Sparse-view CBCT reconstruction via weighted Schatten p-norm minimization

Congcong Xu, Bo Yang, Fupei Guo, Wenfeng Zheng, Philippe Poignet

▶ **To cite this version:**

HAL Id: lirmm-02997654

https://hal-lirmm.ccsd.cnrs.fr/lirmm-02997654

Submitted on 10 Nov 2020

Sparse-view CBCT reconstruction via weighted Schatten *p*-norm minimization

CONGCONG XU,[1] BO YANG,[1,*] 🆔 FUPEI GUO,[1] WENFENG ZHENG,[1] 🆔 AND PHILIPPE POIGNET[2]

[1]*School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu, China*

[2]*LIRMM, University of Montpellier, CNRS, Montpellier, France*

*boyang@uestc.edu.cn

Abstract: A novel iterative algorithm is proposed for sparse-view cone beam computed tomography (CBCT) reconstruction based on the weighted Schatten *p*-norm minimization (WSNM). By using the half quadratic splitting, the sparse-view CBCT reconstruction task is decomposed into two sub-problems that can be solved through alternating iteration: simple reconstruction and image denoising. The WSNM that fits well with the low-rank hypothesis of CBCT data is introduced to improve the denoising sub-problem as a regularization term. The experimental results based on the digital brain phantom and clinical CT data indicated the advantages of the proposed algorithm in both structural information preservation and artifacts suppression, which performs better than the classical algorithms in quantitative and qualitative evaluations.

1. Introduction

Cone beam computed tomography (CBCT) is an important medical imaging technology widely used in clinical diagnosis and treatment. During a CBCT scanning, the cone-beam X-rays are emitted at several equal intervals angles when the X-ray source moves in a circular orbit around the detected object. At the other end of the device, a flat panel detector receives the X-rays that pass through the detected object and a series of projection data (observations) are captured, from which a 3D voxel model can be reconstructed for clinical use. The CBCT has the advantages of fast scanning speed and high spatial resolution. However, since the X-rays are harmful to the human body, which may induce cancer or genetic diseases [1], both doctors and patients are looking for techniques to reduce the radiation dose of CBCT. There are two main approaches to reduce the dose, i.e., decreasing mAs (the product of tube current and exposure time) or reduce projection views [2]. The former introduces noise in the projection data directly, while the latter leads to artifacts in the reconstructed model due to insufficient observations [3]. This paper focuses on the CBCT reconstruction of sparse projection views.

Analytical methods such as the classical FDK [4] and its improved versions [5] have been developed early for full-view CBCT reconstruction. However, sparse views lead to serious artifacts in the analytically reconstructed results due to insufficient projection data. To overcome the shortcoming of the analytical methods, iterative algorithms which treat the CBCT reconstruction as an ill-posed linear problem have been investigated, such as the ART [6], SART [7], OS-SART [8], and EM [9]. Although the reconstruction performance has been improved, it is still very challenging for very sparse views.

The statistical iterative reconstruction (SIR) framework has been studied to further improve the sparse-view reconstruction. Generally, the objective function of SIR is composed of two items. The first item is a data fidelity item, which drives the reconstruction result consistent with the limited projection data as possible. The second item is a regularization (or prior) item, which limits the solution to the prior knowledge of statistics. The regularization term

can be built on various prior models, among which the total variation (TV) [10–12] is the most commonly used. Based on the TV minimization, several CBCT reconstruction algorithms have been proposed [13,14]. However, these TV-based methods suffer from the staircase effect due to the piecewise assumption being used. To overcome this problem, some improved TV terms have been introduced, such as the ASD-POCS algorithm [15], total *p*-variation [16], and prior contour based total variation [17]. Besides the TV, other regularization techniques are also investigated for sparse-view CBCT reconstruction, such as the dictionary learning [18], compressive sensing [19], and Hessian Schatten penalties [20]. In addition, deep learning techniques are also used for the CBCT reconstruction, and good results have been reported [21,22].

Recently, in the field of image and signal denoising, the low-rank matrix approximation (LRMA) has attracted wide attention. In [23] and [24], the nuclear norm minimization (NNM) and the Schatten *p*-norm ($0<p \leq 1$) minimization (SNM) are used to solve the LRMA respectively. As a convex relaxation of rank minimization, the NNM is a special case of the SNM when $p = 1$ and easier to solve, while the SNM performs better for accurate data recovery [25]. However, both NNM and SNM treat all singular values equally and shrink them with the same threshold, which causes poor edge/texture preservation, as larger singular values normally carry more structural information. To overcome this shortcoming, the weighted versions, i.e., WNNM [26] and WSNM [27] are proposed, which shrink the larger singular values less during rank minimization. As reported in [27], compared with the WNNM, the WSNM is a better approximation to the LRMA problem, which can be transformed into a series of independent non-convex l_p-norm minimization problem and solved efficiently with a generalized soft-thresholding (GST) algorithm [28].

In this paper, we propose a novel iterative algorithm for sparse-view cone beam computed tomography (CBCT) reconstruction. The nonlocal self-similarity within the medical CBCT data is utilized to improve the reconstruction. Inspired by [27], the WSNM is introduced to quantify this self-similarity and regularize the reconstructed voxels. The main contributions of this work are as follow: 1) the sparse-view CBCT reconstruction task defined as an ill-posed inverse problem is decomposed into two sub-problems by using the half quadratic splitting (HQS) technique [29], i.e., least-square reconstruction and low-rank image denoising; 2) the WSNM originally proposed for 2D image denoising is extended for CBCT reconstruction, as a regularization term to impose low-rank constraint on the reconstructed voxels. Experimental results on various datasets show that the proposed algorithm performs excellently in both qualitative and quantitative evaluations. It achieves a good balance between noise suppression and structural information preservation. The proposed method is also better than the WNNM-based method, benefiting from the extended WSNM which can effectively utilize the correlations between the non-local similar blocks of reconstructed data to preserve structural information.

2. Reconstruction model and solution framework

2.1. General reconstruction model

Let $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \ldots & x_N \end{bmatrix}^T$ be the vectorized voxels and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \ldots & y_M \end{bmatrix}^T$ be the vectorized observations of several views. The CBCT imaging can be described by [30]

$$\mathbf{y} = \mathbf{Ax}, \tag{1}$$

where $\mathbf{A}$ is a $M \times N$ system matrix, whose element a_{ij} denotes the contribution of the *j*-th voxel x_j to the *i*-th observation y_i and can be scaled by the length of the intersection of the *i*-th projection ray (that generates y_i) with x_j.

The goal of CBCT reconstruction is to recover the non-negative voxels $\mathbf{x}$ from the observations $\mathbf{y}$ with the known system matrix $\mathbf{A}$. To reduce the radiation dose, sparse projection views are adopted, which results in insufficient observations and thus makes the voxels hard to solve from Eq. (1). Regularization items are normally used to provide additional constraints to voxels, such

that the voxels can be estimated by solving the following optimization problem [31]

$$\arg \min_{\mathbf{x}} ||\mathbf{Ax} - \mathbf{y}||_2^2 + \lambda R(\mathbf{x}), \tag{2}$$

where the first item is a data fidelity term to ensure that the projections of the estimated voxels are consistent with the observations captured from the CBCT device, $R(\mathbf{x})$ is the regularization function that derives from inherent characteristics of voxel data for providing additional constraints to $\mathbf{x}$, and λ is a factor to balance the weights of the two terms.

2.2. Solution via half quadratic splitting

To solve Eq. (2), the HQS technique is used to decouple the data fidelity term and the regularization term. By introducing an auxiliary variable $\mathbf{f}$ and replacing the original variable $\mathbf{x}$ in the regularization term, Eq. (2) can be written as

$$\arg \min_{\mathbf{x}} ||\mathbf{Ax} - \mathbf{y}||_2^2 + \lambda R(\mathbf{f}), \quad s.t. \ \mathbf{f} = \mathbf{x}. \tag{3}$$

By using the penalty function method, Eq. (3) can be rewritten as an unconstrained optimization problem of two variables $\mathbf{x}$ and $\mathbf{f}$

$$\arg \min_{\mathbf{x},\mathbf{f}} ||\mathbf{Ax} - \mathbf{y}||_2^2 + \beta||\mathbf{f} - \mathbf{x}||_2^2 + \lambda R(\mathbf{f}), \tag{4}$$

with a large penalty factor β. By fixing one variable and solving another alternatively, the CBCT reconstruction problem defined in Eq. (2) is finally transformed into the alternating iterative optimization of two simple sub-problems

$$\mathbf{f}_{i+1} = \arg \min_{\mathbf{f}} \beta||\mathbf{f} - \mathbf{x}_i||_2^2 + \lambda R(\mathbf{f}), \tag{5}$$

$$\mathbf{x}_{i+1} = \arg \min_{\mathbf{x}} ||\mathbf{Ax} - \mathbf{y}||_2^2 + \beta||\mathbf{f}_{i+1} - \mathbf{x}||_2^2, \tag{6}$$

where β increases with the number of iterations.

The first sub-problem associated with the regularization term can be equivalent to a denoising task under the assumption of additive noise, i.e., restoring the true data $\mathbf{f}$ from the noisy data $\mathbf{x}_i$ of a noise variance proportional to λ/β with a denoiser prior $R(\mathbf{f})$. In this work, the weighted Schatten p-norm of similar voxel blocks is used as the prior to improve the denoising task, and the solution of Eq. (5) based on the WSNM is given in Section 3.

The second sub-problem associated with the data fidelity term is a linear least-squares problem in essence. It can be written as

$$\mathbf{x}_{i+1} = \arg \min_{\mathbf{x}} ||\mathbf{A}'\mathbf{x} - \mathbf{y}'||_2^2, \tag{7}$$

and solved directly by $\mathbf{x}_{i+1} = (\mathbf{A}'^{\mathrm{T}}\mathbf{A}')^{-1}\mathbf{A}'^{\mathrm{T}}\mathbf{y}'$ with the expanded system matrix and observation vector

$$\mathbf{A}' = \begin{pmatrix} \mathbf{A} \\ \sqrt{\beta}\mathbf{I} \end{pmatrix}, \quad \mathbf{y}' = \begin{pmatrix} \mathbf{y} \\ \sqrt{\beta}\mathbf{f}_{i+1} \end{pmatrix}. \tag{8}$$

The initial voxels $\mathbf{x}_0$ can be reconstructed from the observations $\mathbf{y}$ by using the SART algorithm, which continuously reduces the error between the estimated projections and the real observations

by looping over all views until a convergence criterion or an upper limit of the iteration is met. The iteration formula of the SART is given by

$$x_j^{(k+1)} = x_j^{(k)} + \alpha \frac{\sum\limits_{i \in I_\varphi} \mu_i a_{ij}}{\sum\limits_{i \in I_\varphi} a_{ij}}, \mu_i = \frac{y_i - \sum\limits_{j=1}^{N} a_{ij} x_j^{(k)}}{\sum\limits_{j=1}^{N} a_{ij}}, \tag{9}$$

where $x_j^{(k)}$ represents the j-th voxel in the k-th iteration, I_φ represents the set of the projection rays at the view φ, α is the step size parameter that controls the degree of correction.

The proposed sparse-view CBCT reconstruction scheme is summarized in Algorithm 1.

Algorithm 1 Sparse-view CBCT reconstruction

Inputs: System matrix **A**, projection data **y**

1: Calculate initial reconstruction $\mathbf{x}_0$ by using the SART algorithm (see Eq. (9)); $i = 0$

2: Repeat

3:      solving f_{i+1} from Eq. (5) (**Algorithm 2** in Section 3)

4:      calculating x_{i+1} by using Eq. (7)

5:      $i \leftarrow i + 1$

6: Until stopping criterion is satisfied

Output: Reconstruction result x_i

---

3. Solving denoising sub-problem

Data and signals derived from real world normally contain structural information that has significant associations between their elements. Information redundancy caused by the correlations can be used to regularize the reconstructed voxels from insufficient observations. To achieve this, the noisy voxels reconstructed from Eq. (6) are divided into non-overlapping reference blocks of the same size. For each reference block, similar blocks that may overlap are grouped via a sliding-window matching as done in [32]. Each group represents a specific type of structure and forms a 2D matrix by reshaping every block into a column vector, which should be low rank due to the similarity. By using the LRMA to each grouped matrix, the denoising sub-problem defined in Eq. (5) can be solved.

As discussed in Introduction, a state-of-the-art LRMA technique, the WSNM, is employed to solve Eq. (5) in this work. The original WSNM algorithm is designed for 2D image denoising. Although the voxels can be processed with the 2D WSNM slice by slice, the 3D structure information between slices cannot be used for reconstruction. Therefore, we extend the WSNM based on the 3D block grouping.

Since the processing steps for each block group are exactly the same, without loss of generality, we use **X** to denote a 2D matrix grouped from the noisy voxels **x**. Each 2D matrix is denoised independently with the proposed WSNM algorithm, and a complete 3D voxel model can be finally obtained by reassembling the denoised blocks back.

With the assumption of additive white Gaussian noise, the noisy matrix **X** can be written as **X** = **F** + **N**, where **N** denotes the noise matrix, and the true matrix **F** is low rank due to the

similarity of the grouped blocks. Let $||\mathbf{F}||_{\mathbf{w},p}$ be the weighted Schatten p-norm of the matrix $\mathbf{F}$

$$||\mathbf{F}||_{\mathbf{w},p} = \left(\sum_{l=1}^{L} w_l \delta_l^p \right)^{\frac{1}{p}}, \tag{10}$$

where L is the rank of $\mathbf{F}$, δ_l is the l-th singular value of $\mathbf{F}$, and $\mathbf{w}$ is a weight vector composed of the non-negative weights w_l. With the WSNM, the true matrix $\mathbf{F}$ is recovered from the noisy matrix $\mathbf{X}$ by solving

$$\arg \min_{\mathbf{F}} ||\mathbf{X} - \mathbf{F}||_{\mathrm{F}}^2 + ||\mathbf{F}||_{\mathbf{w},p}^p, \tag{11}$$

where the penalty factors in Eq. (5) are merged into the weight vector $\mathbf{w}$, which adjusts not only the weights of various singular values but also the balance between the F-norm fidelity term and the Schatten p-norm regularization term as a whole. In practice, the noise and artifacts introduced by the sparse-view reconstruction are very complex, which do not obey the Gaussian assumption. However, the WSNM can work well, as most of noise and artifacts will increase the weighted Schatten p-norm of the data matrix.

Let σ_N^2 denote the variance of the noise, which can be estimated effectively with the filter-based approaches from the noisy matrix $\mathbf{X}$. The weights in $\mathbf{w}$ should be proportional to σ_N^2 and increase with the size of the matrix to fit the F-norm of the fidelity term. In addition, since larger singular values normally correspond to important structural information that should be preserved more in the denoised voxels, smaller weights should be adopted for larger singular values, i.e. $\omega_l \leq \omega_{l+1}$ for $\delta_l \geq \delta_{l+1}$. As suggested in [27], we set the weights as

$$w_l = c\sqrt{S}\sigma_N^2/(\delta_l^{1/p} + \varepsilon), \tag{12}$$

where c is a constant, S is the number of voxels contained in $\mathbf{X}$, and ε is a very small constant to avoid dividing by zero. Because the true matrix $\mathbf{F}$ is unknown, its singular values δ_l in Eq. (12) can be estimated by

$$\hat{\delta}_l = \sqrt{\max\{\sigma_l^2 - S\sigma_N^2, 0\}}, \tag{13}$$

where σ_l corresponds to the l-th singular values of the noisy matrix $\mathbf{X}$.

Let $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{T}}$ be the SVD of $\mathbf{X}$, where Σ is the diagonal matrix of singular values $\{\sigma_l\}_{l=1}^{L}$ sorted in descending order. As proven in [27], the optimal solution of Eq. (11) can be computed by $\mathbf{F} = \mathbf{U}\Delta\mathbf{V}^{\mathrm{T}}$, where the singular values of $\mathbf{F}$ contained in Δ are solved from

$$\arg \min_{\{\delta_l\}} \sum_{l=1}^{L} [(\delta_l - \sigma_l)^2 + w_l \delta_l^p], \; s. \; t. \;\; \delta_l \geq 0 \text{ and } \delta_l \geq \delta_{l+1}. \tag{14}$$

Equation (14) defines a complex non-convex and non-smooth multi-variable optimization problem with coupled inequality constraints. Fortunately, with the ascending weights ($0<w_l \leq w_{l+1}$) and the descending singular values ($\sigma_l \geq \sigma_{l+1}>0$), as shown in [27], Eq. (14) can be decoupled into L independent single-variable optimizations without any constraints

$$\arg \min_{\delta_l} (\delta_l - \sigma_l)^2 + w_l \delta_l^p, \tag{15}$$

which can be easily solved by using the GST algorithm [28].

Given p and w_l, a threshold is defined as

$$\tau = (2w_l(1-p))^{\frac{1}{2-p}} + w_l p(2w_l(1-p))^{\frac{p-1}{2-p}}. \tag{16}$$

If $\sigma_l<\tau$, $\delta_l = 0$ is the global optimum of Eq. (15), otherwise, the optimum satisfies

$$\delta_l - \sigma_l + w_l p\delta_l^{p-1} = 0 \tag{17}$$

and can be solved iteratively (for more details please refer to [28]).

Finally, by reassembling all denoised blocks back, the denoised voxels **f** can be obtained. The steps of the denoising algorithm is given in Algorithm 2.

Algorithm 2 Denoising voxels by WSNM

Input: Noisy voxels **x**

  1: Divide **x** into non-overlapping reference blocks

  2:    **for** each reference block, **do**

  3:       Search similar blocks to form a 2D matrix **X**

  4:       Singular value decomposition: $\mathbf{X} = \mathbf{U\Sigma V^T}$

  5:       Estimate weight vector **w** by using Eq. (12) and (13)

  6:       Calculate $\mathbf{\Delta} = \mathrm{diag}(\delta_1, \delta_2, \cdots, \delta_L)$ from Eq. (15) by using the GST algorithm [28]

  7:       Get denoised matrix $\mathbf{F} = \mathbf{U\Delta V^T}$

  8:    **end for**

  9: Reassemble denoised blocks to form **f**

Output: Denoised voxels **f**

---

4. Experiments and results

In order to evaluate the reconstruction results, three qualitative indices, the root mean square error (RMSE), peak signal to noise ratio (PSNR), and structural similarity index (SSIM) are used.

The RMSE is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{x}_i - x_i)^2}, \tag{18}$$

where $\hat{x}_i$ and x_i are the voxels of the reconstruction model and ground truth, respectively, and N is the total number of voxels. The PSNR is defined as

$$\mathrm{PSNR} = 20\log_{10}\left(\frac{2^b - 1}{\mathrm{RMSE}}\right), \tag{19}$$

where b is the bit number of the voxel value. Both RMSE and PSNR are based on the deviation of voxels and do not consider the structural information in the CBCT data, which may lead to inconsistent with the human perception. The SSIM measures the structural similarity of two signal, which is defined as

$$\mathrm{SSIM} = \frac{(2\mu_{\hat{\mathbf{x}}}\mu_{\mathbf{x}} + C_1)(2\sigma_{\hat{\mathbf{x}}x} + C_2)}{(\mu_{\hat{\mathbf{x}}}^2 + \mu_{\mathbf{x}}^2 + C_1)(\sigma_{\hat{\mathbf{x}}}^2 + \sigma_{\mathbf{x}}^2 + C_2)}, \tag{20}$$

where $\mu_{\hat{x}}$ and $\sigma_{\hat{x}}$ are the mean value and variance of N reconstructed voxels $\{\hat{x}_i\}$ respectively, μ_x and σ_x are those of the true voxels $\{x_i\}$, $\sigma_{\hat{x}x}$ is the covariance between $\{\hat{x}_i\}$ and $\{x_i\}$, and C_1 and C_2 are constants to avoid instability, which are set to $0.0001 \, (2^b - 1)^2$ and $0.0009 \, (2^b - 1)^2$, respectively.

The power p in the WSNM is set to 0.9 in the experiments. As indicated in [27], the best power p is proportional to the noise level of the data being processed. In terms of the CBCT reconstruction, the power p can be chosen according to the number of projection views used for reconstruction, which means that a smaller p is preferred for more sparse views. The size of blocks is $4 \times 4 \times 4$ and the number of similar blocks is 70.
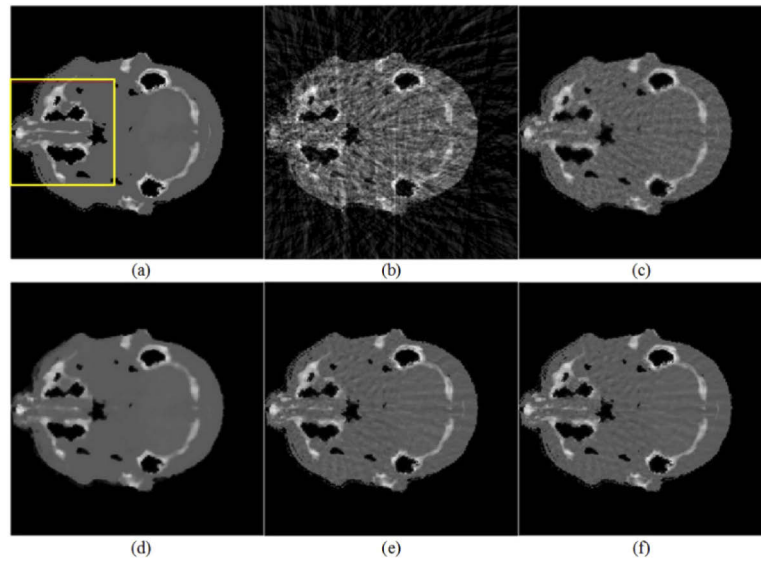
Fig. 1. Brain phantom reconstruction (slice 66): (a) ground truth, (b) FDK, (c) SART, (d) ASD-POCS, (e) WNNM-3D, (f) WSNM-3D

4.1. Reconstruction of digital brain phantom

The digital brain phantom is used for validation, which is generated based on the real MRI data of human brain and widely used for evaluating the CBCT reconstruction [33]. The phantom has $256 \times 256 \times 256$ voxels with the resolution of $1 \times 1 \times 1$ mm³. The Siddon's ray-driven algorithm [34] is used to simulate the projection data from 32 sparse views. The WNNM, proposed for conventional CT slice reconstruction in [35], is extended to 3D (labeled as WNNM-3D) as a benchmark for CBCT reconstruction. Besides, three classical methods discussed in Introduction, the FDK, SART, and ASD-POCS, are also used for comparison.

The reconstruction results are shown in Fig. 1, where the 66th slice in horizontal direction is selected for comparison due to containing rich structural information. The FDK proposed for full-view reconstruction cannot restore a clinically acceptable image from insufficient projection data. The reconstructed slice in Fig. 1(b) is seriously damaged by the streak artifacts. The SART is much better than the FDK, however, as shown in Fig. 1(c) the streak artifacts is still obvious within the tissue region. The ASD-POCS can remove the artifacts and suppress the noise very well, however, meanwhile the edges are blurred and some critical structural details are ignored in Fig. 1(d). The WNNM-3D and the proposed WSNM-3D methods perform excellently in the preservation of structure information. In terms of noise suppression, the proposed method (see Fig. 1(f)) is better than the WNNM-3D (see Fig. 1(e)) and much better than the SART. For a clearer comparison, the region of interest (ROI) marked by the box in Fig. 1(a) are zoomed in and shown in Fig. 2. As indicated by the red arrow, nearly perfect edges are restored by the WNNM-3D and the WSNM-3D methods, while the edges reconstructed by the ASD-POCS are too smooth and compared with the ground truth in Fig. 2(a) some structural details (indicated by the yellow arrow) are obviously lost in Fig. 2(d).

To show the difference between voxels intuitively, Fig. 3 shows the voxel curves and their local magnifications along the coronal axis (128th row, 66 slice) and the vertical axis (128th row, 128th column). As shown in Fig. 3(b) and (e), for the segments where the voxel values change greatly, the curves yielded by the proposed algorithm are the most consistent with the reference curves (ground truth), which means a better edge preservation ability, by contrast, the ASD-POCS is
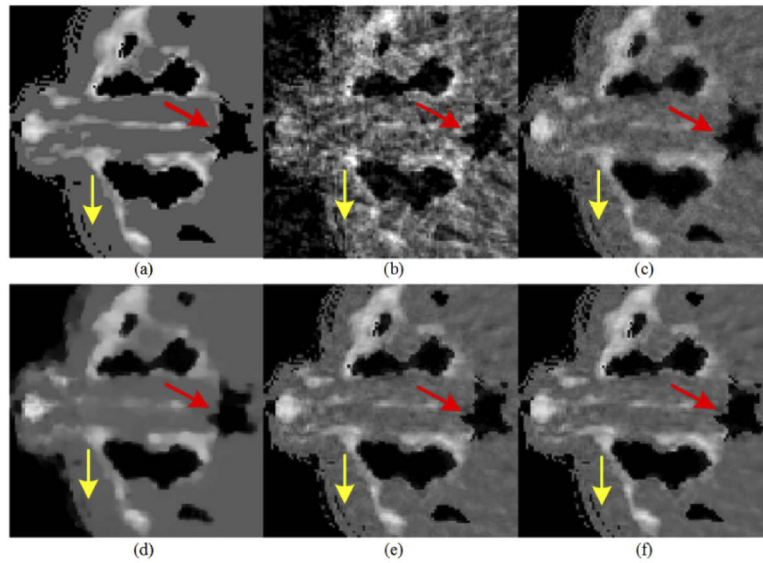
Fig. 2. Zoom in on ROI in Fig. 1: (a) ground truth, (b) FDK, (c) SART, (d) ASD-POCS, (e) WNNM-3D, (f) WSNM-3D

the worst performer, whose curves are too smooth. While for the smooth segments shown in Fig. 3(c) and (f), the ASD-POCS, WNNM-3D, WSNM-3D performs much better than the SART, whose curves are polluted seriously by the noise.
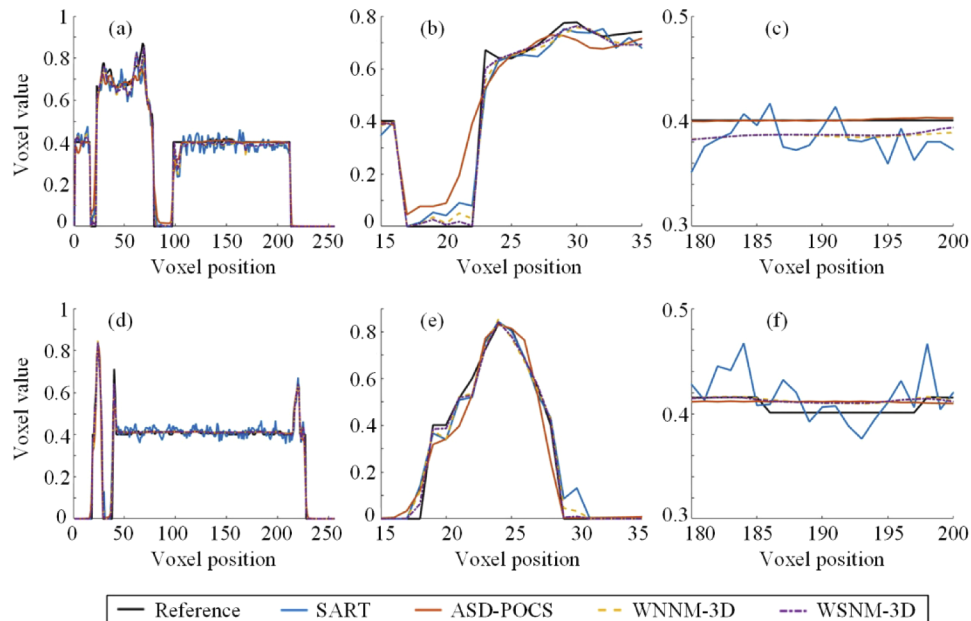

Fig. 3. Voxel curves reconstructed by various algorithms: (a) voxel curve along coronal axis at 128th row, 66th slice, (b)-(c) local magnifications of (a); (d) voxel curve along vertical axis at 128th row, 128th column; (e)-(f) local magnifications of (d).

The results of the quantitative evaluation are given in the Table 1. Consistent with the visual inspection, the proposed algorithm achieves the best scores in all three indices.

Table 1. Quantitative evaluation for brain phantom reconstruction from 32 views

	RMSE	PSNR	SSIM
FDK	0.0766	22.3154	0.9295
SART	0.0219	33.1911	0.9855
ASD-POCS	0.0227	32.8795	0.9846
WNNM-3D	0.0137	37.2656	0.9890
WSNM-3D	**0.0112**	**39.0156**	**0.9898**

4.2. Reconstruction of real clinical data

Clinical voxel data captured from an abdomen routine scan with a Neusoft NeuViz 128 CT system are also used for validation. The X-ray tube voltage is 120kVp, the tube current is 375mA, and the exposure time is 20ms/slice. The original voxels is $512 \times 512 \times 266$, which are down-sampled to $256 \times 256 \times 266$ with a resolution of $1.6367 \times 1.6367 \times 1$ mm^3 as the ground truth. The projection data simulated at 32 views are used for reconstruction.

Figures 4 and 5 show the reconstruction results at the slices 60 and 200, respectively, which are compared with the ground truth. The ROIs marked by the boxes are zoomed in and shown in Fig. 6. The results are consistent with those on the digital brain phantom. As marked by the arrows, the SART preserves edges well but introduces serious artifacts, while the ASD-POCS can suppress artifacts but also blur the edges. The WSNM-3D, slightly better than the WNNM-3D, yields more balanced results.
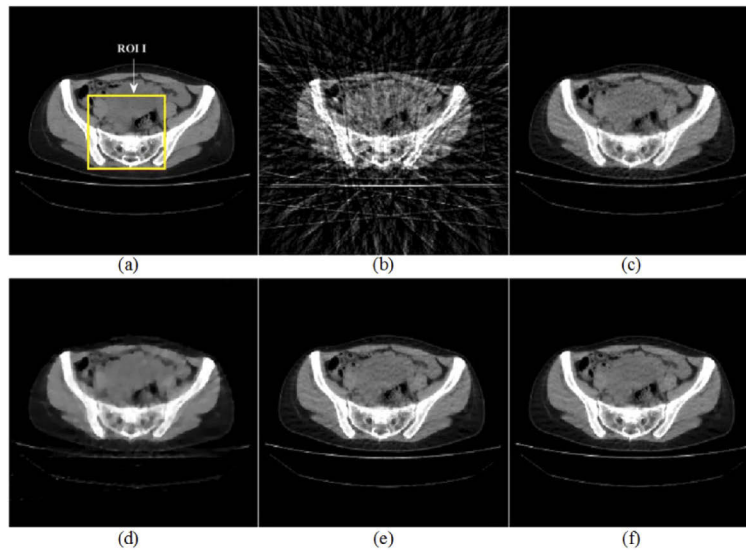

Fig. 4. Clinical data reconstruction (slice 60): (a) ground truth, (b) FDK, (c) SART, (d) ASD-POCS, (e) WNNM-3D, (f) WSNM-3D

Figure 7 shows the absolute difference images between the reconstruction slices and the ground truth. The WSNM-3D achieves the smallest differences at both slices, which further confirms the advantages of the proposed method.
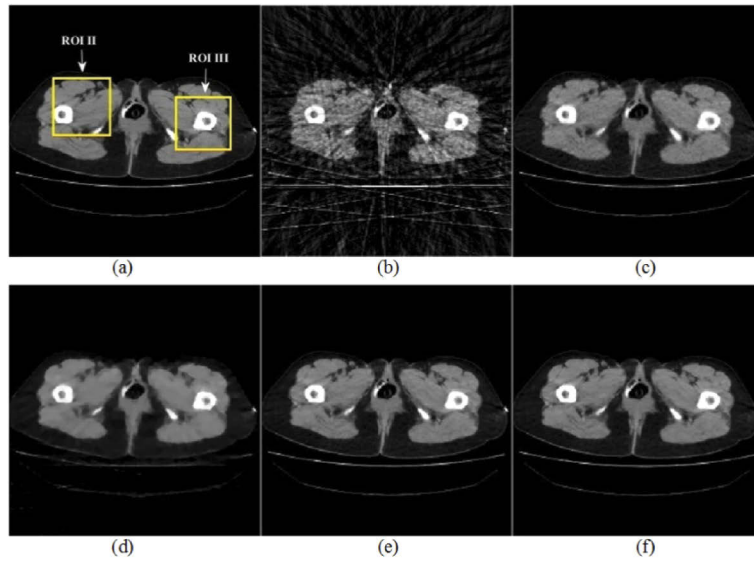
Fig. 5. Clinical data reconstruction (slice 200): (a) ground truth, (b) FDK, (c) SART, (d) ASD-POCS, (e) WNNM-3D, (f) WSNM-3D
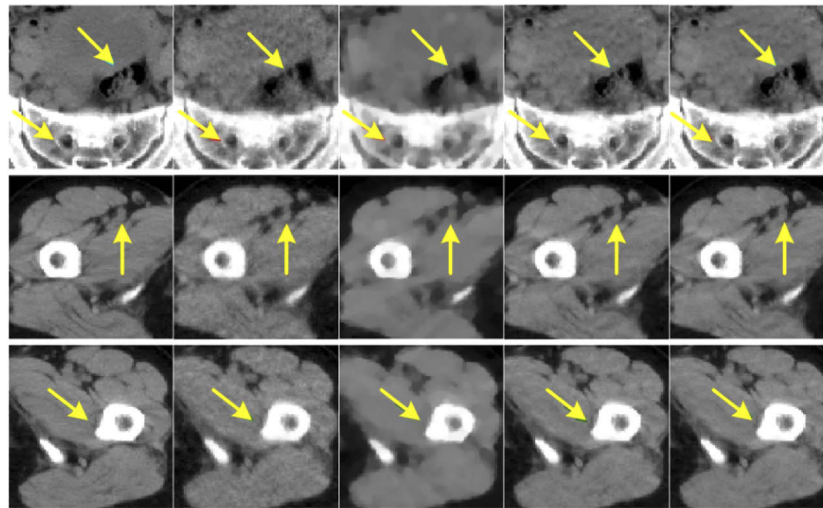

Fig. 6. Zoom in on the ROIs in Figs. 4 and 5. The three rows from top to bottom are ROIs I, II, and III, respectively. The columns from left to right correspond to the ground truth, SART, ASD-POCS, WNNM-3D and WSNM-3D.

For further validation, the reconstruction results in the sagittal and coronal planes are shown in Figs. 8 and 9, respectively. The results are consistent with those in the horizontal plane (see Figs. 4 and 5), the proposed algorithm achieves a good balance between the artifacts suppression and the structural information preservation.

Table 2 shows the quantitative evaluation of the reconstruction results for the clinical data. The best results are also achieved by the proposed algorithm.

Discussion on computational complexity: The most time-consuming step of the proposed algorithm is step 3 in Algorithm 1, i.e., the WSNM for denoising voxels (Algorithm 2), and
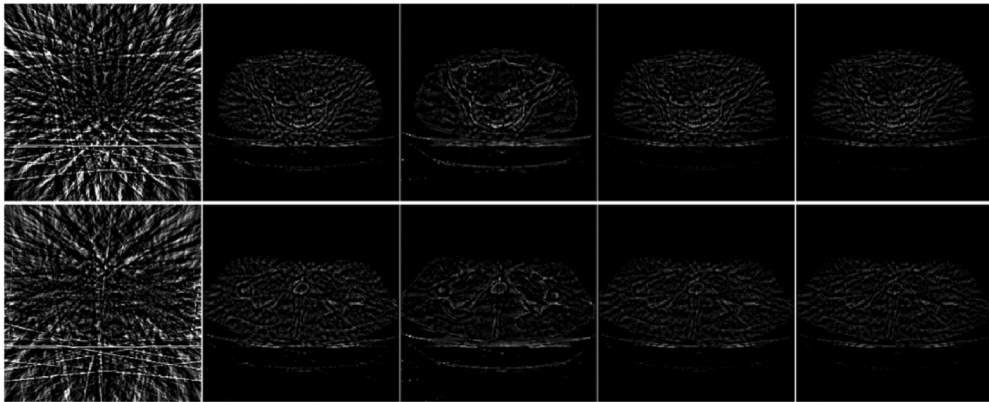
Fig. 7. Absolute difference images between the reconstruction slices and ground truth. The first row shows difference images of slice 60 and the second shows those of slice 200. The columns from left to right are obtained by FDK, SART, ASD-POCS, WNNM-3D, and WSNM-3D.
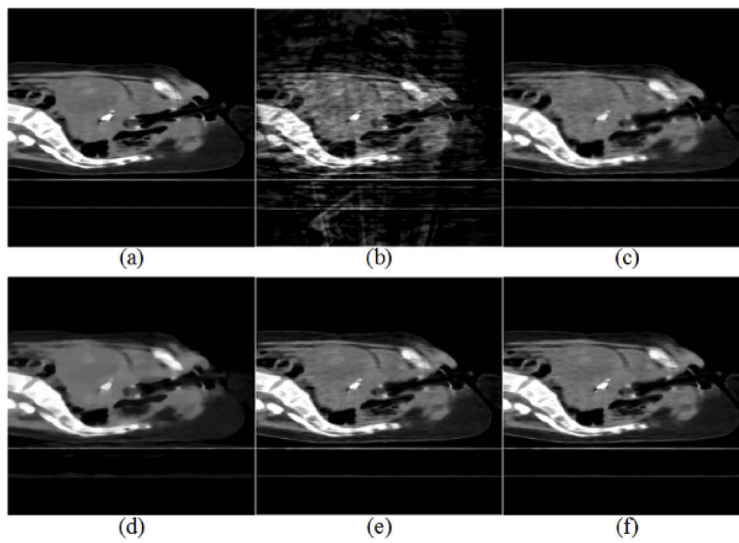

Fig. 8. Reconstruction results in sagittal plane: (a) ground truth, (b) FDK, (c) SART, (d) ASD-POCS, (e) WNNM-3D, (f) WSNM-3D

Table 2. Quantitative evaluation for clinical data reconstruction from 32 views

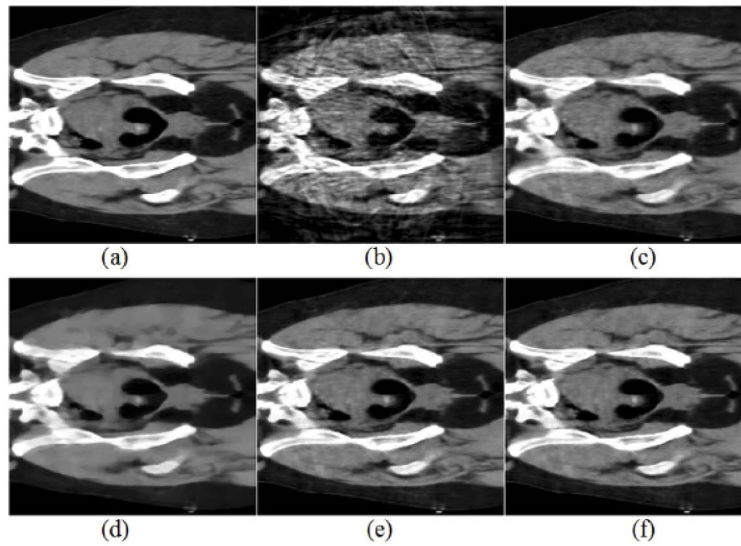|	RMSE	PSNR	SSIM
FDK	0.1350	17.3933	0.8345
SART	0.0354	29.0199	0.9743
ASD-POCS	0.0404	27.8724	0.9684
WNNM-3D	0.0261	31.6672	0.9800
WSNM-3D	**0.0242**	**32.3237**	**0.9811**

Fig. 9. Reconstruction results in coronal plane: (a) ground truth, (b) FDK, (c) SART, (d) ASD-POCS, (e) WNNM-3D, (f) WSNM-3D

the calculation of SVD (step 4) and GST (step 6) dominate the main computation in Algorithm 2. Assume that the size of blocks is $b \times b \times b$, and N similar blocks are searched for each reference block, corresponding to a $b^3 \times N$ noisy matrix $\mathbf{X}$. The computation complexity of SVD is $O(min(b^6N, b^3N^2))$ and the GST is $O(MN)$ [27], where M denotes the number of iterations in the GST algorithm. Letting K denote the total number of reference blocks in a reconstructed model, therefore, $O(K(min(b^6N, b^3N^2) + MN))$ is required for each WSNM iteration.

In our experiments, reconstructing a $256 \times 256 \times 266$ clinical voxel data with the WSNM-3D algorithm takes about 4 hours by using Matlab 2016a on a PC with 3.3GHz CPU and 12GB RAM. The WNNM-3D takes about the same amount of time. The FDK is the most efficient algorithm, taking about 3 minutes. The ASD-POCS and SART take about 20 and 30 minutes respectively. The WSNM-3D is the most time consuming methods in the tests. However, since the WSNM for each reference block is exactly the same, which can be computed in parallel and independently, the reconstruction speed of the WSNM-3D algorithm can be improved greatly by using GPU acceleration and more efficient programming languages (e.g., C/C++).

5. Conclusion

This paper proposes a novel iterative algorithm for sparse-view CBCT reconstruction based on the weighted Schatten *p*-norm minimization. By using the half quadratic splitting, the spare-view CBCT reconstruction problem is decomposed into two sub-problems: simple reconstruction sub-problem and denoising sub-problem, which can be solved alternately. The weighted Schatten *p*-norm minimization proposed in the field of image denoising is extended to 3D to regularize the denoising sub-problem. Benefiting from the extended WSNM, the structural similarity in the 3D voxel can be used effectively to improve the reconstruction results. The digital brain phantom and clinical CT data are used to verify the proposed method. Experimental results show that the proposed algorithm can effectively eliminate the staircase effect, remove artifacts and suppress noise. In terms of edge/structure information preservation, the proposed algorithm performs better than the classical algorithms. The results measured by the three quantitative indicators also confirm the advantage of the proposed algorithm.

In future work, noise models that are more consistent with the CBCT reconstruction than the additive white Gaussian noise will be considered during the voxel denoising. Complex low-dose problem caused by both reducing projection views and decreasing mAs (the product of tube current and exposure time) levels (that introduces noises directly in the projection images) will also be considered based on the proposed reconstruction scheme.

Funding

Disclosures

The authors declare that there are no conflicts of interest related to this article.

References

1. N. Wen, H. Guan, R. Hammoud, D. Pradhan, T. Nurushev, S. Li, and B. Movsas, "Dose delivered from Varian's CBCT to patients receiving IMRT for prostate cancer," Phys. Med. Biol. **52**(8), 2267–2276 (2007)..
2. Y. Gao, Z. Bian, J. Huang, Y. Zhang, S. Niu, Q. Feng, W. Chen, Z. Liang, and J. Ma, "Low-dose X-ray computed tomography image reconstruction with a combined low-mAs and sparse-view protocol," Opt. Express **22**(12), 15190–15210 (2014)..
3. 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. **28**(4), 475–490 (2001)..
4. L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," J. Opt. Soc. Am. A **1**(6), 612–619 (1984)..
5. S. Tang, K. Huang, Y. Cheng, T. Niu, and X. Tang, "Three-dimensional weighting in cone beam FBP reconstruction and its transformation over geometries," IEEE Trans. Biomed. Eng. **65**(6), 1235–1244 (2018).
6. R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography," J. Theor. Biol. **29**(3), 471–481 (1970).
7. A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm," Ultrasonic Imaging **6**(1), 81–94 (1984).
8. G. Wang and M. Jiang, "Ordered-subset simultaneous algebraic reconstruction techniques (OS-SART)," J. X-Ray Sci. Technol. **12**(3), 169–177 (2004).
9. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. Royal Stat. Soc.: Series B (Methodol.) **39**(1), 1–22 (1977)..
10. L. I. Rudin and S. Osher, "Total variation based image restoration with free local constraints," in *Proceedings of IEEE International Conference on Image Processing* (IEEE, 1994), pp. 31–35.
11. H. Bai, W. Zhang, J. Zhao, Y. Wang, and J. Sun, "New reconstruction method for few-view grating-based phase-contrast imaging via dictionary learning," Opt. Express **26**(20), 26566–26575 (2018)..
12. N. Viganò, P. M. Gil, C. Herzog, O. de la Rochefoucauld, R. van Liere, and K. J. Batenburg, "Advanced light-field refocusing through tomographic modeling of the photographed scene," Opt. Express **27**(6), 7834–7856 (2019)..
13. H. M. Huang and 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 **11**(4), e0153421 (2016).
14. T. Li, S. Tang, Z. Qiao, J. Yang, F. Yu, J. Xu, L. Fa, and Y. Zhu, "3-D weighted CBCT iterative reconstruction with TV minimization from angularly under-sampled projection data," Optik **172**, 161–176 (2018)..
15. E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," Phys. Med. Biol. **53**(17), 4777–4807 (2008)..
16. A. Cai, L. Wang, B. Yan, L. Lin, H. Zhang, and G. Hu, "Efficient TpV minimization for circular, cone-beam computed tomography reconstruction via non-convex optimization," Comput. Med. Imaging and Graphics **45**, 1–10 (2015)..
17. Y. Chen, F. F. Yin, Y. Zhang, Y. Zhang, and L. Ren, "Low dose CBCT reconstruction via prior contour based total variation regularization (PCTV): a feasibility study," Phys. Med. Biol. **63**(8), 085014 (2018)..
18. J. Liu, Y. Chen, Y. Hu, and L. Luo, "Low-dose CBCT reconstruction via 3D dictionary learning," in *Proceedings of IEEE International Symposium on Biomedical Imaging* (IEEE, 2016), pp. 735–738.
19. K. Choi, J. Wang, L. Zhu, T. Suh, S. Boyd, and L. Xing, "Compressed sensing based cone-beam computed tomography reconstruction with a first-order method," Med. Phys. **37**(9), 5113–5125 (2010)..
20. L. Liu, X. Li, K. Xiang, J. Wang, and S. Tan, "Low-dose CBCT reconstruction using hessian Schatten penalties," IEEE Trans. Med. Imaging **36**(12), 2588–2599 (2017)..
21. G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," IEEE Trans. Med. Imaging **37**(6), 1289–1296 (2018)..

22. H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, and G. Wang, "Low-dose CT via convolutional neural network," Biomed. Opt. Express **8**(2), 679–694 (2017)..

23. A. Cuadros, X. Ma, and G. R. Arce, "Compressive spectral X-ray tomography based on spatial and spectral coded illumination," Opt. Express **27**(8), 10745–10764 (2019)..

24. L. Liu, W. Huang, and D. R. Chen, "Exact minimum rank approximation via Schatten p-norm minimization," J. Comput. Appl. Math. **267**, 218–227 (2014)..

25. W. Dong, Z. Li, and D. Xiang, "Prior image constrained low-rank matrix decomposition method in limited-angle reverse helical cone-beam CT," J. X-Ray Sci. Technol. **23**(6), 759–772 (2015)..

26. S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," Int. J. Comput. Vis. **121**(2), 183–208 (2017)..

27. Y. Xie, S. Gu, Y. Liu, W. Zuo, W. Zhang, and L. Zhang, "Weighted Schatten p-norm minimization for image denoising and background subtraction," IEEE Trans. on Image Process. **25**(10), 4842–4857 (2016)..

28. W. Zuo, D. Meng, L. Zhang, and X. Feng, "A generalized iterated shrinkage algorithm for non-convex sparse coding," in *Proceedings of IEEE International Conference on Computer Vision* (IEEE, 2013), pp. 217–224.

29. K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2017), pp. 2808–2817.

30. K. Sauer and C. Bouman, "A local update strategy for iterative reconstruction from projections," IEEE Trans. Signal Process. **41**(2), 534–548 (1993)..

31. C. Han and J. Baek, "Multi-pass approach to reduce cone-beam artifacts in a circular orbit cone-beam CT system," Opt. Express **27**(7), 10108–10126 (2019)..

32. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," IEEE Trans. on Image Process. **16**(8), 2080–2095 (2007)..

33. A. Aichert, M. T. Manhart, B. K. Navalpakkam, and R. Grimm, "A realistic digital phantom for perfusion C-arm CT based on MRI data," in *Proceedings of IEEE Nuclear Science Symposium and Medical Imaging Conference* (IEEE, 2013), pp. 1–2.

34. R. I. Siddon, "Fast calculation of the exact radiological path for a three-dimensional CT array," Med. Phys. **12**(2), 252–255 (1985)..

35. Y. Zhang, K. Yang, Y. Zhu, and W. Xia, "NOWNUNM: nonlocal weighted nuclear norm minimization for sparse-sampling CT reconstruction," IEEE Access **6**, 73370–73379 (2018)..