World Academy of Science, Engineering and Technology
International Journal of Electrical and Information Engineering
Vol:13, No:4, 2019

Iterative Image Reconstruction for Sparse-View Computed Tomography via Total Variation Regularization and Dictionary Learning

XianYu Zhao, JinXu Guo

***Abstract*—**Recently, low-dose computed tomography (CT) has become highly desirable due to increasing attention to the potential risks of excessive radiation. For low-dose CT imaging, ensuring image quality while reducing radiation dose is a major challenge. To facilitate low-dose CT imaging, we propose 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) to improve reconstruction performance. We call this method "PWLS-TV-DL". In order to evaluate the PWLS-TV-DL method, we performed experiments on digital phantoms and physical phantoms, respectively. The experimental results show that our method is in image quality and calculation. The efficiency is superior to other methods, which confirms the potential of its low-dose CT imaging.

***Keywords*—**Low dose computed tomography, penalized weighted least squares, total variation, dictionary learning.

I. Introduction

TODAY, X-ray CT provides clear information on the attenuation of X-rays in different tissues of the human body on a millimeter scale, thus providing rich information on human organ organization for the diagnosis and prevention of clinicians. CT has become one of the indispensable tools in the field of radiology diagnosis [1]. However, with the popularity of CT tomography in clinical diagnosis, the problem of radiation dose in CT scans has attracted more and more attention. This is because the dose of radiation in CT is accumulated for life, repeated CT scans significantly increase the probability of carcinogenesis. To reduce the radiation dose in CT examinations, various techniques have been extensively investigated. Among them, statistical iterative reconstruction (SIR) methods by modeling the measurement statistics and imaging geometry can significantly reduce radiation dose while maintaining image quality in various CT applications compared with the filtered back-projection (FBP) reconstruction algorithm [2].

In this paper, we have improved a low-dose CT statistical iterative reconstruction method. Our goal is to reconstruct a sufficiently fine image from a low-dose projection, reconstruct the intermediate image using TV minimization under the PWLS standard [3], and then use post-processing with sparse coding DL to remove residual noise and produce Clinically acceptable CT images. For simplicity, the present method is termed ‘PWLS-TV-DL’. The novelty of the PWLS-TV-DL is that, process through DL the reconstruct image could yield visually pleasant images with more continuous boundaries and less artifacts in smooth regions compared with the PWLS-TV. Qualitative and quantitative evaluations were carried out on the digital phantoms in terms of accuracy and resolution properties.

XianYu Zhao is with the College of Information Engineering, Wuhan University of Technology, China (e-mail: 18684690389@163.com).

JinXu Guo is with the College of Information Engineering, Wuhan University of Technology and Key Laboratory of Fiber Optic Sensing Technology and Information Processing (Wuhan University of Technology), Ministry of Education, China.

II. Methods

A. PWLS Criteria for CT Image Reconstruction

The PWLS approach for iterative reconstruction of X-ray CT images has been studied by Herman and Sauer and Bouman [4]. On the basis of the noise properties of CT projection data, the PWLS criterion for CT image reconstruction can be rewritten as follows:

$$\mu^* = \arg\min_{\mu\geq 0}\{(x - H\mu)^T\Sigma^{-1}(x - H\mu) + \beta R(\mu)\} \qquad (1)$$

where x represents the obtained sinogram data, $x = (x_1, x_2, \dots, x_M)^T$,μ is the vector of attenuation coefficients to be reconstructed, i.e., $\mu = (\mu_1, \mu_2, \dots, \mu_N)^T$, where T denotes the matrix transpose. The operator H represents the system matrix with the size of $M \times N$. The element of H is the length of the intersection of projection ray i with pixel j.Σ is a diagonal matrix with the ith element of σ_i^2 which is the variance of sinogram data x_i. $R(\mu)$ represents a prior term, and β is a hyper-parameter for controlling the strength of prior term as a penalty. The goal for CT image reconstruction is to estimate the attenuation coefficients m 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:

$$\sigma_i^2 = \frac{1}{I_0}\exp(\overline{x_i})(\frac{1}{I_0}\exp(\overline{x_i})(\sigma_e^2 - 1.25)) \qquad (2)$$

where I_0 denotes the incident x-ray intensity, $\overline{x_i}$ is the mean of the sinogram data at bin i and σ_e^2 is the background electronic noise variance.

B. TV Minimization

The TV was first proposed by Rudin in the image denoising model [5], which is used to measure image characteristics up to a certain order of differentiation. Mathematically, the original TV of an image u can be defined as follows:

$$\mathrm{TV}(x) = \int_{\Omega} |\nabla x|\, dx \tag{3}$$

where Ω is a bounded domain, and ∇x is the gradient of image x. Another definition of the TV can be written as follows:

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

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, could promote the performance of iterative CT image reconstruction from noisy or sparse-view projection measurements.

C. DL-Based Image Reconstruction

The DL method has the advantage of adaptability, so that the dictionary can be approximated to the maximum initial signal. An efficient dictionary should also be characterized by multi-scale, geometric invariance and redundancy. An adaptive dictionary with the above advantages is beneficial for sparse representation of the signal, because the richer the features of the dictionary, the closer it is to the signal to be processed, the more likely it is to approximate the signal 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 training sample images can be sparsely represented by the dictionary. To do this, you can construct an objective function:

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

where X is the matrix of training samples, D_0 is the dictionary to be sought, and A is the sparse coefficient matrix in which Y is represented by D_0. $||\cdot||_F^2$ 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 update of the dictionary is updated one by one for each atom.

In summary, the CT reconstruction problem can be stated as:

$$\min\{_{x,a_j} ||x||_{TV} + \textstyle\sum_j ||D_0 a_j - R_j x||_2^2\}\; subject\ to\ ||a_j||_0 \leq \rho\, \forall j, x \geq 0, ||Mx - y||_2^2 < \varepsilon_a \tag{6}$$

where $||x||_{TV}$ is the TV norm of an image x, $||a_j||_0$ is the l_0 norm of a_j, ρ is the threshold of sparsity, M is a system matrix describing the forward projection, y is a measured dataset and ε_a is a small positive value representing the error threshold.

D. PWLS–TV–DL Algorithm

In this section, we will detail the implementation details of the PWLS-TV-DL algorithm.

The PWLS-TV-DL algorithm consists of PWLS reconstruction, TV minimization and DL.PWLS works as the reconstruction algorithm, both TV minimization and DL work 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, the introduction of adaptively learned dictionary alleviates the artifacts caused by the piecewise constant assumption and allows accurate restoration of images with complex structures.

The size of the gradient in TV minimization can be approximated as

$$\tau_{i,j} \approx \sqrt{(x_{i,j} - x_{i-1,j})^2 + (x_{i,j} - x_{i,j-1})^2} \tag{7}$$

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

$$\nabla_{x_{i,j}} ||x||_{TV} = \frac{\partial ||x||_{TV}}{\partial x_{i,j}} \approx \frac{(x_{i,j} - x_{i-1,j}) + (x_{i,j} - x_{i,j-1})}{\tau_{i,j} + \varepsilon} - \frac{(x_{i+1,j} - x_{i,j})}{\tau_{i+1,j} + \varepsilon} - \frac{(x_{i,j+1} - x_{i,j})}{\tau_{i,j+1} + \varepsilon} \tag{8}$$

where ε is a positive number in the denominator. And TV minimization can be stated as follows:

$$x_k = x_{k-1} - \beta \cdot \Delta x \cdot \frac{\nabla ||x_{k-1}||_{TV}}{||\nabla ||x_{k-1}||_{TV}||} \tag{9}$$

where β is the length of each gradient-descent step and q is the iteration index.

The DL process includes atom matching and image updating. In the atom matching, we need to find the sparse representation (SR) of each patch in the target image by the orthogonal matching pursuit (OMP) algorithm. Because our dictionary is relatively large, searching the entire dictionary is impractical. However, we can get it 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

$$\min_{w_j} ||R_j x - D_0 w_j||_2^2 \; subject\ to\ ||w_j||_0 \leq \rho \tag{10}$$

The OMP algorithm stops when $||R_j x - D_0 w_j||_2^2 < \varepsilon_b$ or the maximum iteration number for OMP is reached, 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 $x^j = D_0 w_j$, and then the updated image patches are recorded in a matrix. They are not written back to the target image until all image patches have been updated and recorded. Finally, the image x can be updated as

$$x = (\textstyle\sum_j R_j^T R_j)^{-1} \sum_j R_j^T x^j \tag{11}$$

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

TABLE I
WORKFLOW FOR 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...K:
1. Reconstruct an image by PWLS algorithm

$$\mu^* = \underset{\mu \geq 0}{\arg\min}\{(x - H\mu)^T \Sigma^{-1}(x - H\mu) + \beta R(\mu)\}$$

2. TV-minimization loop
2.1. Initialization: $\Delta x = ||\overline{x_0} - x_k||$;
2.2. TV gradient descent, for k=1,2...K

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

3. For each patch $R_j\overline{x_Q}$ in the image,
(a) find the set $\widetilde{D_0}$ of S nearest atoms in D_0;
(b) compute the weights using OMP so that

$$||R_j x - D_0 w_j||_2^2 < \varepsilon_b,$$

or the maximum iteration number for OMP is reached;
(c) estimate the image patch x_j as

$$x^j = D_0 w_j;$$

4. Update the image simultaneously:

$$x_{k+1} = (\textstyle\sum_j R_j^T R_j)^{-1} \sum_j R_j^T x^j;$$

Repeat beginning with Step 1 until the stopping criteria are satisfied.

III. EXPERIMENTS AND RESULTS

A. Experimental Setup

To evaluate the performance of the PWLS-TV-DL method in CT image reconstruction, we conducted experiments on the digital XCAT phantom.

B. Digital XCAT Phantom

Fig. 1 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 the previous studies (Wang et al., 2006) [6], we first simulated the noise-free sonogram data $\hat{y}$ then generated the noisy transmission measurement I according to the statistical model of the pre-logarithm projection data, that is,

$$y_i = \mathrm{Poisson}(b_i \exp(-\hat{y})) + \mathrm{Normal}(0, \sigma_e^2) \quad (12)$$

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 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 under-sampling the 1,160 views of normal-dose simulation to only 360 views evenly over 2π.

Fig. 1 Digital phantoms used in the studies: a slice of digital XCAT phantom

C. Performance Evaluation on Digital Phantom

In the digital phantom study, the original phantom data were directly used as the ground-truth image. As mentioned before, low-dose CT can be implemented by lowering tube current or reducing projection views. To have a more comprehensive study, we tested our method in both a low-current case and a few-view case.

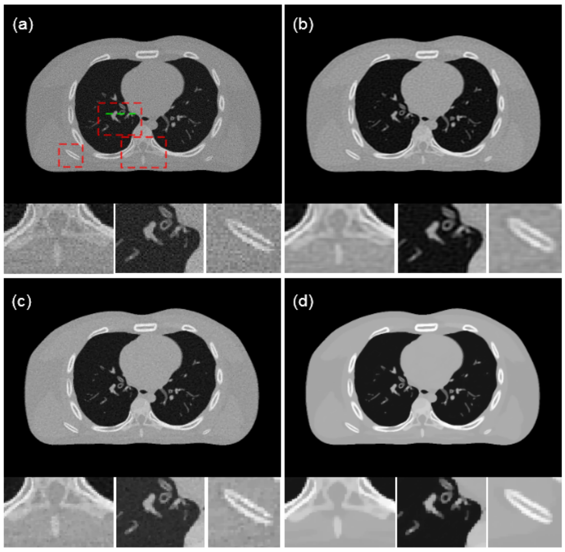

Fig. 2 Imaging results of different methods on digital XCAT: (a) PWLS; (b) PWLS-DL; (c) PWLS-TV; (d) PWLS-TV-DL

In the low-current case, CT scan dose was set to b_i=1.0×10^5 photons per ray and 360 views over 360° were simulated. The imaging results of different methods are shown in Fig. 2. It can be seen that the PWLS result in Fig. 2 (a) contains heavy noise in the whole reconstructed region. Even worse, some small structures are almost covered by noise, which will cause misdiagnosis in the clinic. Fig. 2 (b) shows the image reconstructed by PWLS-DL, which is better than PWLS, but it is still covered by some noise. As shown in Fig. 2 (c), the image becomes much cleaner after being denoised by TV, but the edges in the image are seriously blurred compared to that in the true image in Fig. 1. In contrast, it is easy to see in Fig. 2 (d), the image reconstructed by PWLS-TV-DL, which effectively suppresses noise and artifacts.

The quantitative assessment was carried out by calculating

PSNR, SSIM, and RMSE between the true image and reconstructed images. Table II lists the performance comparison of different methods. The following columns of Table II demonstrate that whether we consider PSNR, SSIM, or RMSE, the method of PWLS-TV-DL obtains a good result.

TABLE II
NUMERIC RESULTS ON XCAT

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	9.704	6.835	4.171	**1.576**

The profile images and residual images were compared in Figs. 3 and 4, respectively. The profiles located at the pixel positions x from 350 to 410 and y = 350. It is not difficult to find that the PWLS-TV-DL curve is closer to the Phantom curve. The results show that the PWLS-TV-DL method can help achieve image quality superior to that of the other comparison methods.

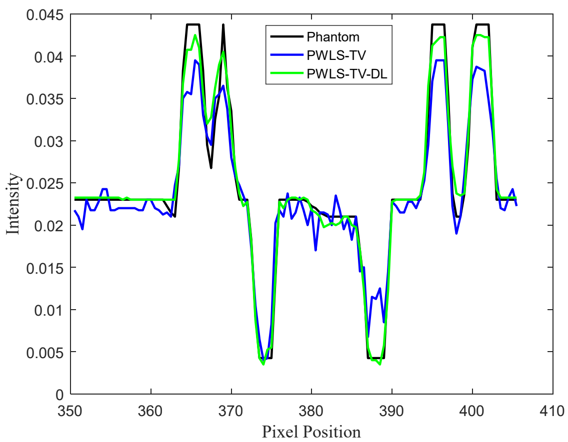

Fig. 3 The profile images: (a) Phantom in black; (c) PWLS-TV in blue; (d) PWLS-TV-DL in green

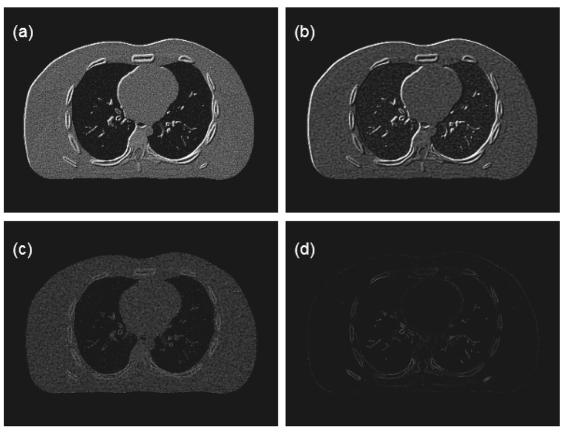

Fig. 4 Residual images of the reconstructed results based on the PWLS method, the PWLS-DL method, the PWLS-TV method, and the PWLS-TV-DL method with 200 iterations for the simulation XCAT data. All images are displayed in the same window

IV. DISCUSSION AND CONCLUSION

In this paper, based on the PWLS standard, we propose a new low-dose CT reconstruction solution by combining TV minimization and sparse DL. The intermediate image is reconstructed using TV minimization and then post-processed using DL to remove residual noise and produce a clinically acceptable CT image. It can be seen from simulation experiments that compared with reconstruction methods such as PWLS, PWLS-DL and PWLS-TV, this method can improve the quality of reconstructed images and produce smaller RMSE and larger PSNR and SSIM values. However, the main shortcoming of the PWLS-TV-DL algorithm is that the update of the matrix in DL increases the computational burden and requires a long running time. To solve this problem, a fast computer and dedicated hardware are needed. We believe that in the future, most iterative-based image reconstructions including the PWLS-TV-DL algorithm can be widely used in medical clinics.

REFERENCES

[1] Yan, H., Cervino, L., Jia, and X., "A comprehensive study on the relationship between the image quality and imaging dose in low-dose cone beam CT," Phys. Med.Biol., vol. 57, pp. 2063-2080, 2012.
[2] Zhanli.Hu, Yunwan.Zhang, Jianbo.Liu, Jianhua.Ma, Hairong.Zheng, and Dong.Liang, "A feature refinement approach for statistical interior CT reconstruction," Physics in Medicine and Biology, vol. 61, pp. 5311-5334, 2016.
[3] 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., vol. 56, pp. 5535-5552, 2011.
[4] C. H. McCollough, M. R. Bruesewitz, and J. M. K. Jr, "CT dose reduction and dose management tools: overview of available options," Radiographics, vol. 26, pp. 503-512, 2006.
[5] Ginat, D. T., Gupta, and R., "Advances in computed tomography imaging technology," Ann. Rev. Biomed. Eng., vol. 16, pp. 431-453, 2014.
[6] W. J., L. T., L. H., and L. Z., "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose x-ray computed tomography," IEEE Trans. Med. Imag., vol. 24, pp. 1272-83, 2006.