq. (13) can be solved through the following iterative process:

$$u_{sol} = \arg \min_u \|x - Gu\|_2^2 + \xi \|z - u\|_2^2, \quad (14)$$

$$\begin{aligned} z_{sol} &= \arg \min_z \xi \|z - u\|_2^2 + \eta R(z) \\ &= \arg \min_z \|z - u\|_2^2 + \varphi R(z), \end{aligned} \quad (15)$$

where  $\varphi = \eta/\xi$ . Eq. (14) is a quadratic regularized least square problem, which can be solved using many methods. We selected a Fast Fourier Transform (FFT) method for this problem because of its computational efficiency. More details

are reported in Appendix A. Eq. (15) corresponds to a denoising problem, which we solved by using a neural network method. To facilitate and reinforce the training phase, we combined the residual learning [34] and the image transfer net [46], [54]. The structure of our network is illustrated in Fig. 2.

The proposed network consists of one down-sampling, five residual blocks, and one up-sampling block. In the down-sampling and up-sampling blocks, we used the stride and fractional convolution instead of pooling layers, and used a reflection pad to keep the size. All non-residual convolutional layers were followed by instance normalization [54]. More details about the network are described in the following sections.

**1) Dataset:** Many datasets are available to train our network, including T91, BSDS200, General100 [39], and Imagenet [55]. We used the COCO dataset [56], which contains nearly 80 thousand images. We converted color images to gray images through the formula  $\text{Gray} = 0.299R + 0.587G + 0.114B$ .

**2) Inputs and Outputs:** Inputs included images from the COCO dataset, with the Gaussian noise of a certain level, and outputs included the corresponding Gaussian noise. Inputs and outputs were both randomly cropped gray images of  $1 \times 256 \times 256$ .

**3) Down-Sampling and Up-Sampling:** The down-sampling process used a large stride to shrink the input image from 256 to 64. The up-sampling process used the nearest neighboring pixel value to enlarge the image from 64 to 256. A reflection pad was used to keep the size in the convolution phase.

4) *Residual Blocks*: The residual blocks can help train a very deep network and enable the network to learn how to identify the function [57].

5) *Instance Normalization and Residual Learning*: With the help of instance normalization, we trained our network in different batch sizes, which is important to fully use GPU memory and obtain a similar performance in different batch sizes. Residual learning helped us focus on noise rather than the image itself, reducing training convergence time and also helping to design a general denoiser.

6) *Loss Function*: We used the simple least square form:

$$L = \frac{1}{M} \sum_{k=1}^M \|F(x_k; \Theta) - (x_k - y_k)\|_2^2, \quad (16)$$

where  $F$  is the neural network operation,  $M$  is the batch size,  $x_k, y_k$  are the  $k^{\text{th}}$  noise image, and the  $k^{\text{th}}$  original image in the batch, respectively. This loss function corresponds to the peak signal to noise (PSNR), and is also recommended to define the loss function in the feature layers [44], [45].

Our network presents several advantages when compared to DnCNN [33], [34]. First, the number of weights in the middle process was reduced by using the down-sampling and up-sampling processes, allowing us to design a deeper and more powerful network. Second, using instance normalization instead of batch normalization, we can use a smaller batch size, lightening our burden to select the image size and dividing the image into many patches. Third, other pre-processes were not needed in the images, including image enhancement or flipping. By using the ADAM optimization method [58] without any tricks applied to the training phase, our network achieved the same performance level as that of DnCNN. Finally, our network converged quickly, and the training took only 4 hours on GTX1080Ti.

Iteration-based deblurring needs various denoisers with various noise levels, but considering all the cases is challenging and time-consuming. Using an inexact sub-problem solution can be a reasonable method [59]. In our network, every single denoiser in the deblurring process does not need to be fully functional in each iteration. The deblurring process gradually improved its performance during the iterative process. The iterative method can also potentially help us use the Gaussian denoiser to handle the deblurring problem with other noise types because different types of noise tend to be statistically similar at low noise levels. We trained 25 denoisers in the noise variation level [0, 50], as previously reported by Zhang *et al.* [34]. Our experiment results showed that this range of noise level worked well.

#### F. Parameter Selection

The parameters that need to be set up in our algorithm can be divided into two groups: (a) Parameter for pure reconstruction,  $\lambda$ ; (b) Parameters for deblurring,  $\xi$  and  $\eta$ .

The objective function of the pure reconstruction sub-problem is  $\Phi_1(x)$  (see Eq. (8)), and  $\lambda$  is a parameter that balances the reconstruction  $(p - Ax)^T \Sigma^{-1} (p - Ax)$  and the deblurred part  $\|x - Gu\|_2^2$ . Theoretically, if  $\lambda$  is too large, the result of the Gauss-Seidel iterative algorithm  $x$  will be

enforced to approach  $Gu$ . If  $\lambda$  is too small, the reconstruction sub-problem and the deblurring sub-problem tend to become de-coupled. In our experiments, we empirically set  $\lambda$  in the near order of magnitude as  $s_j$ . Specifically, parameter  $\lambda$  was set to 75000 for the CS phantom, and 30000 for the physical phantom and clinical patient data.

The objective function of the deblurring sub-problem is  $\Phi_2(x)$  (see Eq. (12)), and this function can be divided between Eq. (14) and Eq. (15). The minimizer for the optimization problem in Eq. (15) corresponds to a Gaussian denoiser with noise level  $\sqrt{\eta/(2\xi)}$ . In addition, the noise level of a reconstructed image would decrease gradually during the whole iteration process, indicating that a different denoiser trained over a different noise level should be used in each iteration. To lighten the burden of parameter selection, deblurring parameters  $\xi$  and  $\eta$  were set so that the ratio  $\varphi = \eta/\xi$  would decrease linearly during reconstruction [34].

#### G. Algorithm

The pseudo code of our algorithm is reported as follows:

---

##### Algorithm 1

---

**Initialize:**  $x_0 = FDK(p)$ ,  $m = 0$ ,  $k = 0$ ,  $N_1 = 30$ ,  $N_2 = 30$   
**While** ( $m < N_1$ )  
  {  
     $k = 0$ ,  $z_0 = x_m$ ;  
    **While** ( $k < N_2$ )  
      {  
         $u_{k+1} \leftarrow \arg \min_{u_k} \|x_m - Gu_k\|_2^2 + \beta \|z_k - u_k\|_2^2$ ; (See Eq.14)  
         $z_{k+1} \leftarrow \arg \min_{z_k} \|z_k - u_{k+1}\|_2^2 + \varphi R(z_k)$ ; (See Eq.15)  
         $k = k + 1$ ;  
      }  
    **If** ( $m = N_1 - 1$ )  
      **break**  
     $x_{m+1} \leftarrow \arg \min_{x_m} (p - Ax_m)^T \Sigma^{-1} (p - Ax_m) + \lambda \|x_m - Gz_{N_2}\|_2^2$  (See Eq.8)  
     $m = m + 1$ ;  
  }  
**Output:**  $z_{N_2}$   
**End**

---

Note:  $n_1$  is the iteration number for the whole algorithm,  $n_2$  is the iteration number inside the deblurring process consisting of the least square problem (Eq. (14)) and the denoise problem (Eq. (15)). We used the Gauss-Seidel update strategy, FFT, and the denoising network to update  $x$ ,  $u$ , and  $z$  respectively.

---

### III. MATERIALS AND EVALUATION

Two simulation phantoms (Compressed Sensing (CS) phantom, a modified Shepp-Logan Phantom), two physical phantoms (CatPhan 600, anthropomorphic head phantom), and data from a clinic patient were used to test our reconstruction algorithm.

### A. Synthetic Phantom

The CS and the 3D Shepp-Logan phantoms were used for simulation. Their projection data covered 360 angles acquired by the SF algorithm [53]. Both phantoms contain  $350 \times 350 \times 16$  pixels, with an individual physical size of  $0.776 \times 0.776 \times 0.776 \text{ mm}^3$ . Different Gaussian noise levels were simulated as different phantoms at different dose levels. The distance from source to detector is 150 cm and the distance from the object axis to detector is 50 cm for both simulated phantoms.

### B. Physical Phantom

Projection data of the CatPhan 600 and anthropomorphic head, two physical phantoms, were acquired using a Varian Acuity system (Varian Medical System, Palo Alto, CA) with 125 kVp voltage. The X-ray tube current was set to 10 mA for low doses and 80mA for high doses. The projection image was of size  $1024 \times 768$ . The number of projections was 634 covering 360 degrees. Sizes were  $350 \times 350 \times 16$  for the CatPhan phantom and  $550 \times 550 \times 32$  for the head phantom. The physical size of each pixel was  $0.776 \times 0.776 \times 0.776 \text{ mm}^3$  and  $0.388 \times 0.388 \times 0.388 \text{ mm}^3$  for the CatPhan phantom and the head phantom, respectively. The distance between source and detector was 150 cm and the distance between axis and detector was 50 cm for both phantoms.

### C. Patient Data

CBCT images were also acquired from a head and neck patient with a Varian Acuity system using a high-dose protocol of 80mA and 25ms, and a low-dose protocol of 20mA and 20ms. The X-ray tube was set to 100 kVp voltage. The reconstruction image was  $500 \times 500 \times 16$ , the size of each pixel was  $0.5 \times 0.5 \times 0.5 \text{ mm}^3$ , and the number of projections was 365 covering over 200 degrees.

### D. Evaluation Criteria

PSNR, improvement signal to noise ratio (ISNR), structural similarity (SSIM), and contrast-to-noise ratio (CNR) were adopted to evaluate the reconstructed images. Since noise-free images were unavailable for the physical phantoms, we chose the FDK reconstruction images of high-dose level as reference. The quality of the reconstructed images depended on the selected parameters. To make the reconstructed images comparable across different methods, their noise levels were adjusted to be similar to one another. For the proposed algorithm, we fixed the pure reconstruction parameter  $\lambda$  and adjusted the noise level by using different deblurring parameters. Noise level was calculated as the average standard deviation of the selected regions. The definitions of PSNR, ISNR, SSIM, and CNR have been described [19].

## IV. EXPERIMENTS

### A. CS Phantom

A representative slice of the noise-free image of the CS Phantom is shown in Fig. 3. Five blue rectangles were used to

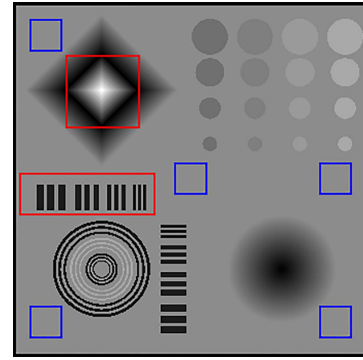


Fig. 3. The original noise-free CS phantom.

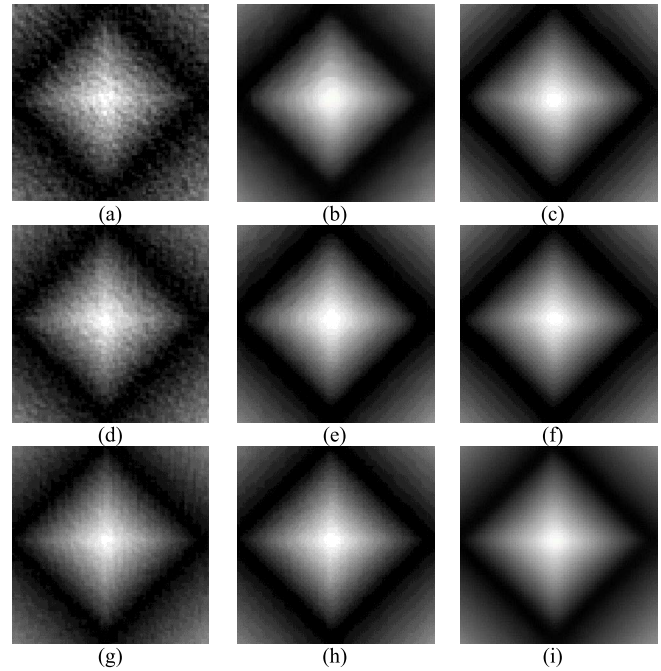


Fig. 4. Zoom-in visual inspections of ROI (rhombus) of images reconstructed by different algorithms with different incident numbers. FDK: (a) high noise level; (d) medium noise level; (g) low noise level; TV: (b) high noise level; (e) medium noise level; (h) low noise level; Our method: (c) high noise level; (f) medium noise level; (i) low noise level.

calculate noise levels. All the images reconstructed by different methods were adjusted to have the same noise level for fair comparison. The red rectangles were chosen as the region of interest (ROI) to reflect the quality of different methods.

The enlarged rhombus ROI of the reconstruction results by FDK, TV, and our proposed method with incident photon number  $5 \times 10^3$  is shown in Fig. 4(a-c). The ROI results with incident photon number  $1 \times 10^4$  are shown in Fig. 4(d-f), and the ROI results with incident photon number  $5 \times 10^4$  are shown in Fig. 4(g-i). As observed in the reconstruction results, the FDK algorithm failed to reduce noise and caused significant artifacts. The TV penalty led to a serious staircase effect and several artificial piecewise constant regions. The images reconstructed using our method were more natural and smooth, indicating the ability of our method to preserve smooth regions in addition to suppressing noise. The codes and projection data of the CS phantom are available at: <https://github.com/HUST-Tan/Deblur-CBCT>.

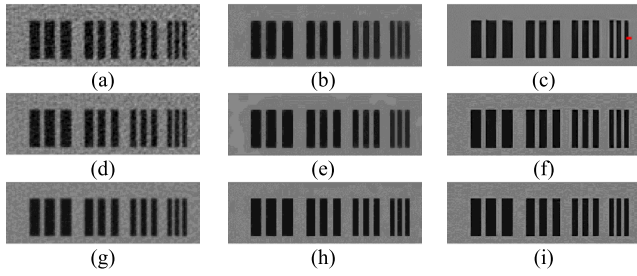


Fig. 5. Zoom-in visual inspections of ROI (Barcode) of images reconstructed by different algorithms with different incident numbers. FDK: (a) high noise level; (d) medium noise level; (g) low noise level; TV: (b) high noise level; (e) medium noise level; (h) low noise level; Our method: (c) high noise level; (f) medium noise level; (i) low noise level.

TABLE II  
PSNR, ISNR, AND SSIM VALUES OF IMAGES BY DIFFERENT RECONSTRUCTION ALGORITHMS FOR THE CS PHANTOM

Incident photon	Penalty	Noise level ( $\times 10^{-5}$ )	PSNR	ISNR	SSIM
$5 \times 10^3$	FDK	120	23.85	/	0.42
	TV	3.51	29.82	5.97	0.96
	Ours	3.46	<b>33.95</b>	<b>10.10</b>	<b>0.97</b>
$1 \times 10^4$	FDK	95.20	25.66	/	0.54
	TV	2.79	31.32	5.66	0.97
	Ours	2.74	<b>35.18</b>	<b>9.52</b>	<b>0.99</b>
$5 \times 10^4$	FDK	52.10	27.93	/	0.77
	TV	1.96	36.97	9.04	<b>0.99</b>
	Ours	1.96	<b>38.27</b>	<b>10.34</b>	<b>0.99</b>

TABLE III  
PSNR, ISNR, AND SSIM VALUES OF ROI (RHOMBUS) BY DIFFERENT RECONSTRUCTION ALGORITHMS FOR THE CS PHANTOM

Incident photon	Penalty	Noise level ( $\times 10^{-5}$ )	PSNR	ISNR	SSIM
$5 \times 10^3$	FDK	120	25.76	/	0.66
	TV	3.51	35.38	9.61	0.95
	Ours	3.46	<b>38.32</b>	<b>12.55</b>	<b>0.98</b>
$1 \times 10^4$	FDK	95.20	28.84	/	0.79
	TV	2.79	36.72	7.87	0.96
	Ours	2.74	<b>39.36</b>	<b>10.52</b>	<b>0.98</b>
$5 \times 10^4$	FDK	52.10	33.93	/	0.93
	TV	1.96	39.98	6.05	<b>0.98</b>
	Ours	1.96	<b>40.87</b>	<b>6.94</b>	<b>0.98</b>

To evaluate our method in the edge area, the strip area (Barcode) was selected (Fig. 5). For all dose levels, visual inspection showed that both the TV penalty and the proposed method preserved the edges and suppressed the noise well. In contrast, the FDK method generated noise. A FWHM check along the red line in Fig. 5(c) indicated that our method had better spatial resolution (0.59 vs. 1.55) than the TV penalty, at low-dose levels.

To quantify the reconstruction image with different methods, PSNR, ISNR, and SSIM of the whole slice and also the two selected ROIs were listed in Table II, III, and IV, respectively. The valuation indexes of our proposed method were found to be the highest at all dose levels.

TABLE IV  
PSNR, ISNR, AND SSIM VALUES OF ROI (BARCODE) BY DIFFERENT RECONSTRUCTION ALGORITHMS FOR THE CS PHANTOM

Incident photon	Penalty	Noise level ( $\times 10^{-5}$ )	PSNR	ISNR	SSIM
$5 \times 10^3$	FDK	120	17.61	/	0.76
	TV	3.51	21.91	5.97	0.96
	Ours	3.46	<b>28.89</b>	<b>11.29</b>	<b>0.99</b>
$1 \times 10^4$	FDK	95.20	19.22	/	0.82
	TV	2.79	24.39	5.17	0.97
	Ours	2.74	<b>30.78</b>	<b>11.56</b>	<b>0.99</b>
$5 \times 10^4$	FDK	52.10	20.82	/	0.90
	TV	1.96	32.65	11.83	<b>1.00</b>
	Ours	1.96	<b>33.26</b>	<b>12.43</b>	<b>1.00</b>

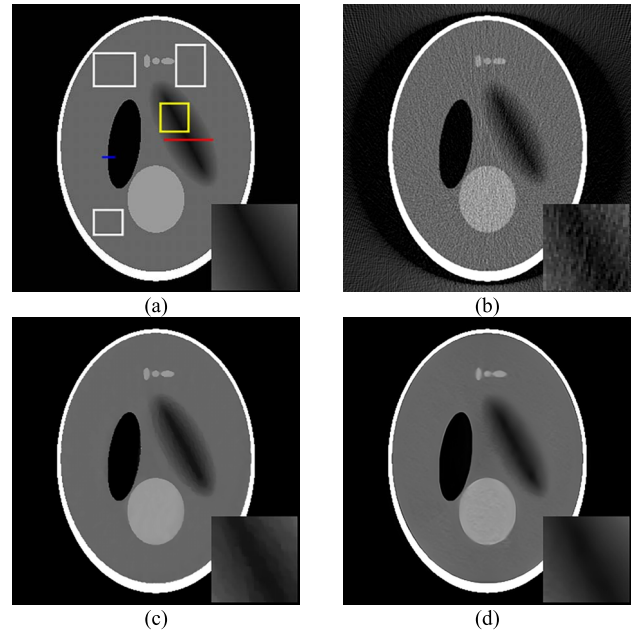


Fig. 6. A representative slice of the 3D Shepp-Logan Phantom reconstructed by different methods: (a) Original image; (b) FDK; (c) TV; (d) Ours.

### B. Modified 3D Shepp-Logan Phantom

The 3D Shepp-Logan phantom was modified by replacing the constant ellipsoid with one containing linearly changed intensity. A representative slice of the modified 3D Shepp-Logan phantom obtained by different methods is displayed in Fig. 6. The original image is illustrated in Fig. 6(a), and the image reconstructed by different methods is illustrated in Fig. 6(b-d). The white rectangles were selected as background to calculate noise levels. The yellow rectangle indicates the ROI, which is displayed in the lower right corner in Fig. 6(a-d). Our proposed method can successfully avoid the staircase effect and preserve the smooth region. In contrast, the FDK reconstruction image contains too much noise, and the TV penalty shows an obvious staircase effect.

The change of image intensity along the red line in Fig. 6(a) is illustrated in Fig. 7. These plots indicate that the TV penalty produces piecewise constant artifacts, while the proposed method exhibits better reconstruction quality.



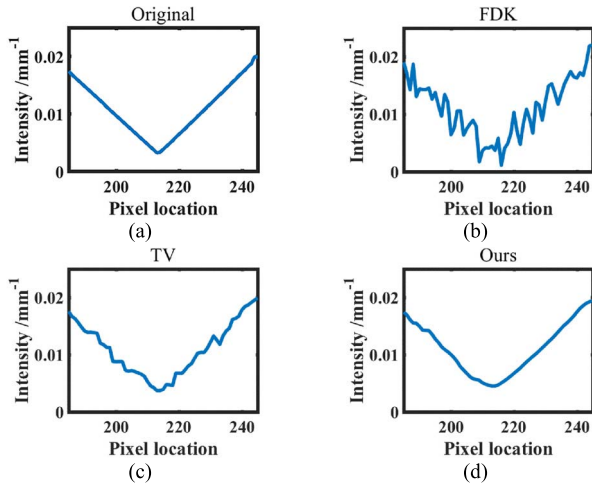


Fig. 7. Profiles through the ellipsoid object (the red line in Fig. 6(a)): (a) Original image; (b) FDK; (c) TV; (d) Ours.

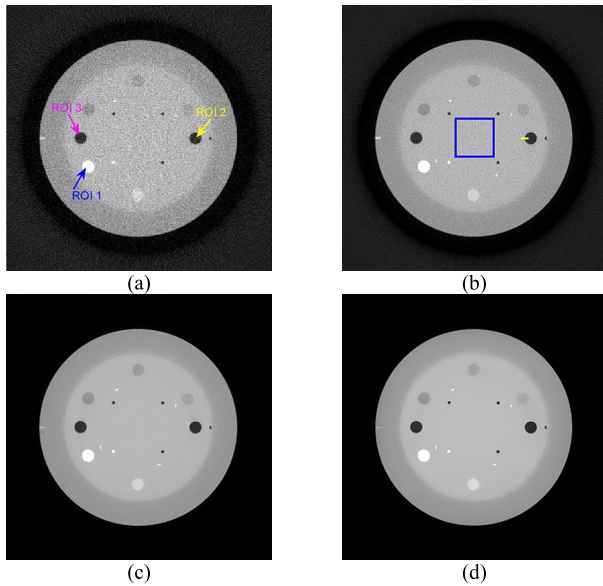


Fig. 8. Representative slices from the reconstructed CatPhan 600 images. (a) FDK with high-dose protocol (80mA/10ms); (b) FDK with Low-dose protocol (10mA/10ms); (c) TV (low-dose); (d) Ours (low-dose).

As for image resolution, we selected the blue line in Fig. 6(a) to calculate the FWHM. The FWHM value of the TV penalty was 1.8597, and that of our method was 0.5821, indicating a higher image resolution for our method.

### C. CatPhan 600 Phantom

A representative slice of the reconstructed CatPhan600 phantom images generated with different methods are shown in Fig. 8. The blue rectangle area was used to calculate noise level. Since we cannot obtain noise-free images, we selected three ROIs, to measure and quantitatively compare CNR between different reconstruction methods (indicated by the arrows in Fig. 8(a)).

The evaluation indexes are listed in Table V. Our method performed better than the TV penalty in denoising.

TABLE V  
CNRs OF DIFFERENT ROIs OF THE CATPHAN 600 PHANTOM

	Noise( $\times 10^{-4}$ )	ROI1	ROI2	ROI3
FDK_10mA	24	3.55	5.37	5.22
FDK_80mA	9.15	9.64	12.76	12.63
TV	1.16	70.40	87.49	74.95
Ours	1.15	<b>82.58</b>	<b>113.41</b>	<b>94.47</b>

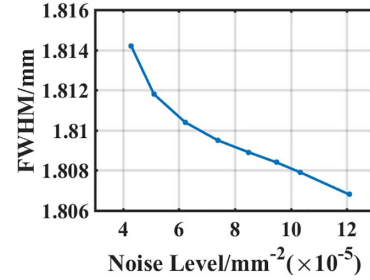


Fig. 9. FWHM Values-Noise Level curves.

Since CNR depends highly on the background noise level, its values present limitations in evaluating the quality of the reconstructed images. To emphasize the importance of our method in deblurring, we drew a yellow line to calculate FWHM (Fig. 8(b)). The FWHM of the TV penalty is 1.9477 and that of our method is 1.7932, indicating higher image resolution for our method. To check if the proposed algorithm is universal and adaptive, we calculated FWHM values for images reconstructed with different noise levels (Fig. 9). The image resolution of our method changed only slightly (FWHM values decreased by 0.4% from 1.8143 to 1.8068) when the noise level was increased by 140% from  $4 \times 10^5$  to  $12 \times 10^5$  (Fig. 9). Table V and Fig. 9 indicated that our method can efficiently suppress both noise and blurring effect, providing high image resolution (Table V and Fig. 9).

### D. Anthropomorphic Head Phantom

A representative slice of the results obtained using different reconstruction methods for the head phantom are illustrated in Fig. 10. The red rectangle area was selected as the ROI, and the white rectangle area was used to measure noise levels (Fig. 10(b)). The reconstructions of the proposed algorithm and TV were within the same noise level. The high dose FDK reconstruction image was chosen as reference because the original noise-free image was unavailable. We used the SSIM to evaluate the similarity degree of the local structures between the reconstruction images and the high dose FDK image. Our method had a whiter SSIM map of the anthropomorphic head phantom (Fig. 11), indicating a better ability to preserve structures. The mean SSIM (MSSIM) value in the selected ROI is listed in Table VI, revealing that our method had the best structure-preserving ability.

### E. Patient Results

The results of the FDK reconstruction on patient data acquired with different dose levels are shown in Fig. 12 (a-b).



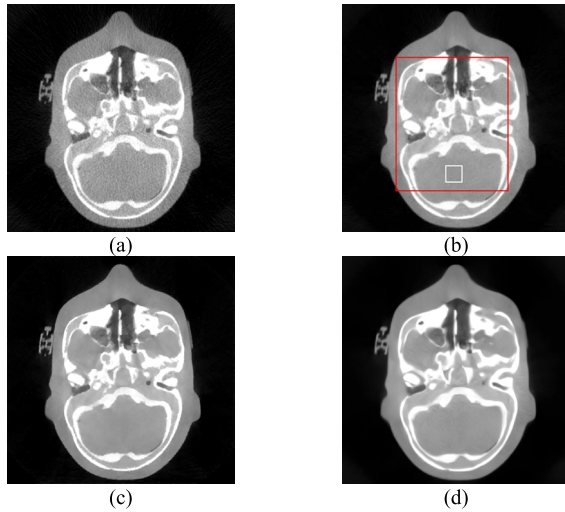


Fig. 10. A representative slices of the reconstructed anthropomorphic head phantom image using different penalties: (a) FDK low-dose protocol (10mA/10ms); (b) FDK high-dose protocol (80mA/10ms); (c) TV (low-dose); (d) Ours (low-dose).

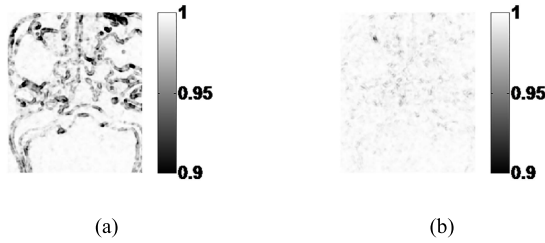


Fig. 11. SSIM map of the head phantom using different methods. (a) TV (SSIM = 0.9861), (b) Ours (SSIM = 0.9964).

TABLE VI  
MSSIM OF THE RED ROI SHOWN IN Fig. 10 (B)

Penalty	Noise level( $\times 10^{-4}$ )	SSIM
FDK_10mA	17	0.74
FDK_80mA	8.54	N/A
TV	4.88	0.97
Ours	4.87	1.00

The low dose results using the TV penalty and the proposed method are shown in Fig. 12 (c-d). The white rectangle on the background was selected to measure noise levels. The noise levels for the

FDK low-dose protocol, FDK high-dose protocol, SIR with the TV penalty and our algorithm were found to be  $29 \times 10^4$ ,  $14 \times 10^4$ ,  $6.52 \times 10^4$ , and  $6.75 \times 10^4$ , respectively. The red line in the image was selected to calculate the FWHM value and test the deblurring effect. The FWHM of our method is 1.7145, while that of the TV penalty is 3.0090 with. These results suggest that our method had higher spatial resolution than TV. The red rectangle is displayed in the lower right corner to illustrate the performance of different reconstruction methods. Visual results obtained by our proposed method were more

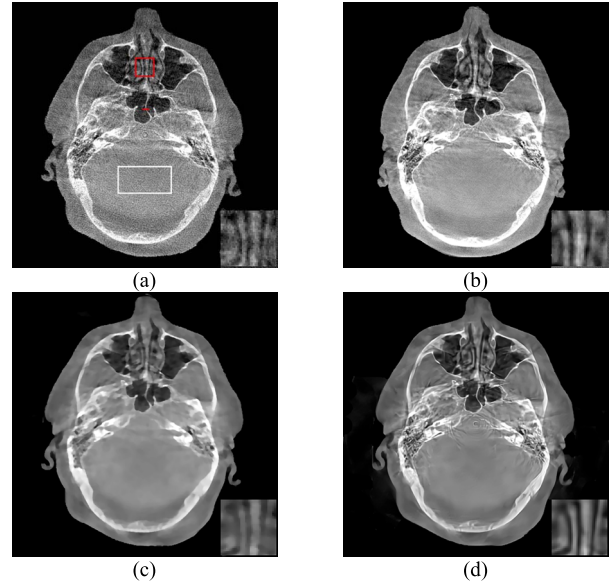


Fig. 12. A representative slice of the reconstructed patient data using different penalties: (a) FDK low-dose protocol (20mA/20ms, nose level:  $29 \times 10^4$ ); (b) FDK high-dose protocol (80mA/25ms, nose level:  $14 \times 10^4$ ); (c) TV (nose level:  $6.52 \times 10^4$ ) and (d) Ours (nose level:  $6.75 \times 10^4$ ), both using the low-dose protocol.

similar to those of the reference high dose FDK reconstruction, than to those of the traditional TV penalty.

## V. DISCUSSION

In statistical iterative reconstruction, the **regularization term** represents prior information on the target image, and plays an important role for reconstruction quality. Many regularization terms such as TV and Hessian have been proposed because of their good performance. However, because the **manually designed regularization term** is based primarily on the researchers' knowledge about the problem, effective regularization is often hard to design. We proposed using the network to learn the "potential regularization term" through data. This shifts the problem of designing the regularization term to the problem of training the neural network.

**The major challenge of applying CNN to SIR is where and what role it should play in the iteration process.** Considering that the amount of normal-dose CBCT images is low and that these images are no way to be noiseless in practice, we used transfer learning [33] to train the network on natural images. Instead of training a network end-to-end [43], we arranged for the neural network to play the role of a regularization term, which can be embedded into different inverse problems. To adapt to the SIR process, which might have different blurring kernels and noise levels in every iterative phase, we used an iterative deblurring method consisting of a CNN denoiser and a quadratic regularized least-square problem. The iterative deblurring method possibly can also deal with different kinds of noise that benefit CBCT reconstruction problems in which the noise type may be unknown. One of the major advantages of the proposed reconstruction method is integrating CNN into the iterative reconstruction, adding extra

flexibility and extendibility compared to other CNN-based CBCT reconstruction methods [43], [44].

Eq. (15) corresponds to a Gaussian denoising problem, and the regularization term in Eq. (15) plays a role in enforcing the solution to be more similar to a noise-free CBCT image. Traditionally, the key is to design a proper regularization term  $R(u)$  for the denoising problem. We used a neural network to solve the denoising problem. The corresponding form of a neural network solution can be represented as  $z = F(u)$  (Please see Eq. (15)). The neural network method solves the denoising problem by learning  $F$ . The learnt denoiser-using network can be viewed as learning a “potential regularization term” through data. Ulyanov *et al.* [60] found that the network is sufficient to capture rich low-level image statistics. The regularization term in Eq. (15) was exactly the same as that in Eq. (4). Thanks to the variable splitting technique, the optimization of Eq. (15) (or denoising) using the neural network is equivalent to learning the regularization term for the entire algorithm (Eq. (4)) via the neural network.

Our findings indicate that our network-based SIR can successfully eliminate the staircase effect and preserve both edges and image regions with smooth intensity transition. The traditional SIR with a specially designed penalty term often needs to adjust the parameters to balance noise levels and image resolution, which is typically time-consuming. In contrast, our method can potentially provide reconstruction images with low noise level, while maintaining high image resolution.

Although our network-based SIR showed good performance, several aspects need to be improved. First, the deblurring process might not be thorough due to the rough PSF estimation (using a linear decreasing way), which may reduce image resolution. This limitation may be addressed by designing a dynamic PSF estimation process. Second, as shown in the CS experiment, different regions of the images exhibited different noise levels. Also, the network did not work well for local regions where the background was close to the foreground, indicating the need to design a more “intelligent” network. Third, we trained the network using 25 different noise levels, which are not elegant enough, indicating the need to design a more common denoiser. Finally, as shown in our CatPhan600 experiments, the small bright structure on the left of the phantom (Fig. 8 (a)) became dim when our algorithm was used (Fig. 8 (d)). Interestingly, reconstruction using the TV penalty underwent a similar phenomenon (Fig. 8(c)). In our algorithm, parameter  $\lambda$  in Eq. (6) still played a similar role to the regularization parameter in a standard inverse problem formulation. In this formulation, a large regularization parameter will generally make the solution smoother. The objective function in Eq. (6) used a quadratic “penalty” that was introduced by the variable splitting technique. A quadratic penalty is known to over-smooth the solution in a standard inverse problem. In this sense, the quadratic “penalty” might be an inherent shortcoming of the proposed algorithm. One way to improve the current version of our proposed algorithm is to use other variable splitting techniques that lead to other “penalties” (e.g. non-quadratic) that better preserve small and faint structures.

## VI. CONCLUSION

We proposed a new network-based iterative CBCT reconstruction method that addresses issues such as data deficiency and unknown noise level to train the network. Our proposed SIR method solves the CBCT reconstruction problem innovatively. The data-driven approach that learns the “potential regularization term” overcomes the challenge of designing penalty terms in traditional SIR methods. Our studies conducted in simulated and physical phantoms, and also on data from a clinic patient showed that our method can eliminate the staircase effect and preserve both edges and regions with smooth intensity transition.

## VII. APPENDIX

### A. Solve Quadratic Regularized Least Squares

We defined several symbols as follows:

$\otimes$ : convolution,  $\circ$ : dot product,  $k$ : frequency domain coordinates,  $\bar{X}$ : conjugate transpose.

Rewrite Eq. (14) as follow:

$$\Phi_1 = \|x - G \otimes u\|_2^2 + \xi \|z - u\|_2^2. \quad (\text{A.1})$$

Introduce a new matrix  $M$ : the same size as  $G$  and every element in  $M$  is 0 excluding the center element which is 1. Eq. (A-1) now became:

$$\Phi_1 = \|x - G \otimes u\|_2^2 + \xi \|M \otimes u - z\|_2^2. \quad (\text{A.2})$$

Letting  $\partial \Phi_1 / \partial u = 0$  and applying Fourier transform, we obtained

$$u(k) = \frac{\bar{G}(k) \circ x + \xi \bar{M}(k) \circ z}{|G(k)|^2 + \xi |M(k)|^2}. \quad (\text{A.3})$$

Using IFFT, we can obtain the solution of  $u$ .

## ACKNOWLEDGEMENT

The authors would like to thank Dr. Damiana Chiavolini for editing the paper.

## REFERENCES

- [1] A. B. de González *et al.*, “Projected cancer risks from computed tomographic scans performed in the United States in 2007,” *Arch. Internal Med.*, vol. 169, no. 22, pp. 2071–2077, 2009.
- [2] N. Wen *et al.*, “Dose delivered from Varian’s CBCT to patients receiving IMRT for prostate cancer,” *Phys. Med. Biol.*, vol. 52, no. 8, p. 2267, 2007.
- [3] J. Hsieh, “Adaptive streak artifact reduction in computed tomography resulting from excessive X-ray photon noise,” *Med. Phys.*, vol. 25, no. 11, pp. 2139–2147, 1998.
- [4] M. Kachelrieß, O. Watzke, and W. A. Kalender, “Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT,” *Med. Phys.*, vol. 28, no. 4, pp. 475–490, 2001.
- [5] J. Wang *et al.*, “An experimental study on the noise properties of X-ray CT sinogram data in Radon space,” *Phys. Med. Biol.*, vol. 53, no. 12, p. 3327, 2008.
- [6] H.-M. Huang and I.-T. Hsiao, “Accelerating an ordered-subset low-dose X-ray cone beam computed tomography image reconstruction with a power factor and total variation minimization,” *PLoS ONE*, vol. 11, no. 4, p. e0153421, 2016.
- [7] X. Jia *et al.*, “GPU-based fast low-dose cone beam CT reconstruction via total variation,” *J. X-Ray Sci. Technol.*, vol. 19, no. 2, pp. 139–154, 2011.
- [8] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.

- [9] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.*, vol. 53, no. 17, p. 4777, Sep. 2008.
- [10] Y. Zhang, J. N. Tehrani, and J. Wang, "A biomechanical modeling guided CBCT estimation technique," *IEEE Trans. Med. Imag.*, vol. 36, no. 2, pp. 641–652, Feb. 2017.
- [11] X. Jia, Y. Lou, R. Li, W. Y. Song, and S. B. Jiang, "GPU-based fast cone beam CT reconstruction from undersampled and noisy projection data via total variation," *Med. Phys.*, vol. 37, no. 4, pp. 1757–1760, 2010.
- [12] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *J. X-Ray Sci. Technol.*, vol. 14, no. 2, pp. 119–139, 2006.
- [13] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 492–526, 2010.
- [14] Y.-L. You and M. Kaveh, "Fourth-order partial differential equations for noise removal," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1723–1730, Oct. 2000.
- [15] Y. Hu and M. Jacob, "Higher degree total variation (HDTV) regularization for image recovery," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2559–2571, May 2012.
- [16] S. Lefkimmatis, J. P. Ward, and M. Unser, "Hessian schatten-norm regularization for linear inverse problems," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1873–1888, May 2013.
- [17] J. Liu, T.-Z. Huang, I. W. Selesnick, X.-G. Lv, and P.-Y. Chen, "Image restoration using total variation with overlapping group sparsity," *Inf. Sci.*, vol. 295, pp. 232–246, Feb. 2015.
- [18] Q. Shi, N. Sun, T. Sun, J. Wang, and S. Tan, "Structure-adaptive CBCT reconstruction using weighted total variation and Hessian penalties," *Biomed. Opt. Exp.*, vol. 7, no. 9, pp. 3299–3322, 2016.
- [19] T. Sun, N. Sun, J. Wang, and S. Tan, "Iterative CBCT reconstruction using Hessian penalty," *Phys. Med. Biol.*, vol. 60, no. 5, p. 1965, 2015.
- [20] X. Jia, B. Dong, Y. Lou, and S. B. Jiang, "GPU-based iterative cone-beam CT reconstruction using tight frame regularization," *Phys. Med. Biol.*, vol. 56, no. 13, p. 3787, 2011.
- [21] 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. Imaging*, vol. 31, no. 9, pp. 1682–1697, Sep. 2012.
- [22] Y. Lu, J. Zhao, and G. Wang, "Few-view image reconstruction with dual dictionaries," *Phys. Med. Biol.*, vol. 57, no. 1, p. 173, 2011.
- [23] K. Choi, J. Wang, L. Zhu, T.-S. Suh, S. Boyd, and L. Xing, "Compressed sensing based cone-beam computed tomography reconstruction with a first-order method," *Med. Phys.*, vol. 37, no. 9, pp. 5113–5125, 2010.
- [24] S. Ravishanker, B. E. Moore, R. R. Nadakuditi, and J. A. Fessler, "Low-rank and adaptive sparse signal (LASSI) models for highly accelerated dynamic imaging," *IEEE Trans. Med. Imag.*, vol. 36, no. 5, pp. 1116–1128, May 2017.
- [25] K. Kim *et al.*, "Sparse-view spectral CT reconstruction using spectral patch-based low-rank penalty," *IEEE Trans. Med. Imag.*, vol. 34, no. 3, pp. 748–760, Mar. 2016.
- [26] H. Zhang *et al.*, "Iterative reconstruction for X-ray computed tomography using prior-image induced nonlocal regularization," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 9, pp. 2367–2378, Sep. 2014.
- [27] S. Niu, G. Yu, J. Ma, and J. Wang, "Nonlocal low-rank and sparse matrix decomposition for spectral CT reconstruction," *Inverse Problems*, vol. 34, no. 2, p. 024003, 2017.
- [28] J. Hao, L. Zhang, L. Li, and K. Kang, "An improved non-local means regularized iterative reconstruction method for low-dose dental CBCT," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. (NSS/MIC)*, Oct./Nov. 2012, pp. 3422–3425.
- [29] Y. Chen *et al.*, "Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior," *Comput. Med. Imag. Graph.*, vol. 33, no. 7, pp. 495–500, 2009.
- [30] P. F. Feruglio, C. Vinegoni, J. Gros, A. Sbarbati, and R. Weissleder, "Block matching 3D random noise filtering for absorption optical projection tomography," *Phys. Med. Biol.*, vol. 55, no. 18, p. 5401, 2010.
- [31] Y. Choi, J.-Y. Koo, and N.-Y. Lee, "Image reconstruction using the wavelet transform for positron emission tomography," *IEEE Trans. Med. Imag.*, vol. 20, no. 11, pp. 1188–1193, Nov. 2001.
- [32] J.-F. Cai, X. Jia, H. Gao, S. B. Jiang, Z. Shen, and H. Zhao, "Cine cone beam CT reconstruction using low-rank matrix factorization: Algorithm and a proof-of-principle study," *IEEE Trans. Med. Imag.*, vol. 33, no. 8, pp. 1581–1591, Aug. 2014.
- [33] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
- [34] K. Zhang, W. Zuo, S. Gu, and L. Zhang. (2017). "Learning deep CNN denoiser prior for image restoration." [Online]. Available: <https://arxiv.org/abs/1704.03264>
- [35] J. Zhang, J. Pan, W.-S. Lai, R. Lau, and M.-H. Yang. (2016). "Learning fully convolutional networks for iterative non-blind deconvolution." [Online]. Available: <https://arxiv.org/abs/1611.06495>
- [36] S. Lefkimmatis. (2016). "Non-local color image denoising with convolutional neural networks." [Online]. Available: <https://arxiv.org/abs/1611.06757>
- [37] A. Chakrabarti, "A neural approach to blind motion deblurring," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 221–235.
- [38] P. Wieschollek, B. Schölkopf, H. P. Lensch, and M. Hirsch, "End-to-end learning for image burst deblurring," in *Proc. Asian Conf. Comput. Vis.*, 2016, pp. 35–51.
- [39] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang. (2017). "Deep Laplacian pyramid networks for fast and accurate super-resolution." [Online]. Available: <https://arxiv.org/abs/1704.03915>
- [40] C. Ledig *et al.* (2016). "Photo-realistic single image super-resolution using a generative adversarial network." [Online]. Available: <https://arxiv.org/abs/1609.04802>
- [41] G. Wang, "A perspective on deep imaging," *IEEE Access*, vol. 4, pp. 8914–8924, 2016.
- [42] E. Kang, J. Min, and J. C. Ye. (2016). "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction." [Online]. Available: <https://arxiv.org/abs/1610.09736>
- [43] H. Chen *et al.*, "Low-dose CT via convolutional neural network," *Biomed. Opt. Exp.*, vol. 8, no. 2, pp. 679–694, 2017.
- [44] Q. Yang, P. Yan, M. K. Kalra, and G. Wang. (2017). "CT image denoising with perceptive deep neural networks." [Online]. Available: <https://arxiv.org/abs/1702.07019>
- [45] L. A. Gatys, A. S. Ecker, and M. Bethge. (2015). "A neural algorithm of artistic style." [Online]. Available: <https://arxiv.org/abs/1508.06576>
- [46] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 694–711.
- [47] M. Arjovsky, S. Chintala, and L. Bottou. (2017). "Wasserstein gan." [Online]. Available: <https://arxiv.org/abs/1701.07875>
- [48] Q. Yang *et al.* (2017). "Low dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss." [Online]. Available: <https://arxiv.org/abs/1708.00961>
- [49] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Trans. Signal Process.*, vol. 4, no. 7, pp. 932–946, Jul. 1995.
- [50] T. Li *et al.*, "Nonlinear sinogram smoothing for low-dose X-ray CT," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 5, pp. 2505–2513, Oct. 2004.
- [51] J. Wang, T. Li, and L. Xing, "Iterative image reconstruction for CBCT using edge-preserving prior," *Med. Phys.*, vol. 36, no. 1, pp. 252–260, 2009.
- [52] S. Hashemi *et al.*, "Simultaneous deblurring and iterative reconstruction of CBCT for image guided brain radiosurgery," *Phys. Med. Biol.*, vol. 62, no. 7, p. 2521, 2017.
- [53] Y. Long, J. A. Fessler, and J. M. Balter, "3D forward and back-projection for X-ray CT using separable footprints," *IEEE Trans. Med. Imag.*, vol. 29, no. 11, pp. 1839–1850, Nov. 2010.
- [54] D. Ulyanov, A. Vedaldi, and V. Lempitsky. (2016). "Instance normalization: The missing ingredient for fast stylization." [Online]. Available: <https://arxiv.org/abs/1607.08022>
- [55] O. Russakovsky *et al.*, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
- [56] T.-Y. Lin *et al.*, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 740–755.
- [57] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
- [58] D. P. Kingma and J. Ba (2014). "Adam: A method for stochastic optimization." [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [59] Z. Lin, M. Chen, and Y. Ma. (2010). "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices." [Online]. Available: <https://arxiv.org/abs/1009.5055>
- [60] D. Ulyanov, A. Vedaldi, and V. Lempitsky. (2017). "Deep image prior." [Online]. Available: <https://arxiv.org/abs/1711.10925>