Iterative image reconstruction for sparse-view CT via total variation regularization and dictionary learning

Xianyu Zhao[a,b], Changhui Jiang[a], Qiyang Zhang[a], Yongshuai Ge[a], Dong Liang[a], Xin Liu[a], Yongfeng Yang[a], Hairong Zheng[a] and Zhanli Hu[a,*]

[a]*Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China*
[b]*School of Information Engineering, Wuhan University of Technology, Wuhan, China*

Abstract. Recently, low-dose computed tomography (CT) has become highly desirable due to the increasing attention paid to the potential risks of excessive radiation of the regular dose CT. However, ensuring image quality while reducing the radiation dose in the low-dose CT imaging is a major challenge. Compared to classical filtered back-projection (FBP) algorithms, statistical iterative reconstruction (SIR) methods for modeling measurement statistics and imaging geometry can significantly reduce the radiation dose, while maintaining the image quality in a variety of CT applications. To facilitate low-dose CT imaging, we in this study proposed an improved statistical iterative reconstruction scheme based on the penalized weighted least squares (PWLS) standard combined with total variation (TV) minimization and sparse dictionary learning (DL), which is named as a method of PWLS-TV-DL. To evaluate this PWLS-TV-DL method, we performed experiments on digital phantoms and physical phantoms, and analyzed the results in terms of image quality and calculation. The results show that the proposed method is better than the comparison methods, which indicates the potential of applying this PWLS-TV-DL method to reconstruct low-dose CT images.

Keywords: Low dose computed tomography, penalized weighted least squares, total variation, dictionary learning

1. Introduction

X-ray computed tomography (CT) provides clear information on the attenuation of X-rays in different human body tissues at a millimeter scale, thus providing rich information on human organs for clinical diagnosis and prevention. CT has become an indispensable tool in the radiology diagnosis field [2]. However, as popularity of CT tomography has grown for clinical diagnosis, the problem of the radiation dose delivered during CT scans has attracted increasing attention. A large number of clinical studies have shown that CT radiation doses beyond the normal range can cause metabolic abnormalities and diseases such as cancer [3]. This is because a CT radiation dose accumulates over time; thus, repeated CT scans significantly increase the probability of carcinogenesis, which highlights the importance of low-dose CT. Low-dose CT reduces the radiation delivered to the patient by reducing the X-ray

---

*Corresponding author: Zhanli Hu, Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China. E-mail: zl.hu@siat.ac.cn.

tube load setting or reducing the number of projections. However, lowering the dose has the side effect of adding noise and stripe artifacts to the reconstructed image, which reduces the quality of CT reconstruction and affects the clinician's ability to diagnose abnormal tissues. Therefore, obtaining high-quality diagnostic CT images from low-dose or sparse view acquisitions is an important and much-studied research topic in the field of medical imaging.

At present, two main strategies exist to achieve low-dose CT (LDCT). The first strategy is to reduce the X-ray flux towards each detector element while the number of X-ray attenuation measurements is sufficient, and the second strategy is to decrease the number of X-ray attenuation measurements while the X-ray flux towards each detector element is normal. The first strategy is implemented by adjusting the tube current, tube voltage and exposure time of an X-ray tube. The second strategy produces insufficient projection data, suffering from few-view or limited-angle. Both of these strategies will lead to reconstruction results with artifacts. Thus, both strategies have high requirements for reconstruction algorithms and represent a major opportunity for algorithmic research.

Various techniques have been extensively investigated to reduce the radiation dose in CT examinations. Among them, statistical iterative reconstruction (SIR) methods that model the measurement statistics and imaging geometry allow the radiation dose to be reduced significantly while maintaining image quality in various CT applications compared with the filtered back-projection (FBP) reconstruction algorithm [4, 5]. Total variation (TV) has become a more popular form of prior knowledge than FBP. TV prior knowledge is based on the piecewise constant property of the CT image, which indicates that the image should have a small TV value. In practice, TV-based methods such as the adaptive steepest descent projection to convex set (ASD-POCS) method reconstruct CT images by solving TV minimization problems and use iterative optimization to find the minimal TV solutions [6–12].

Dictionary learning (DL) has attracted much attention in recent years. Some methods have incorporated dictionary learning into an iterative reconstruction framework [13–16], such as adaptive dictionary-based statistical iterative reconstruction (ADSIR) and 3D dictionary learning (3D-DL) [17]. The combination of dictionary learning with iterative reconstruction in these methods not only helps to improve the imaging performance but also increases the computational complexity due to their nested iterations. Other methods treat dictionary learning as postprocessing to reduce artifacts in images reconstructed with FBP, such as artifact suppression dictionary learning (ASDL) algorithms [18–25].

The statistical iterative reconstruction (SIR) method for modeling measurement statistics and imaging geometry can significantly reduce the radiation dose while maintaining image quality in various CT applications. The cost function of SIR consists of data fidelity and regularization terms. We use the TV regularization model based on the PWLS standard in which the weighted least squares (WLS) fidelity term considers the exact relationship between the variance and the mean of the projection data in the presence of electronic background noise. TV regularization uses the piecewise constant assumption (PCA), which can reconstruct high-quality CT images by TV minimization measured by sparse view without introducing significant artifacts. However, TV minimization cannot distinguish between real structure and image noise. Therefore, some structures may be lost or distorted, and block artifacts are generated in the reconstructed image. The inherent flaws of the TV minimization method prompted us to study other sparse representations, combined with dictionary learning (DL), and the sparse representation in redundant dictionaries achieved significant improvements. The DL method has significant effects in maintaining structure and suppressing image noise.

In this paper, we present an improved low-dose CT statistical iterative reconstruction method. Our goal is to reconstruct a sufficiently detailed image from a low-dose projection, reconstruct the intermediate image using TV minimization under the Penalized Weighted Least Squares (PWLS) standard [26], and postprocess the result using sparse coding dictionary learning to remove residual noise and produce clinically acceptable CT images. For simplicity, we term the proposed method "PWLS-TV-

DL." The novelty of PWLS-TV-DL is twofold. First, the weighted least squares fidelity term considers the exact relationship between the variance and the mean of the projection data in the presence of electronic background noise. Second, we present a novel algorithm for image reconstruction from few-view data utilizing the PWLS coupled with dictionary learning, sparse representation and TV minimization on two interconnected levels. The sparse constraint in terms of TV minimization has already led to good results for low-dose CT reconstruction and it is a global requirement, but it cannot directly reflect object structures. Compared to the discrete gradient transform used in the TV method, dictionary learning has been shown to be an effective approach to sparse representation. The sparse constraint in terms of a redundant dictionary is incorporated into an objective function in a statistical iterative reconstruction framework, and the dictionary can be adaptively defined during the reconstruction process. Moreover, unlike conventional image processing techniques that process an image pixel by pixel, the dictionary-based methods process an image patch by patch. Thus, it naturally and adaptively imposes strong structural constraints. By integrating TV and DL into the same frame to achieve a sparser representation of the signal, the introduction of an adaptively learned dictionary alleviates the artifacts caused by the piecewise constant assumption and allows accurate restoration of images with complex structures. Qualitative and quantitative evaluations were carried out on digital and physical phantoms, and the results are presented in terms of accuracy and resolution.

The rest of this paper is organized as follows. Section 2 states the PWLS image reconstruction model, TV minimization, DL-based CT reconstruction, and the PWLS-TV-DL algorithm workflow, along with the experimental setup and evaluation metrics. In Section 3, we present the results of the PWLS-TV-DL algorithm and comparison methods on simulation experiments and physical experiments, respectively. Finally, a discussion and conclusion are given in Section 4.

2. Methods

2.1. PWLS criteria for CT image reconstruction

The penalized weighted least-squares (PWLS) approach for iterative reconstruction of X-ray CT images was studied by Herman, Sauer and Bouman [27]. Based on the noise properties of CT projection data, the PWLS criterion for CT image reconstruction can be written as follows:

$$x^* = \arg \min_{x \geq 0} \left\{ (y - Hx)^T \sum^{-1} (y - Hx) + \beta R(x) \right\} \tag{1}$$

where y represents the obtained sinogram data, $y = (y_1, y_2, \ldots, y_M)^T$, and x is the image to be reconstructed, i.e., $x = (x_1, x_2, \ldots, x_N)^T$, where T denotes the matrix transpose. The operator H represents the system matrix which has a size of $M \times N$. The element H is the length of the intersection of projection ray i with pixel j. Σ is a diagonal matrix whose ith element is σ_i^2, which is the variance of the sinogram data x_i. $R(x)$ represents a prior term, and β is a hyperparameter penalty that controls the strength of the prior term. The goal for CT image reconstruction is to estimate the attenuation coefficients x from the measurement y with H.

Based on our previous works, in this study, the variance of σ_i^2 is determined by the following mean-variance relationship:

$$\sigma_i^2 = \frac{1}{I_0} \exp(\bar{y}_i) \left(\frac{1}{I_0} \exp(\bar{y}_i) \left(\sigma_e^2 - 1.25 \right) \right), \tag{2}$$

where I_0 denotes the incident X-ray intensity, $\bar{y}_i$ is the mean of the sinogram data in bin i, and σ_e^2 is the background electronic noise variance.

2.2. TV minimization

Total variation was first proposed by Rudin in the image denoising model in [28], which was used to measure image characteristics with up to a certain order of differentiation. Mathematically, the original TV of an image x can be defined as follows:

$$\mathrm{TV}(x) = \int_{\Omega} \left| \sqrt{(x_{i,j} - x_{i-1,j})^2 + (x_{i,j} - x_{i,j-1})^2} \right| \, dx, \tag{3}$$

where Ω is a bounded domain, and i and j represent the number of the row and column in which the image pixels are located, respectively. Another definition of TV can be written as follows:

$$\mathrm{TV}(x) = sup \left\{ \int_{\Omega} x \, div \, v \, dx \, \Big| v \in C_c^1 \left(\Omega, \, \mathbb{R}^d \right), \, \|v\|_{\infty} \leq 1 \right\} \mathrm{C}, \tag{4}$$

where *div* represents the divergence operator, v denotes the dual variable of the exact TV definition, and $\mathbb{R}^d$ denotes the d-dimensional real space. TV, as an edge-preserving penalty, promotes the performance of iterative CT image reconstruction from noisy or sparse-view projection measurements.

2.3. DL-based image reconstruction

Dictionary learning has the advantage of adaptability, allowing the dictionary to be approximated to the maximum initial signal. An efficient dictionary should also be characterized by multiscale, geometric invariance and redundancy. An adaptive dictionary with the above advantages is beneficial for sparse signal representation because richer dictionary features with greater similarity to the signal to be processed increase the probability of accurate signal approximation with fewer atoms.

The main purpose of the K-SVD algorithm is to train a suitable dictionary based on the training image samples so that the dictionary sparsely represents the training sample images. The objective function can be stated as [29]:

$$\min_{D_0, A} \|X - D_0 A\|_F^2 \ \textit{Subject to} \ \forall i, \ \|a_i\|_0 \leq T, \tag{5}$$

where X is the matrix of training samples, D_0 is the target dictionary, and A is the sparse coefficient matrix in which X is represented by D_0. Here, $\|\cdot\|_F^2$ represents the square of the Frobenius norm, defined as the sum of the squares of each atom in the matrix.

When the sparse coefficient calculation is completed, the K-SVD algorithm begins to update the dictionary to further reduce the error. The dictionary is updated individually for each matrix atom.

The DL heuristic uses a group of pixels, instead of a single pixel, as the minimum unit in the process. Under DL, the CT reconstruction problem can be stated as follows:

$$\min \left\{ _{x,w_j} \|x\|_{TV} + \sum_j \left\| D_0 w_j - R_j x \right\|_2^2 \right\} \ \textit{Subject to} \ \|w_j\|_0 \leq \rho \, \forall j, x \geq 0, \ \|Mx - g\|_2^2 < \varepsilon_a, \tag{6}$$

where $\|x\|_{TV}$ is the TV norm of an image x, $\|w_j\|_0$ is the l_0 norm of w_j, $R_j x = x^j$, R_j is the operator to extract an image patch from image x, ρ is the threshold of sparsity, M is a system matrix describing the forward projection, g is a measured dataset and ε_a is a small positive value.

Mathematically, Equation (6) contains three subproblems.

The first subproblem is as follows:

$$\min \left\{ _{w_j} \sum_j \left\| D_0 w_j - R_j x \right\|_2^2 \right\} \ \textit{subject to} \ \|w_j\|_0 \leq \rho \, \forall j. \tag{7}$$

This subproblem is called atom matching. With an image x, we must find the sparse representation w_j with the dictionary D_0, and the number of nonzero elements in w_j is fewer than that in ρ. We use the orthogonal matching pursuit to solve this problem.

The second subproblem is as follows:

$$\min_x \|x\|_{TV} \ subject \ to \ x \geq 0, \ \|Mx - g\|_2^2 < \varepsilon_a. \tag{8}$$

This subproblem is a TV minimization problem involving two constraints: nonnegativity $x \geq 0$ and data fidelity $\|Mx - g\|_2^2 < \varepsilon_a$. TV minimization has demonstrated its power for reconstructing relatively sparse images. Thus, the TV term here can help reduce image artifacts and reduce the probability of incorrect atom matching.

The third subproblem is as follows:

$$\min_x \sum_j \|D_0 w_j - R_j x\|_2^2. \tag{9}$$

This subproblem involves updating the image. The closed-form solution to this problem is as follows:

$$x = \left(\sum_j R_j^T R_j\right)^{-1} \sum_j R_j^T D_0 w_j, \tag{10}$$

where $\sum_j R_j^T R_j$ is a matrix whose diagonal elements represent the number of 'overlapped patches' at a given position.

2.4. PWLS–TV–DL algorithm

We formulate the CT reconstruction problem using DL and discuss the strategy for dictionary construction. In this section, we provide the implementation details of our PWLS–TV–DL algorithm.

The proposed algorithm consists of PWLS reconstruction, TV minimization and DL. PWLS serves as the reconstruction algorithm, and both TV minimization and DL serve as regularization terms. PWLS with TV minimization can reconstruct high-quality CT images by sparse view measurement, but the real structure and image noise cannot be distinguished, causing some structures to be lost or distorted, and block artifacts are generated in the reconstructed image. Integrating TV and DL into the same frame to achieve a sparser representation of the signal and the introduction of an adaptively learned dictionary alleviate the artifacts caused by the piecewise constant assumption and allow accurate restoration of images with complex structures.

TV minimization can be implemented using the gradient descent method. The magnitude of the gradient can be approximately expressed as follows:

$$\tau_{i,j} \approx \sqrt{\left(x_{i,j} - x_{i-1,j}\right)^2 + \left(x_{i,j} - x_{i,j-1}\right)^2}. \tag{11}$$

The image TV can be defined as $\|x\|_{TV} = \sum_i \sum_j \tau_{i,j}$. The steepest descent direction is then defined by

$$\nabla_{x_{i,j}} \|x\|_{TV} = \frac{\partial \|x\|_{TV}}{\partial x_{i,j}} \approx \frac{\left(x_{i,j} - x_{i-1,j}\right) + \left(x_{i,j} - x_{i,j-1}\right)}{\tau_{i,j} + \varepsilon} - \frac{\left(x_{i+1,j} - x_{i,j}\right)}{\tau_{i+1,j} + \varepsilon} - \frac{\left(x_{i,j+1} - x_{i,j}\right)}{\tau_{i,j+1} + \varepsilon}, \tag{12}$$

where ε is a small positive number in the denominator that allows any singularity to be avoided.

Then, TV minimization can be stated as follows:

$$x_k = x_{k-1} - \beta \cdot \Delta x \cdot \frac{\nabla \|x_{k-1}\|_{TV}}{\|\nabla \| x_{k-1} \ \|TV\|}, \tag{13}$$

where β is the length of each gradient-descent step, Δx is the difference between the results at the kth and (k–1)th iterations, and k is the iteration index.

The DL process includes atom matching and image updating. In the atom matching part, we need to find the sparse representation (SR) for each patch in the target image by the orthogonal matching pursuit (OMP) algorithm [30]. Because our dictionary is relatively large, searching the entire dictionary is impractical. However, a practical solution can be obtained in two steps.

(a) Initial dictionary D_0;
(b) Use the OMP algorithm to find a local w_j that minimizes the local reconstruction error:

$$\min_{w_j} \left\| R_j x - D_0 w_j \right\|_2^2 \ \textit{subject to} \ \left\| w_j \right\|_0 \leq \rho. \tag{14}$$

The OMP algorithm terminates when $\left\| R_j x - D_0 w_j \right\|_2^2 < \varepsilon_b$ or when the algorithm reaches the maximum number of iterations, where ε_b is a small positive value.

In the image updating part, the w_j in terms of the dictionary D_0 is used for updating an image patch $x^j = D_0 w_j$; then, the updated image patches are recorded in a matrix. The image patch values are not written back to the target image until all the image patches have been updated and recorded. Finally, image x is updated as follows:

$$x = \left(\sum_j R_j^T R_j \right)^{-1} \sum_j R_j^T x^j. \tag{15}$$

The workflow for the PWLS–TV–DL algorithm is summarized in Table 1.

3. Experiments and results

3.1. Experimental setup

To evaluate the performance of the PWLS-TV-DL method for CT image reconstruction, we conducted experiments on the digital extended cardiac-torso (XCAT) phantom and the Catphan physical phantom.

3.1.1. Digital XCAT phantom

Figure 1(a) shows a slice of the XCAT phantom. We chose a geometry that was representative of a monoenergetic fan-beam CT scanner setup with a circular orbit to acquire 1160 projection views over 2π. The number of channels per view was 672. The distance from the detector arrays to the X-ray source was 1040 mm, and the distance from the rotation center to the X-ray source was 570 mm. The reconstructed images were composed of 512×512 square pixels. Each projection datum along an X-ray through the sectional image was calculated based on the known densities and intersection areas of the ray with the geometric shapes of the objects in the sectional image.

Similar to previous studies (Wang et al. 2006) [31], we first simulated the noise-free sonogram data $\hat{y}$ and then generated the noisy transmission measurement I according to the statistical model of the prelogarithm projection data, that is,

$$y_i = Poisson \left(b_i \ \exp \left(-\hat{y} \right) \right) + Normal \left(0, \ \sigma_e^2 \right), \tag{16}$$

where b_i is the incident X-ray intensity, and σ_e^2 is the background electronic noise variance. In the simulation, b_i and σ_e^2 were set to 1.0×10^5 and 10.0, respectively, for low-dose scan simulation. Finally, the noisy sinogram data y were calculated by performing the logarithm transformation on the transmission data y_i. For the digital XCAT phantom experiment, the sparse-view projections were generated by undersampling the 1,160 views of normal-dose simulation to only 360 views evenly over 2π.

Table 1
Workflow for the PWLS-TV–DL algorithm

Input: x_0 - measured projections.
Output: x - reconstructed image.

Parameters: $\sqrt{n} \times \sqrt{n}$ - patch size, β - length of each gradient-descent step, k - iteration index, K - the maximum iteration number for main loop.

Initialization: Set x to 0.

Main loop for $k = 1, 2, \ldots$ K:
1. Reconstruct an image using the PWLS algorithm
$$x^* = \arg \min_{x \geq 0} \left\{ (y - Hx)^T \sum{}^{-1} (y - Hx) + \beta R(x) \right\}$$
2. TV-minimization loop
   2.1. Initialization: $\Delta x = \|\overline{x_0} - x_k\|$;
   2.2. TV gradient descent
$$x_k = x_{k-1} - \beta \cdot \Delta x \cdot \frac{\nabla \|x_{k-1}\|_{TV}}{\|\nabla \| x_{k-1} \|_{TV} \|};$$
3. For each patch $R_j \overline{x_K}$ in the image,
   (a) find the set $\widetilde{D_0}$ of nearest atoms in D_0;
   (b) compute the weights using OMP such that
$$\left\| R_j x - \widetilde{D_0} w_j \right\|_2^2 < \varepsilon_b$$
   or OMP reaches the maximum number of iterations;
   (c) estimate the image patch x_j as
$$x^j = D_0 w_j;$$
4. Update the image simultaneously:
$$x_{k+1} = \left(\sum_j R_j^T R_j \right)^{-1} \sum_j R_j^T x^j;$$
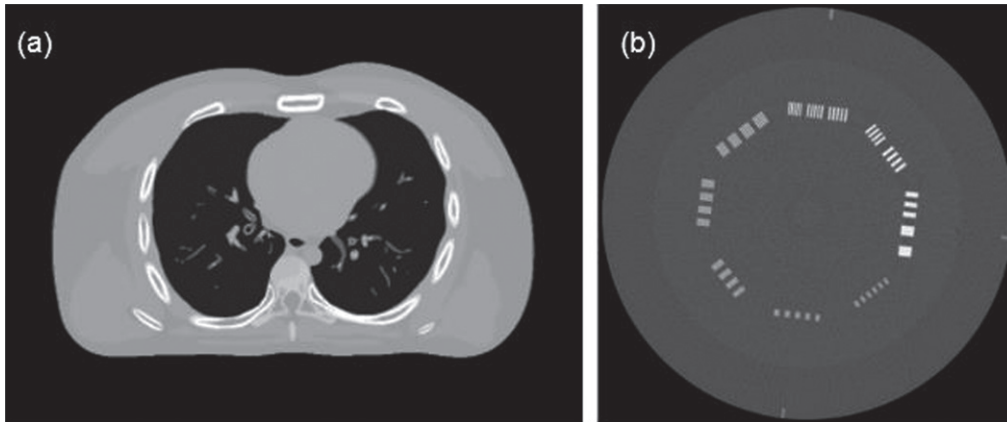Repeat from Step 1 until a termination criterion is satisfied.

Fig. 1. Digital and physical phantoms used in the studies: (a) a slice of the digital XCAT phantom; (b) the standard image reconstructed by the FBP method of the CTP714 module in the CatPhan-700.

3.1.2. Catphan physical phantom

Figure 1(b) shows the standard image reconstructed by the FBP method of the CTP714 module in the CatPhan-700 (Phantom Laboratory Inc., Salem, NY, USA). The experiment was performed on an in-house CT imaging bench. The system had a rotating-anode tungsten target diagnostic level X-ray tube (Varex G-242, Varex Imaging Corporation, UT, USA) and was operated in 120.00 kV continuous fluoroscopy mode with a 0.40 mm nominal focal spot. The X-ray tube current was set to 11.00 mA. The collimated beam had a vertical height of 25.00 mm at the rotation center. The X-ray detector was an energy-resolving photon-counting detector (XC-Hydra FX50, XCounter AB, Sweden) made from cadmium telluride (CdTe) with an imaging area of 512.00 mm × 6.00 mm and a native element dimension of 0.10 mm × 0.10 mm. The detector signal accumulation period was 0.10 seconds. The projection datasets were acquired with a 1 × 1 detector binning mode and were rebinned to 6 × 6 during the postprocessing procedures. Because there was no need to discriminate the photon energies, the detector was operated and calibrated with only a single energy threshold (10.00 keV). The source-to-detector distance was 1500.00 mm, and the source-to-rotation center was 1000.00 mm.

3.1.3. Performance evaluation

We also applied numeric metrics for objective assessment. The first metric is the peak signal-to-noise ratio (PSNR), which is defined based on the mean square error (MSE), where x is the reconstructed image, y is the reference image, N is the number of pixels, and MAX is the maximum pixel value [32]. PSNR is measured in decibel (dB) units; a larger value represents less image distortion.

$$\begin{cases} \text{MSE} = \frac{1}{N} \sum_i^N (x(i) - y(i))^2 \\ PSNR = 10 \, lg\left(\frac{MAX^2}{MSE}\right) \end{cases}. \tag{17}$$

The second metric is the structural similarity index (SSIM) [33], which measures the similarity between two images by considering the luminance l, contrast c, and structural information s, as follows:

$$\begin{cases} l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \\ c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \\ s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \end{cases}. \tag{18}$$

$$\text{SSIM}(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y),$$

where μ_x and σ_x are the mean and variance of image x, respectively, μ_y and σ_y are the mean and variance of image y, respectively, σ_{xy} is the covariance of x and y, and C_1, C_2, and C_3 are constants. The SSIM index yields a value from 0 to 1; values closer to 1 denote greater image similarity.

The third metric is the root mean square error (RMSE), which indicates the difference between the reconstructed image and the ground truth image and characterizes the reconstruction accuracy.

$$\text{RMSE} = \sqrt{\frac{\sum_{x=1}^{X} (\mu_x - \mu_x^*)^2}{X}}, \tag{19}$$

where μ_x is the true image, μ_x^* is the reconstructed image, and X is the number of image pixels.

In addition to image quality comparisons, the computational cost of the proposed method was evaluated in our study. All the experiments were implemented using MATLAB 2016b and executed on a PC with an Intel Core i7-6700 CPU @ 3.40 GHz and 16.0 GB of memory.

3.1.4. Reconstruction parameter selection

As observed from the objective function, parameter β is the penalty factor that controls the proportion of the penalty term. In this experiment, the constant step size was set to 1. In the DL process, the size of the patch is set to 6*6. For the method of this paper, parameter β is set by observing the experimental results, and its value will also be discussed in subsequent experiments.

3.2. Performance evaluation on digital phantom

In the digital phantom study, the original phantom data were directly used as the ground-truth image. As mentioned above, low-dose CT can be implemented by lowering the tube current or reducing projection views. To achieve a more comprehensive study, we tested the proposed method using a low-current case and a sparse-view case.

3.2.1. Experiment in a low-current case

In the low-current case, the incident photon count was set to $b_i = 1.0 \times 10^5$ photons per ray, and 360 views over $360°$ were simulated. The imaging results of the tested methods are shown in Fig. 2. The PWLS result in Fig. 2(a) contains heavy noise in the entire reconstructed region. Even worse,
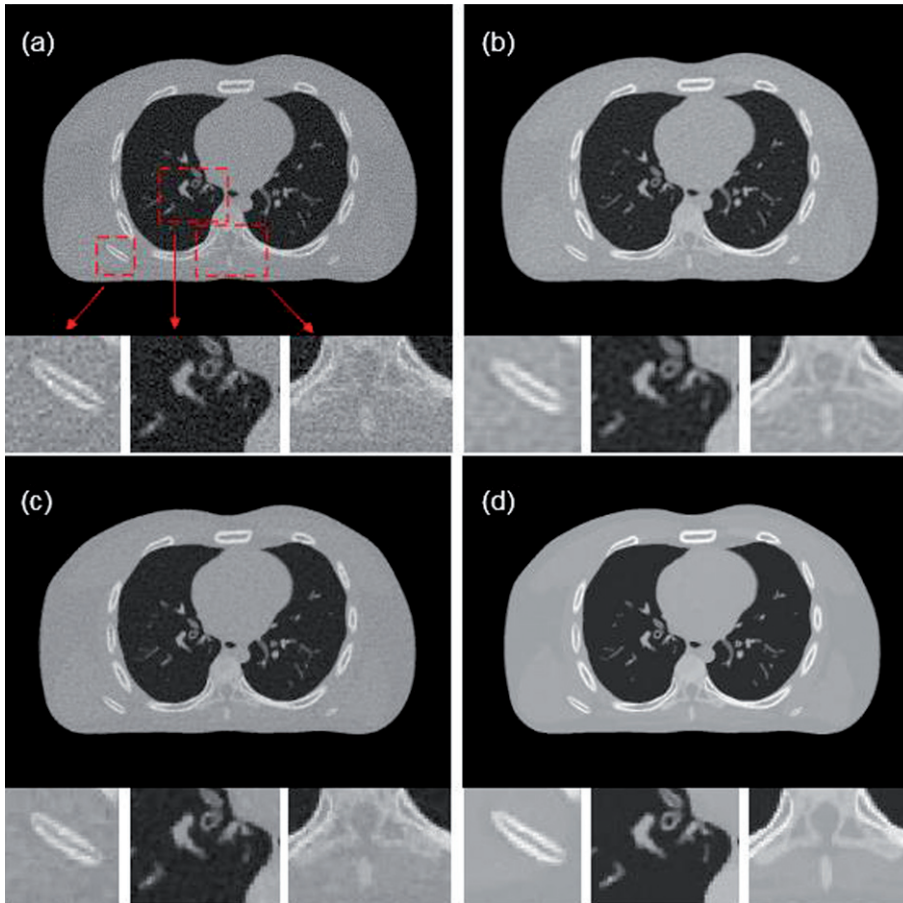

Fig. 2. Imaging results of different methods on the digital XCAT: (a) PWLS; (b) PWLS-DL; (c) PWLS-TV ($\beta = 6$); (d) PWLS-TV-DL ($\beta = 8$).

some small structures are covered almost entirely by noise, which can cause misdiagnosis in the clinic. Figure 2(b) shows the image reconstructed by PWLS-DL. This result is better than that of PWLS, but it still contains some noise. As shown in Fig. 2(c), the image is much cleaner after being denoised by TV, but the edges in the image are seriously blurred compared those in the true image in Fig. 1(a). In contrast, it is easy to see in Fig. 2(d) that the image reconstructed by PWLS-TV-DL effectively suppresses noise and artifacts. As the number of iterations increases, the RMSE value decreases. Obviously, the proposed method is better than the other methods.

The quantitative assessment was carried out by calculating the PSNR, SSIM, and RMSE between the true image and the reconstructed images. Table 2 shows a performance comparison of the different methods. The column values in Table 2 demonstrate that regardless of whether we consider PSNR, SSIM, or RMSE, the proposed method achieves a good result.

The profile images and residual images are compared in Figs. 3 and 4, respectively. Clearly, the PWLS-TV-DL curve is closer to the Phantom curve. The results show that the proposed method can achieve better-quality images than those of the compared methods.

3.2.2. Experiment in a sparse-view case

In the sparse-view case, the incident photon count was set to $b_i = 1.0 \times 10^6$ photons per ray, and 180 projections were used for sparse-view reconstruction [1]. The imaging results of the tested methods are shown in Fig. 6. The PWLS result in Fig. 6(a) contains heavy noise in the entire reconstructed region. Figure 6(b and c) show the image reconstructed by PWLS-DL and PWLS-TV, respectively.

Table 2
Numeric results on XCAT (view = 360 $b_i = 1.0 \times 10^5$)

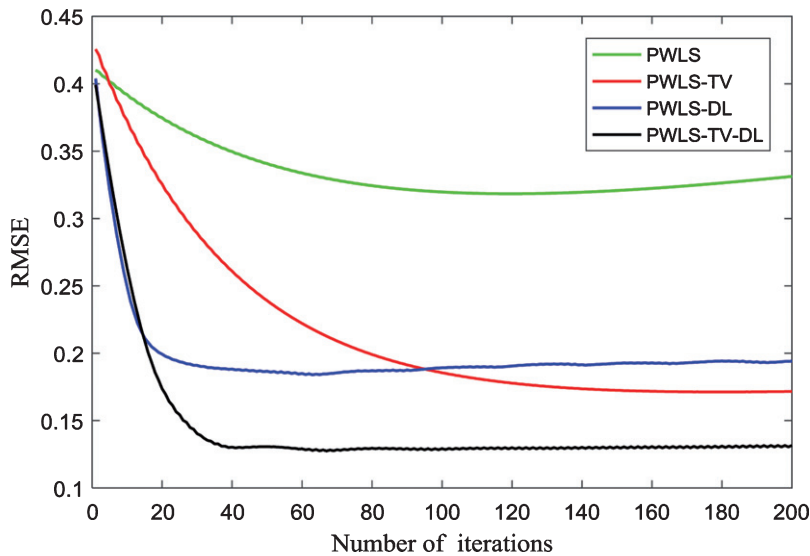| Method	PWLS	PWLS-DL	PWLS-TV	PWLS-TV-DL
PSNR	23.892	25.731	33.112	**44.854**
SSIM	0.936	0.948	0.977	**0.991**
RMSE	0.331	0.194	0.172	**0.131**

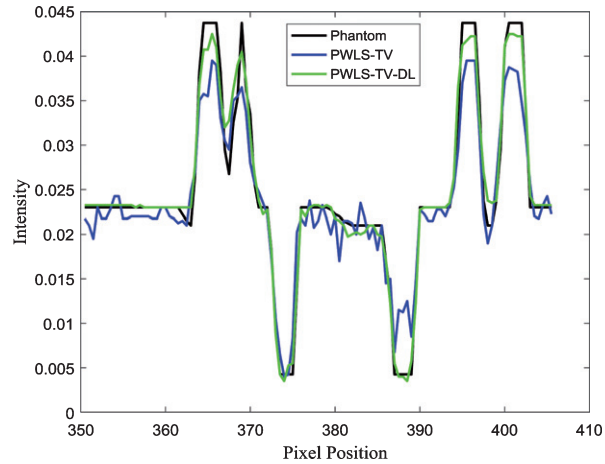Fig. 3. The curves of RMSE values versus the number of iterations.

Fig. 4. The profiles were located at the pixel positions x from 350 to 410 and *y* = 350. The Phantom curve represents the profile of the true image in Fig. 1(a). The PWLS-TV curve represents the profile of the reconstructed image using the PWLS-TV method in Fig. 2(c). The PWLS-TV-DL curve represents the profile of the reconstructed image using the PWLS-TV-DL method, as shown in Fig. 2(d).
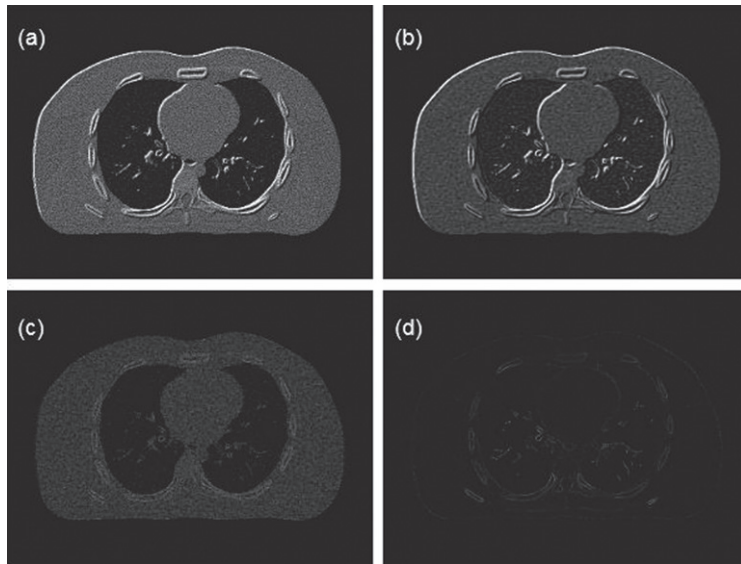

Fig. 5. Residual images of the reconstructed results based on the PWLS method, the PWLS-DL method, the PWLS-TV method, and the PWLS-TV-DL method with 200 iterations on the simulation XCAT data. All the images are displayed in the same window.

Both results are better than that of PWLS but still contain some noise. In contrast, as shown in Fig. 6(d), the image reconstructed by PWLS-TV-DL suppresses noise and artifacts.

The quantitative assessment was carried out by calculating the PSNR, SSIM, and RMSE between the true image and the reconstructed images. Table 3 shows a performance comparison of the different methods. The column values in Table 3 demonstrate that regardless of whether we consider PSNR, SSIM, or RMSE, the proposed method achieves a good result.

The residual images are compared in Fig. 7. Clearly, the results show that the proposed method can achieve better-quality images.
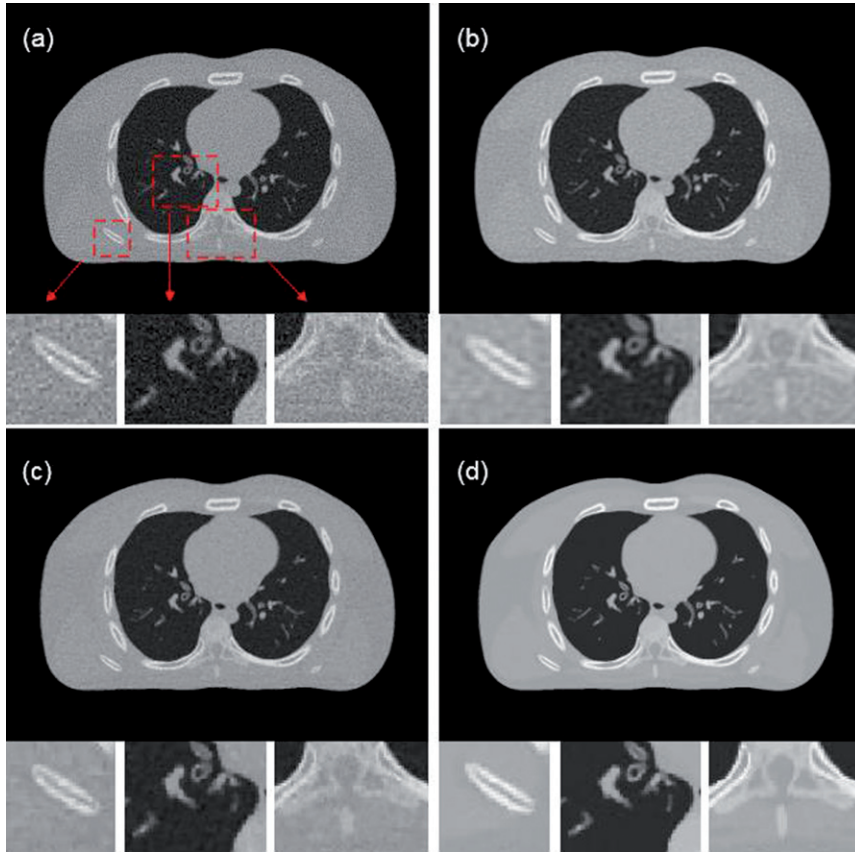
Fig. 6. Imaging results of different methods on the digital XCAT: (a) PWLS; (b) PWLS-DL; (c) PWLS-TV ($\beta = 2$); and (d) PWLS-TV-DL ($\beta = 4$).

Table 3
Numeric results on XCAT (view = 180 $b_i = 1.0 \times 10^6$)

Method	PWLS	PWLS-DL	PWLS-TV	PWLS-TV-DL
PSNR	25.767	28.751	32.136	**37.491**
SSIM	0.936	0.954	0.962	**0.987**
RMSE	0.397	0.198	0.177	**0.151**

3.3. Performance evaluation on physical phantom

3.3.1. Experiment in a sparse-view case

In the physical phantom study, an image reconstructed with 720 views by the FBP method was used as the true image. To further evaluate the proposed method, we also investigated a much worse case, where the view was lowered to 120.

The results are shown in Fig. 8. The PWLS image Fig. 8(a) contains obvious artifacts. In the PWLS-DL Fig. 8(b) and PWLS-TV Fig. 8(c) images, some artifacts are almost completely missing. As shown in Fig. 8(d), PWLS-TV-DL achieves an acceptable result—almost the same as the true image. Based on Fig. 9, the image quality index obviously improves as the number of iterations increases.
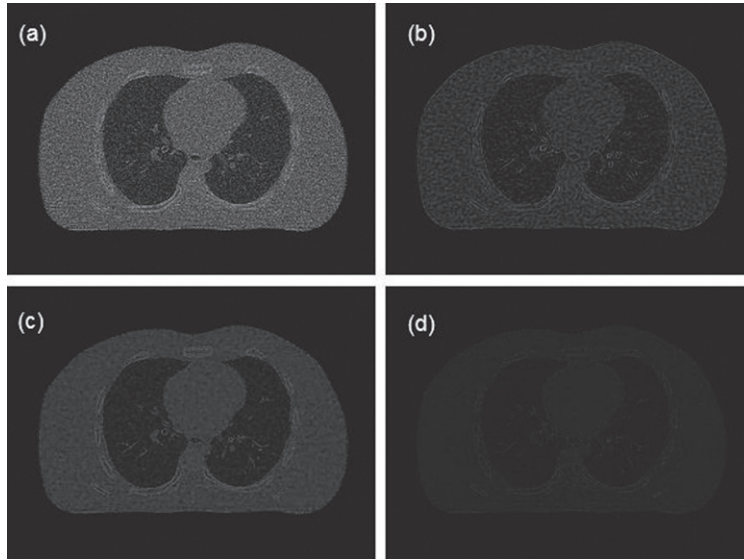
Fig. 7. Residual images of the reconstructed results based on the PWLS method: (a) PWLS; (b) PWLS-DL; (c) PWLS-TV; and (d) PWLS-TV-DL. All the images are displayed in the same window.
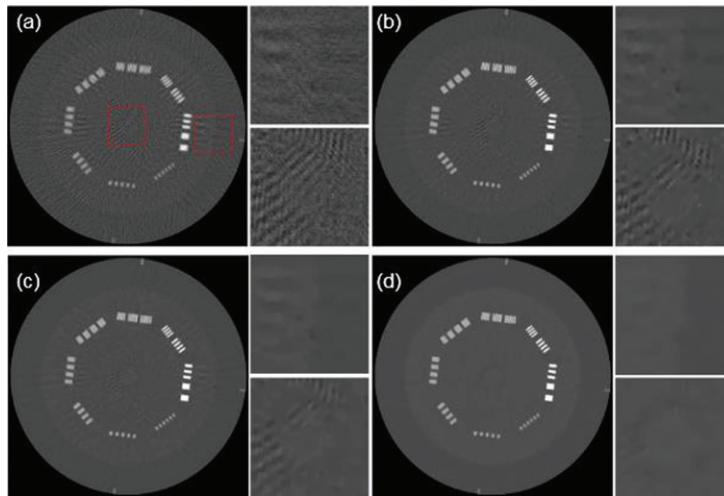

Fig. 8. Imaging results of the CTP714 module in the CatPhan-700: (a) Image reconstructed from the PWLS algorithm; (b) Image reconstructed from the PWLS-DL algorithm; (c) Image reconstructed from the PWLS-TV algorithm; (d) Image reconstructed from the PWLS-TV-DL algorithm.

Table 4 compares the numeric results of these methods. Notably, the PWLS-TV-DL method obtains the best results on all the numeric metrics. The residual images were compared in Fig. 10. We can conclude that the proposed method has a clear performance advantage over the compared methods.

3.3.2. Algorithm performance comparison at a fixed total dose with different view angle sampling

In reality, an interesting question of practical importance is as follows: given a total radiation dose, what is the best manner to distribute the dose—would delivering the dose to more view angles with a
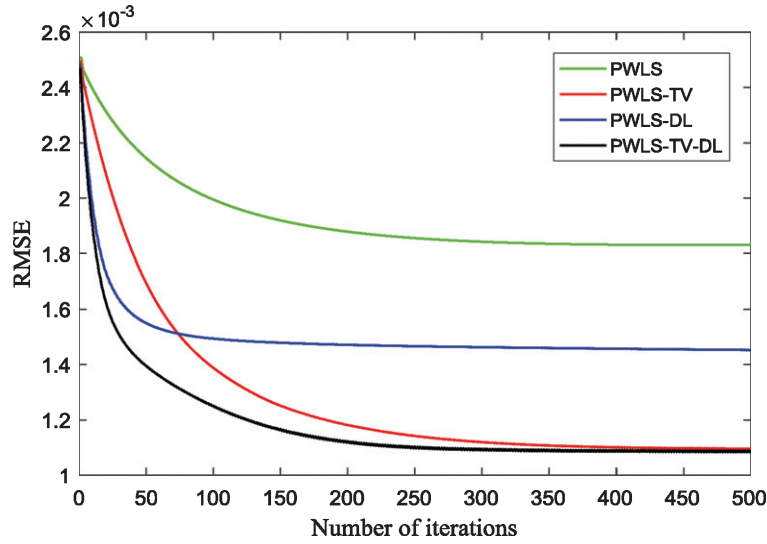
Fig. 9. The curves of the RMSE values versus the number of iterations.

Table 4

Numeric results on Catphan (120 views, 11 mA, 0.1 s)

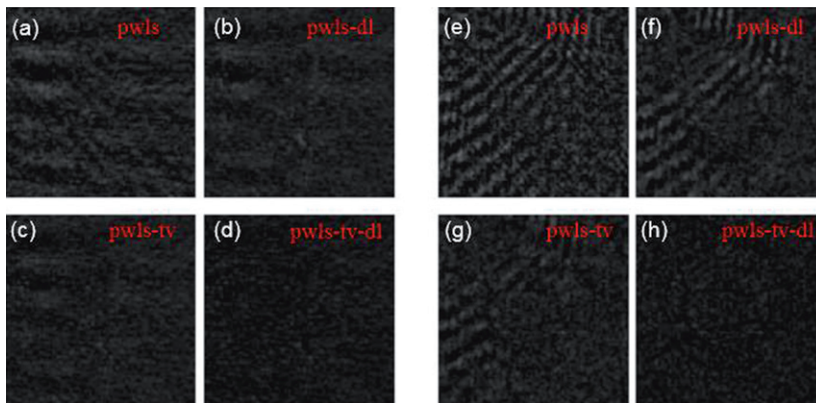| Method	PWLS	PWLS-DL	PWLS-TV	PWLS-TV-DL
PSNR	27.747	30.653	32.546	**32.621**
SSIM	0.794	0.897	0.925	**0.929**
RMSE	0.0018	0.0015	0.0011	**0.0010**

Fig. 10. Residual images of the reconstructed results based on the PWLS method, the PWLS-DL method, the PWLS-TV method, and the PWLS-TV-DL method with 500 iterations. All the images are displayed in the same window.

low mAs/view be better, or would distributing the total dose with a higher mAs/view be better [34]? In this study, we fixed the total radiation dose level to 99 mAs. Two cases for distributing this total dose were used: (180 views, 5.5 mA, 0.1 s) and (90 views, 11 mA, 0.1 s). The reconstructed images are shown in Figs. 11 and 12, respectively. In these images, we can see that in both cases, PWLS-TV-DL provides better reconstruction results than the other methods, with fewer streaking artifacts.
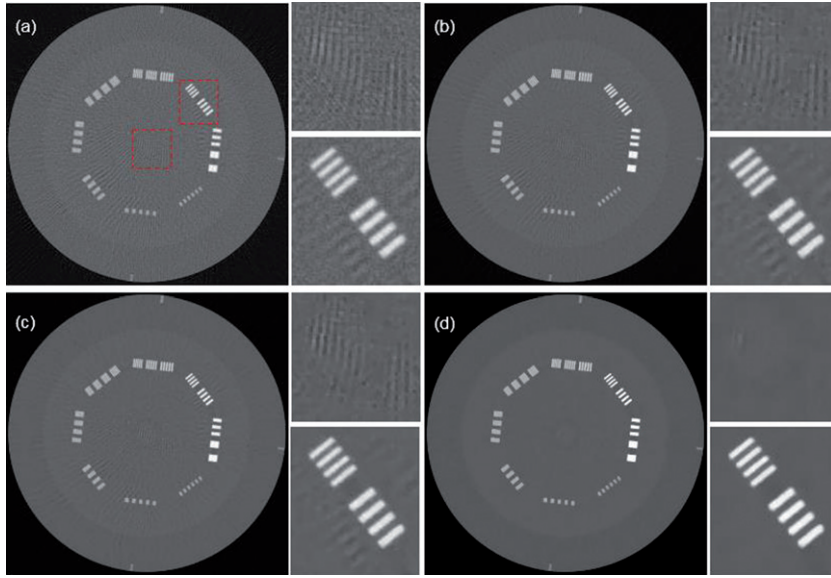
Fig. 11. Imaging results of different methods in the case of 180 views and 5.5 mA: (a) PWLS; (b) PWLS-DL; (c) PWLS-TV ($\beta = 0.01$); and (d) PWLS-TV-DL ($\beta = 0.02$).
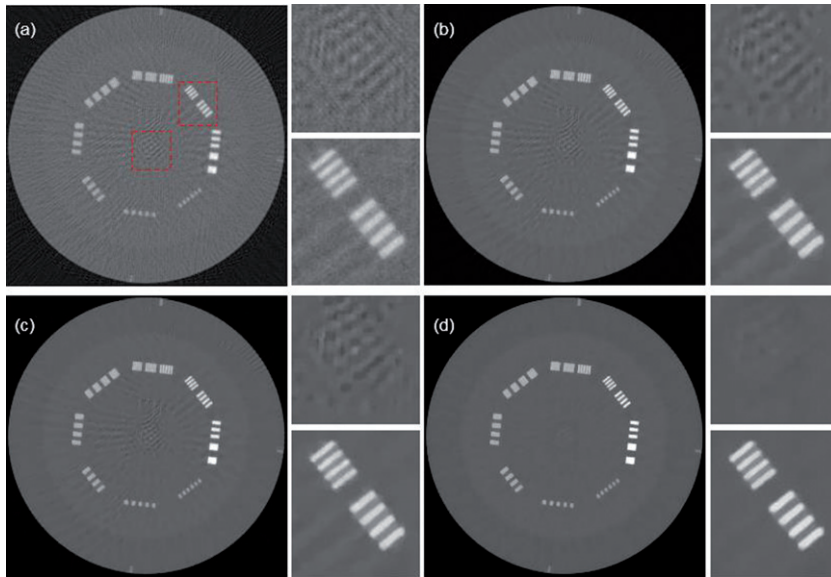

Fig. 12. Imaging results of different methods in the case of 90 views and 11 mA: (a) PWLS; (b) PWLS-DL; (c) PWLS-TV ($\beta = 0.01$); and (d) PWLS-TV-DL ($\beta = 0.02$).

Tables 5 and 6 show a performance comparison of the different methods in the two cases, respectively. The quantitative assessment was carried out by calculating the PSNR, SSIM, and RMSE between the true image and the reconstructed images. In the case of 180 views, the overall reconstruction results are little better than those with 90 views.

Table 5

Numeric results on Catphan (180 views, 5.5 mA)

Method	PWLS	PWLS-DL	PWLS-TV	PWLS-TV-DL
PSNR	26.212	29.961	30.911	**31.884**
SSIM	0.811	0.884	0.902	**0.913**
RMSE	0.0019	0.0015	0.0012	**0.0010**

Table 6

Numeric results on Catphan (90 views, 11 mA)

Method	PWLS	PWLS-DL	PWLS-TV	PWLS-TV-DL
PSNR	25.787	29.333	30.468	**31.290**
SSIM	0.776	0.871	0.892	**0.904**
RMSE	0.0019	0.0015	0.0013	**0.0011**

4. Discussion and conclusion

In this study, based on the PWLS standard, we proposed and demonstrated a new low-dose CT reconstruction solution by combining TV minimization and sparse dictionary learning. First, an intermediate image is reconstructed using TV minimization; then, it is postprocessed using dictionary learning to remove residual noise and produce a clinically acceptable CT image. As demonstrated by the simulation and physical experiment results, compared with reconstruction methods such as PWLS, PWLS-DL and PWLS-TV, the proposed method improves the quality of reconstructed images and produces smaller RMSE and larger PSNR and SSIM values. However, the main shortcoming of the PWLS-TV-DL algorithm is that matrix update in dictionary learning increases the computational burden and requires a long running time. To solve this problem, a fast computer and dedicated hardware are required. In the future, we believe that iterative-based image reconstructions such as the PWLS-TV-DL algorithm will be widely used in medical clinics [35].

References

[1] H. Yan, L. Cervino, X. Jia and S.B. Jiang, A comprehensive study on the relationship between the image quality and imaging dose in low-dose cone beam CT, _Phys Med Biol_ **57** (2012), 2063–2080.
[2] L. Yu, X. Liu, S. Leng, et al., Radiation dose reduction in computed tomography: Techniques and future perspective, _Imaging Med_ **1** (2009), 65–84.
[3] H.M. Zhang and e. al., Constrained Total Generalized p-Variation Minimization for Few-View X-Ray Computed Tomography Image Reconstruction, _Plos One_ **11** (2016), 98–99.

[4] S. Niu, Y. Gao, Z. Bian, et al., Sparse-view X-ray CT reconstruction via total generalized variation regularization, *Phys Med Biol* **59** (2014), 2997–3017.

[5] Z. Hu, Y. Zhang, J. Liu, et al., A feature refinement approach for statistical interior CT reconstruction, *Phys Med Biol* **61** (2016), 5311–5334.

[6] Y. Zhang, W. Zhang and J. Zhou, Accurate sparse-projection image reconstruction via nonlocal TV regularization, *The Scientific World Journal* **22** (2014), 458–496.

[7] C. Z., J. X., L. L. and W. G., A limited-angle CT reconstruction method based on anisotropic TV minimization, *Phys Med Biol* **58** (2013), 2119–2141.

[8] Y. Wang and Z. Qi, A new adaptive-weighted total variation sparse-view computed tomography image reconstruction with local improved gradient information, *J Xray Sci Technol* **26** (2018), 957–975.

[9] E.Y. Sidky, C.M. Kao, X.C. Pan, Image reconstruction in circular cone-beam computed tomogrphy by constrained, total-variation minimization, *Phys Med Biol* **53** (2008), 4777–4807.

[10] Y. Liu, J. Ma, Y. Fan and Z. Liang, Adaptive-weighted Total Variation Minimization for Sparse Data toward Low-dose X-ray Computed Tomography Image Reconstruction, *Phys Med Biol* **57** (2012), 7923–7956.

[11] B. Chen and e. al., A new Mumford–Shah total variation minimization based model for sparse-view x-ray computed 291 tomography image reconstruction, *Neurocomputing* **285** (2018), 74–81.

[12] Z. Hu, Q. Liu, N. Zhang, et al., Image Reconstruction from Few-view CT Data by Gradient-domain Dictionary Learning, *J Xray Sci Technol* **24** (2016), 627–638.

[13] Q. Xu, H. Yu, X. Mou, et al., Low dose X-ray reconstruciton via dictionary learning, *IEEE Trans Med Imaging* **31** (2012), 1682–1697.

[14] Y. Chen, L. Shi, Q. Feng, et al., "Artifact suppressed dictionary learning for low-dose CT image processing," *IEEE Trans Med Imaging* **33** (2014), 2271–2292.

[15] M. Aharon, M. Elad and A. Bruckstein, The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation, *IEEE Trans Signal Process* **54** (2006), 4311–4322.

[16] J. Liu, Y. Chen, Y. Hu and e. al., Low-dose CBCT reconstruction via 3D dictionary learning, *IEEE 13th Int. Symp. on Biomedical Imaging* (2016), 735–738.

[17] Z. Hu, D. Liang, D. Xia and H. Zheng, Compressive sampling in computed tomography: Method and application, *Nuclear Instruments and Methods in Physics Research A* **748** (2014), 26–32.

[18] D. Trinca and E. Libin, Performance of the sinogram-based iterative reconstruction in sparse view X-ray computed tomography, *J Xray Sci Technol* **27** (2019), 37–49.

[19] S. Singh, M.K. Kalra, M.D. Gilman, J. Hsieh, H.H. Pien, S. R., et al., Adaptive statistical iterative reconstruction technique for radiation dose reduction in chest CT: A pilot study, *Radiology* **259** (2011), 565–573.

[20] J. Ma, J. Huang, Q. Feng, et al., Lowdose computed tomography image restoration using previous normal-dose scan, *Med Phys* **38** (2011), 5713–5731.

[21] J. Bian, J.H. Siewerdsenm, X. Han, et al., Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT, *Phys Med Biol* **55** (2010), 75–99.

[22] J. Huang and et al., Projection data restoration guided non-local means for low-dose computed tomography reconstruction, *IEEE International Symposium on Biomedical Imaging* **48** (2011), 1167–1170.

[23] X. Han, J.G. Bian, E.L. Ritman, et al., Optimizationbased reconstruction of sparse images from few-view projections, *Phys Med Biol* **57** (2012), 5245–5273.

[24] I.A. Elbakri and J.A. Fessler, Statistical image reconstruction for polyenergetic x-ray computed tomography, *IEEE Trans Med Imaging* **21** (2002), 89–99.

[25] L. Ouyang, T. Solberg and J. Wang, Effects of the penalty on the penalized weighted least-squares image reconstruction for low-dose CBCT, *Phys Med Biol* **56** (2011), 5535–5552.

[26] C.H. McCollough, M.R. Bruesewitz, et al., CT dose reduction and dose management tools: Overview of available options, *Radiographics* **26** (2006), 503–512.

[27] D.T. Ginat and R. Gupta, Advances in computed tomography imaging technology, *Ann Rev Biomed Eng* **16** (2014), 431–453.

[28] Q. Xu, H. Yu, X. Mou, et al., Low-Dose X-ray CT Reconstruction via Dictionary Learning, *IEEE trans Med Imaging* **31** (2012), 1682–1697.

[29] Y. Pati, R. Rezaiifar and P. Krishnaprasad, Orthogonal Matching Pursuit: Recursive function approximation with application to wavelet decomposition, *Asilomar Conf. on Signals, Systems and Comput* 1993.

[30] J. Wang, T. Li, H. Lu and Z. Liang, Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose x-ray computed tomography, *IEEE Trans Med Imaging* **24** (2006), 1272–83.

[31] J. Jorgensen, E. Sidky and X. Pan, Quantifying admissible undersampling for sparsity-exploiting iterative image reconstruction in x-ray CT, *IEEE Trans Med Imaging* **32** (2013), 460–73.

[32] Z. Wang, A.C. Bovik, H.R. Sheikh, et al., "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process* **13** (2004), 600–612.

[33] X. Han, J. Bian, T.L. Kline, et al., Algorithm-Enabled Low-Dose Micro-CT Imaging, *IEEE Trans Med Imaging* **30** (2011), 606–620.

[34] J. Tang, B.E. Nett and G.-H. Chen, Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms, *Phys Med Biol* **54** (2009), 5781–5804.

[35] https://github.com/whutzxy/zxy