

CBCT reconstruction via a penalty combining total variation and its higher-degree term

Nanbo Sun¹, Tao Sun¹, Jing Wang², and Shan Tan¹

¹Key Laboratory of Image Processing and Intelligent Control of Ministry of Education of China, School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China;

²Department of Radiation Oncology, The University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75390, USA

ABSTRACT

Penalized weighted least-squares (PWLS) iterative algorithm with a total variation penalty (PWLS-TV) has shown potential to improve cone-beam CT (CBCT) image quality, particularly in suppressing noise and preserving edges. However, it sometimes suffers from the well-known staircase effect, which produces piece-wise constant areas in images. In order to remove the staircase effect, there is an increasing interest in replacing TV by higher-order derivative operations such as Hessian. Unfortunately, Hessian tends to blur the edges in the reconstruction results. In this study, we proposed a new penalty, namely the TV-H penalty, which combines the TV penalty and the Hessian penalty for CBCT reconstruction. The TV-H penalty retains some of the most favorable properties of the TV penalty like suppressing noise and preserving edges and has a better ability in preserving the structures of gradual intensity transition in images. The penalized weighted least-squares (PWLS) criterion with the majorization-minimization (MM) approach was used to minimize the objective function. Two simulated digital phantoms were used to compare the performance of TV, Hessian penalty and TV-H penalties. Our experiments indicated that the TV-H penalty outperformed the TV penalty and the Hessian penalty.

Keywords: cone-beam CT, penalized weighted least-squares, total variation, Hessian norm

1. INTRODUCTION

On-board cone-beam computed tomography (CBCT) is extensively used in clinics in image-guided radiation therapy nowadays [1]. However, a repeated exposure to high x -ray radiation dose during the CBCT scanning brings the patient a lifetime risk of cancer [2], and increases the probability of stochastic risk of genetic defects as well [3].

A main method to reduce radiation dose is to lower the mAs level in CBCT acquisition. The conventional FDK method with lower mAs level [4] raises degraded reconstructed CBCT images. Recently, several algorithms have been proposed to improve CBCT imaging quality. Statistical iterative reconstruction (SIR) algorithms have shown a superior performance [5] over the FDK method. Algorithms based on compressed sensing have also been applied to the CBCT reconstruction problem [6-9]. The total variation (TV) methods [10, 11] have demonstrated their power in edge preservation and noise suppressing. Projections on convex sets (POCS) iterative algorithm has been developed that leads to encouraging reconstruction results [12]. The penalized weighted least-squares (PWLS) iterative algorithm with the TV penalty (PWLS-TV) has shown the state-of-the-art performance in CBCT reconstruction [5]. Although the TV penalty is successful in many aspects, sometimes it produces the annoying staircase effect [13, 14]. The TV penalty tends to penalize the image gradient irrespective of the image structures, leading to piecewise-constant results.

The TV penalty has also been applied to image restoration or denoising field. In this field, higher-order derivative operations have been used to replace the TV penalty for reducing the staircase effect [13]. Having a weaker penalty on the absolute difference between neighboring pixels, the higher-order derivatives allow for piecewise-smooth restored results, and have shown improved performance in suppressing the staircase effect. To prevent the staircase effect of TV penalty in CBCT, we previously proposed to use a second-order derivative penalty — the Hessian penalty — for PWLS CBCT reconstruction [15]. However, the Hessian penalty introduces another negative effect, which tends to blur edges in reconstruction results.

In this study, we proposed a new TV-H penalty, which combines TV with Hessian by an exponential term, for CBCT reconstruction. The proposed TV-H penalty has a better ability in preserving both sharp edges and areas with gradual intensity transition. We developed an effective algorithm to minimize the objective function based on the penalized weighted least-square (PWLS) criterion [16]. The objective function was upper-bounded by a quadratic majorizer and minimized with the Gauss-Seidel update strategy. We conducted a systematic comparison among the TV, Hessian and TV-H for CBCT reconstruction on two simulated phantoms. The results indicated that the TV-H penalty not only outperformed the TV penalty in suppressing the staircase effect, but also outperformed the Hessian penalty in preventing blurring edges.

2. METHODS

2.1 PWLS Reconstruction

In CBCT imaging, the line integral of attenuation coefficients can be calculated through the logarithm transform of the ratio of incident and detected photon numbers

$$p_i = \ln \frac{N_{i0}}{N_i}, \quad (1)$$

where N_{i0} and N_i are the incident photon number and detected number at the detector bin i respectively. In the following of this paper, we refer the value of p_i as the sinogram datum at the detector bin i .

Mathematically, the statistical noise property of x-ray CT projection data after logarithm transform approximately follows a Gaussian distribution [17]. The variance of the sinogram datum can be estimated by

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

where \bar{p}_i and σ_i^2 are the mean and variance of projection datum p_i at detector bin i .

To calculate the sinogram datum variance at the detector bin i via Eq. (2), we need to estimate the incident photon number N_{i0} for calculation of the sinogram variance. Ideally, the incident x-ray flux from the tube would be calibrated as uniform as possible across a field of view (FOV), i.e. N_{i0} is a constant for all the detector bins [5].

Because the projection data p_i follows the Gaussian distribution, the PWLS cost function in the sinogram domain can be written as [18]

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

The first term in Eq. (3) is a data-fidelity term. p is the vector of log-transformed projection measurements. A represents the projection matrix. μ is the attenuation map and Σ is a diagonal matrix with its i th element σ_i^2 . The symbol T denotes the transpose operator. The second term is a penalty term, where β is a parameter controlling the trade-off between the data-fidelity term and the penalty term.

2.2 Penalty Terms in PWLS

The total variation (TV) penalty consists of the first-order derivative of the reconstructed images. Despite the fact that the TV penalty in the CBCT reconstruction has shown good properties in noise suppressing and edge preserving [17], it sometimes leads to the well-known staircase effect. We extended from the first-order derivative to the second-order derivative and defined the Hessian-based norm penalty in our previous work [15].

Though the Hessian penalty can overcome the staircase effect of the TV penalty, it tends to reduce the resolution of the image and blur the edges of an object. The TV penalty, in contrast, can preserve the edges. The intuitive idea in this study was to combine the two penalties together. The combined penalty was defined as TV-H penalty as follows:

$$R_{TV-H}(\mu) = \int_{\Omega} \|\nabla \mu\|_2 d\mathbf{x} + \int_{\Omega} \lambda \|H_{\mu}(\mathbf{x})\|_F d\mathbf{x}, \quad (4)$$

where $\Omega \subset R^3$, $\nabla \mu = (\mu_x, \mu_y, \mu_z)^T$. The first term in Eq. (4) is the TV penalty and the second term in Eq. (4) is a modified Hessian penalty, where

$$H_{\mu} = \begin{pmatrix} \mu_{xx} & \mu_{xy} & \mu_{xz} \\ \mu_{yx} & \mu_{yy} & \mu_{yz} \\ \mu_{zx} & \mu_{zy} & \mu_{zz} \end{pmatrix}, \quad (5)$$

and the symbol $\|\cdot\|_F$ denotes the Frobenius norm of a matrix, and λ is the ratio which balances the TV and Hessian penalty.

In discrete case, let us denote the index of a voxel by the triplet (x, y, z) and all possible filters yielding second-order derivatives as $\{L_i(x, y, z); i = 1, \dots, 6\}$. The TV-H penalty in Eq. (4) can be re-written as

$$R_{TV-H} = \sum_{x,y,z} \left\{ \sqrt{\mu_x^2(x, y, z) + \mu_y^2(x, y, z) + \mu_z^2(x, y, z)} + \lambda \sqrt{\sum_{i=1}^6 [L_i(x, y, z) * \mu(x, y, z)]^2} \right\}. \quad (6)$$

2.3 Implementation

The PWLS-TV reconstruction algorithm [5] can be performed by the Gauss-Seidel update strategy. To adopt the Gauss-Seidel update strategy, the objective function Eq. (3) must be quadratic. However, the proposed TV-H penalty is not quadratic. To minimize the object function we used the majorization-minimization (MM) approach, which can transform the TV-H penalty into a series of quadratic terms [19, 20], thus allowing to use the Gauss-Seidel update strategy. The main idea of the approach is to obtain a series of quadratic functions that upper-bound the original target function.

According to the MM approach, we developed a quadratic majorizer of the TV-H penalty in Eq. (6). We employed the inequality holding for a general function

$$\sqrt{g(x)} \leq \frac{\sqrt{g(y)}}{2} + \frac{g(x)}{2\sqrt{g(y)}}, \quad \forall x, \forall y: g(y) \neq 0, \quad (7)$$

with equality if and only if $g(x) = g(y)$. Using the fact that the inequality relation is closed under the formation of sums and non-negative products, we then obtained that the functions

$$\sqrt{\mu_x^2(x, y, z) + \mu_y^2(x, y, z) + \mu_z^2(x, y, z)} \leq \left\{ \frac{1}{2} \sqrt{[\mu_x^{(t)}(x, y, z)]^2 + [\mu_y^{(t)}(x, y, z)]^2 + [\mu_z^{(t)}(x, y, z)]^2} + \frac{\mu_x^2(x, y, z) + \mu_y^2(x, y, z) + \mu_z^2(x, y, z)}{2\sqrt{[\mu_x^{(t)}(x, y, z)]^2 + [\mu_y^{(t)}(x, y, z)]^2 + [\mu_z^{(t)}(x, y, z)]^2}} \right\}, \quad (8)$$

$$\sqrt{\sum_{i=1}^6 [L_i(x, y, z) * \mu(x, y, z)]^2} \leq \frac{1}{2} \sqrt{\sum_{i=1}^6 [L_i * \mu^{(t)}(x, y, z)]^2} + \frac{\sum_{i=1}^6 [L_i * \mu(x, y, z)]^2}{2\sqrt{\sum_{i=1}^6 [L_i * \mu^{(t)}(x, y, z)]^2}}. \quad (9)$$

Substituting the Eq. (8) and Eq. (9) into Eq. (6), the

$$Q_{R_{TV-H}}(\mu, \mu^{(t)}) = \sum_{x,y,z} \left\{ \begin{aligned} & \frac{1}{2} \sqrt{[\mu_x^{(t)}(x,y,z)]^2 + [\mu_y^{(t)}(x,y,z)]^2 + [\mu_z^{(t)}(x,y,z)]^2} \\ & + \frac{\mu_x^2(x,y,z) + \mu_y^2(x,y,z) + \mu_z^2(x,y,z)}{2 \sqrt{[\mu_x^{(t)}(x,y,z)]^2 + [\mu_y^{(t)}(x,y,z)]^2 + [\mu_z^{(t)}(x,y,z)]^2}} \\ & + \lambda \left[\frac{1}{2} \sqrt{\sum_{i=1}^6 [L_i * \mu^{(t)}(x,y,z)]^2} + \frac{\sum_{i=1}^6 [L_i * \mu(x,y,z)]^2}{2 \sqrt{\sum_{i=1}^6 [L_i * \mu^{(t)}(x,y,z)]^2}} \right] \end{aligned} \right\} = \frac{1}{2} \mu^T (R_{TV}^{(t)} + \lambda R_{Hessian}^{(t)}) \mu + const \quad (10)$$

is a valid estimation of $R_{TV-H}(\mu)$ at fixed point $\mu^{(t)}$, where $R_{TV}^{(t)}$ and $R_{Hessian}^{(t)}$ are asymmetric square matrixes, and the constants are independent of μ , but depend only upon $\mu^{(t)}$.

The majorizers in Eq. (10) can be equivalently stated as [5]

$$Q_{R_{TV-H}}(\mu, \mu^{(t)}) = \frac{1}{2} \sum_j \sum_{i \in N_j} (w_{TV_j} + \lambda w_{Hessian_j}) (\mu_j - \mu_i)^2 + const, \quad (11)$$

where N_j is the set of neighbor voxels of voxel j and $w_{TV_j} + \lambda w_{Hessian_j}$ is determined by $\mu^{(t)}$.

Now, the key problem is how to preserve the good properties of both penalties. To take the advantages of the TV penalty, which preserves the edges, and to take the advantages of the Hessian penalty, which preserves the gradual transition regions of the image, the λ is better to be an adapted variable. In [16], Wang added an exponential term to an isotropic penalty, leading to an anisotropic penalty as follows:

$$R(\mu) = \frac{1}{2} \sum_j \sum_{m \in N_j} \exp \left[- \left(\frac{\mu_j - \mu_m}{\delta} \right)^2 \right] w_{jm} (\mu_j - \mu_m)^2 \quad (12)$$

where N_j is the set of neighbor voxels of voxel j . The parameter w_{jm} was set to 1 for first-order neighbors and $1/\sqrt{2}$ for second-order neighbors. The μ in the exponential term changes at each iteration. The results showed that the anisotropic penalty has a better ability in preserving the edges than the isotropic penalty, which indicated that the exponential term plays a role in detecting image edges in the process of iteration.

In this study, we also adopted the exponential term to detect image edges so that one can use the TV penalty when it comes to the sharp edges, and the Hessian penalty when it comes to gradual transition regions. After replacing λ with an exponential term, Eq. (11) can be expressed as

$$Q_{R_{TV-H}}(\mu, \mu^{(t)}) = \frac{1}{2} \sum_j \sum_{i \in N_j} \left(w_{TV_j} + \alpha \exp \left(- \frac{[\mu_x^{(t)}(x,y,z)]^2 + [\mu_y^{(t)}(x,y,z)]^2 + [\mu_z^{(t)}(x,y,z)]^2}{\sigma^2} \right) w_{Hessian_j} \right) (\mu_j - \mu_i)^2 + const, \quad (13)$$

where α is a constant, σ determines the strength of Hessian penalty, $\mu_x^{(t)}, \mu_y^{(t)}, \mu_z^{(t)}$ represents the derivation of μ in x, y, z direction at the t iteration. When it comes to image edges, the difference between two pixels is large. So, $\nabla \mu^{(t)}$ is large, and thus the exponential term is small. In this way, the TV penalty plays a major role at edges, which has an advantage to preserve the edges effectively. By contrast, when it comes to the uniform regions or the gradual transition regions, the difference between two pixels is small. Thus, $\nabla \mu^{(t)}$ is small, which means that the exponential term is large. Therefore, at the uniform or the gradual transition regions the Hessian penalty plays an important role as the TV penalty does at edges, which is able to suppress the staircase effect of the TV penalty.

When α increases, the second equation plays a major role, leading to a better performance in gradual transition regions but blurring edges at the same time. In this study, we choose different α for different images: α was set to 1 at first; if the result was blurring, we set it to 1/2, 1/4 and so on. As for the δ , we set the value of δ to be 90% of the histogram of the gradient magnitude of the initial FDK reconstructed image (which was used as the initial during iterative reconstruction).

Substituting Eq. (13) into Eq. (3), we can get the following equation:

$$\Phi(\mu) = \frac{1}{2}(p - A\mu)^T \Sigma^{-1}(p - A\mu) + \beta \left(\frac{1}{2} \mu^T \mathbf{R}_{TV-H}^{(i)} \mu \right) \quad (14)$$

where $\mathbf{R}_{TV-H}^{(i)}$ is a asymmetric matrix.

Minimizing the objective function can be achieved by solving the following equation

$$\frac{\partial \Phi(\mu)}{\partial \mu} = 0. \quad (15)$$

The minimization of the resulting majorizers is equivalent to solving the following series of liner equations

$$\left(A^T \Sigma^{-1} A + \beta \mathbf{R}_{TV-H}^{(i)} \right) \mu = A^T \Sigma^{-1} p. \quad (16)$$

Eq. (16) can be efficiently solved by the Gauss-Seidel update strategy [16].

The method to calculate the projection matrix A is another important part in the implement of the algorithm. Different methods can affect the efficiency and accuracy of the whole algorithm. The Siddon's algorithm [21] is a popular one in the implementation of several SIR CBCT reconstruction methods in the literature [5, 22]. However, this line intersection model is imprecise. Therefore, we adopted a more precise and more efficient forward projection algorithm, Separable Footprints (SF) [23], instead. The method approximates the voxel footprint functions as 2D separable functions. Specifically we chose SF-TR projector which uses trapezoid functions in the transaxial direction and rectangular functions in the axial direction.

3. EXPERIMENTS

We tested our reconstruction algorithm on two cases: a Compressed Sensing (CS) phantom and a modified 3D Shepp-Logan phantom.

The two digital phantoms were forward projected mathematically using the SF forward projection algorithm. The total angle samples over 360° was 360. The dimensions of each projection image were 800×200 pixels with the pixel size of $0.776 \times 0.776 \text{ mm}^2$. The distance of source-to-axial distance is 100 cm and the source-to-detector distance is 150 cm. Randomly generated Gaussian noises were added the projections according to the noise model of Eq. (2). The incident photon number in Eq. (2) was set to 5×10^3 , 1×10^4 and 5×10^4 in order to simulate low-dose, medium-dose and high-dose protocol, respectively. The simulated projection data at different noise level was used for reconstruction. The reconstructed images had dimensions of $350 \times 350 \times 16$ voxels, where the voxel size is $0.776 \times 0.776 \times 0.776 \text{ mm}^3$.

3.1 Evaluation Indexes

For each group experiments with different incident photon numbers, the trade-off parameter β in Eq. (3) were chosen so that the iterative reconstructed images with three penalties were at same noise level. We calculated the structural similarity (SSIM) [24] and the full width at half maximum (FWHM) of a selected region of interest (ROI) and the whole image. Higher SSIM means the same region of two images is more similar. In other words, the reconstructed image is more consistent with the original image. The FWHM can be calculated using the fitted normal distribution of the gradient of an object's edge. Lower FWHM means that the object's edge is sharper.

3.2 CS Phantom

We first performed a simulation study using the CS phantom. Figure 1(a) shows one typical slice of the CS phantom. The phantom has a uniform background with attenuation coefficient of 0.0125 mm^{-1} . It contains an octahedral with linearly gradual transitioning intensity. Figure 1(b) plots the horizontal profile through the center of the octahedral along the shown slice. It also contains a set of line objects for studying the image resolution, a set of circular cylinders with different diameters and intensity and a ball with linearly gradual transitioning intensity similar to the octahedral. The noise level was characterized by the mean standard derivations of 5 uniform regions with size of 30×30 pixels (indicated by the black squares in Fig. 1(a)). We selected a ROI of size 76×76 pixels (indicated by a white square in Fig. 1(a)) at the octahedral that we were particularly interested in.

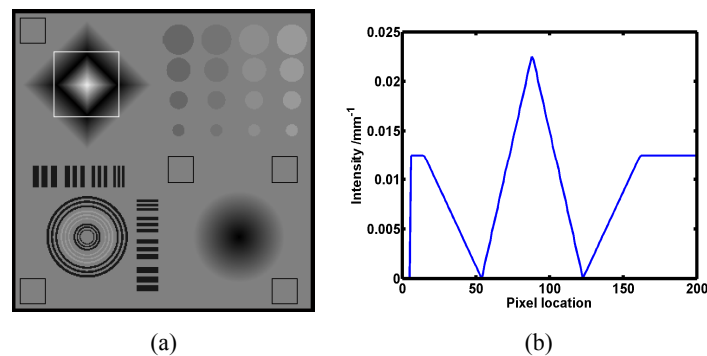
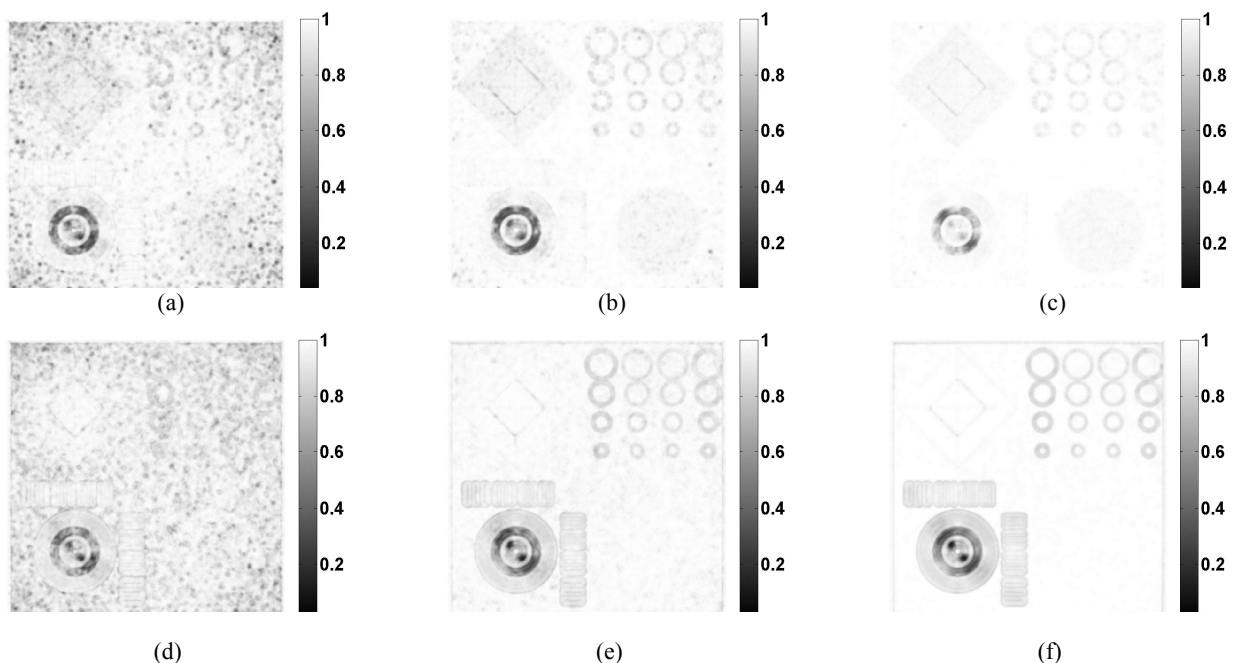


Figure 1. (a) One representative slice of the CS phantom (b) horizontal profile through the center of the octahedral at the shown slice

Figure 2 shows the *SSIM* map images between the iterative reconstructed images and the original image of the reconstruction algorithms with TV, Hessian and TV-H penalties. It is obvious that the regions near the octahedral and the ball in the map images reconstructed using TV-H and Hessian have higher values than those reconstructed using TV, indicating a better preservation of the gradual transition regions. In contrast, the annular and strip-type regions reconstructed using the TV-H penalty and the TV penalty have higher values than those of the Hessian penalty, implying a better preservation of the edges. The results indicated that the TV-H penalty indeed took the advantages of both the TV penalty and the Hessian penalty.



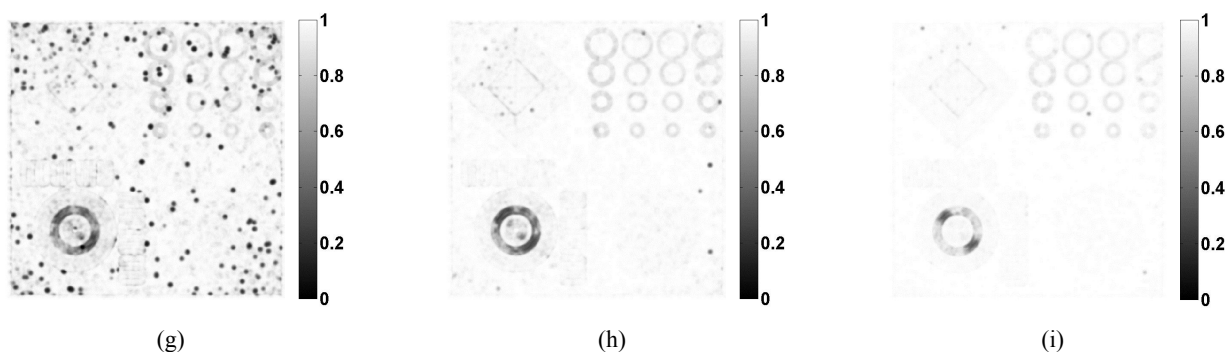


Figure 2. Comparison of *SSIM* maps of the CS phantom reconstructed using PWLS iterative reconstruction algorithms with different penalties and the original image at different noise levels. (a) high noise projection data using TV; (b) medium noise projection data using TV; (c) low noise projection data using TV; (d) high noises projection data using Hessian; (e) medium noises projection data using Hessian; (f) low noises projection data using Hessian; (g) high noises projection data using TV-H; (h) medium noises projection data using TV-H; (i) low noises projection data using TV-H

3.3 Modified 3D Shepp-Logan Phantom

For the modified 3D Shepp-Logan phantom, we made some modification on the original Shepp-Logan phantom by substituting the constant ellipsoid with a linearly transitioning ellipsoid. Our modification was also based on the consideration that we were focused on the preservation of structure of transitioning in the image.

Figure 3 shows a representative slice of the modified 3D Shepp-Logan phantom obtained by reconstruction algorithms with different penalties. Fig. 3(a) shows the original image, Fig. 3(b) shows the analytical FDK reconstructed image, Fig. 3(c) shows the reconstructed image using TV penalty, Fig. 3(d) shows the reconstructed image using Hessian penalty, and Fig. 3 (e) shows the reconstructed image using TV-H penalty. For better illustration, a ROI (indicated by white square in Fig. 3(a)) at the ellipsoid was enlarged for each image. The display window was set to 0-0.05mm⁻¹. The reconstructed image using the TV-H penalty and the Hessian penalty is smoother while the reconstructed image using the TV penalty shows the obvious staircase effect.

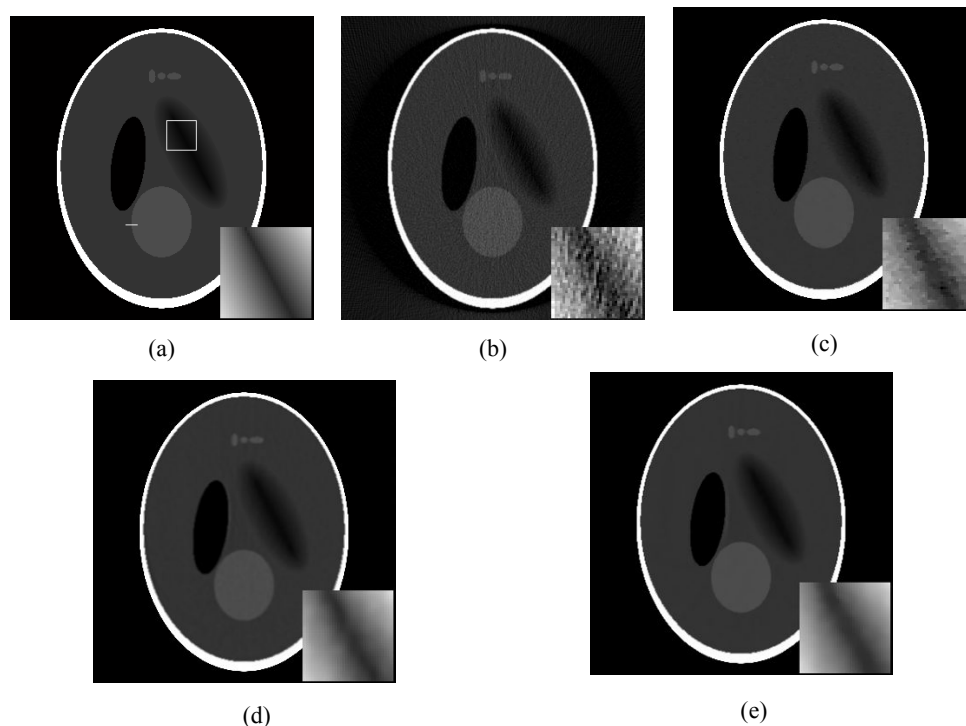


Figure (3). Representative slices of the modified 3D Shepp-Logan phantom (a) original image; (b) FDK reconstructed image; (c) PWLS reconstructed image using TV; (d) PWLS reconstructed image using Hessian; (e) PWLS reconstructed image using TV-H

As shown in Fig. 4, to measure the blurring degree of the edges, we got the derivation of the white line in Fig. 3 (a), and then fitted it with the Gaussian function and obtained the FWHM. The FWHM of TV-H penalty was 1.59, which was close to that of TV (FWHM=1.61) and smaller than that of Hessian(FWHM=3.16). The results showed that the TV-H penalty can preserve the edges as the TV penalty.

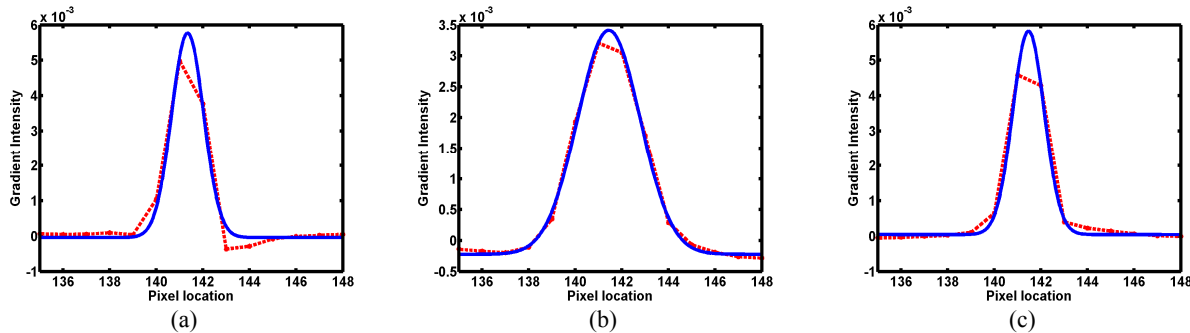


Figure 4. Profiles (Red curves) of the gradient of white line in Fig. 3, fitted with the Gaussian function (Blue curves) to computed FWHM. (a) PWLS reconstruction using TV (FWHM=1.61); (b) PWLS reconstruction using Hessian (FWHM=3.16); (c) PWLS reconstruction using TV-H (FWHM=1.59)

4. CONCLUSION

In this study, we proposed a new TV-H penalty for CBCT reconstruction, and developed an effective algorithm to minimize the objective function using the majorization-minimization strategy. Comparison studies indicated that the proposed TV-H penalty not only outperformed the TV penalty in suppressing the staircase effect, but also outperformed the Hessian penalty in preventing blurring edges.

Acknowledgement

This work was supported in part by National Natural Science Foundation of China (NNSFC), under Grant Nos. 60971112 and 61375018, and Fundamental Research Funds for the Central Universities, under Grant No. 2012QN086. J. Wang was supported in part by grants from the Cancer Prevention and Research Institute of Texas (RP130109 and RP110562-P2) and a grant from the American Cancer Society (RSG-13-326-01-CCE).

REFERENCES

- [1] D. A. Jaffray, J. H. Siewerdsen, J. W. Wong, and A. A. Martinez, "Flat-panel cone-beam computed tomography for image-guided radiation therapy," *International Journal of Radiation Oncology* Biology* Physics*, vol. 53, pp. 1337-1349, 2002.
- [2] A. B. de González, M. Mahesh, K.-P. Kim, M. Bhargavan, R. Lewis, F. Mettler, *et al.*, "Projected cancer risks from computed tomographic scans performed in the United States in 2007," *Archives of internal medicine*, vol. 169, pp. 2071-2077, 2009.
- [3] N. Wen, H. Guan, R. Hammoud, D. Pradhan, T. Nurushev, S. Li, *et al.*, "Dose delivered from Varian's CBCT to patients receiving IMRT for prostate cancer," *Physics in medicine and biology*, vol. 52, p. 2267, 2007.

- [4] L. Feldkamp, L. Davis, and J. Kress, "Practical cone-beam algorithm," *JOSA A*, vol. 1, pp. 612-619, 1984.
- [5] L. Ouyang, T. Solberg, and J. Wang, "Effects of the penalty on the penalized weighted least-squares image reconstruction for low-dose CBCT," *Physics in medicine and biology*, vol. 56, p. 5535, 2011.
- [6] K. Choi, J. Wang, L. Zhu, T.-S. Suh, S. Boyd, and L. Xing, "Compressed sensing based cone-beam computed tomography reconstruction with a first-order method," *Medical physics*, vol. 37, pp. 5113-5125, 2010.
- [7] H. Lee, L. Xing, R. Davidi, R. Li, J. Qian, and R. Lee, "Improved compressed sensing-based cone-beam CT reconstruction using adaptive prior image constraints," *Physics in medicine and biology*, vol. 57, p. 2287, 2012.
- [8] S. Lefkimmiatis, A. Bourquard, and M. Unser, "Hessian-based regularization for 3-D microscopy image restoration," in *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, 2012, pp. 1731-1734.
- [9] J. C. Park, B. Song, J. S. Kim, S. H. Park, H. K. Kim, Z. Liu, *et al.*, "Fast compressed sensing-based CBCT reconstruction using Barzilai-Borwein formulation for application to on-line IGRT," *Medical physics*, vol. 39, pp. 1207-1217, 2012.
- [10] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-ray Science and Technology*, vol. 14, pp. 119-139, 2006.
- [11] X. Jia, Y. Lou, R. Li, W. Y. Song, and S. B. Jiang, "GPU-based fast cone beam CT reconstruction from undersampled and noisy projection data via total variation," *Medical physics*, vol. 37, pp. 1757-1760, 2010.
- [12] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Physics in medicine and biology*, vol. 53, p. 4777, 2008.
- [13] T. Chan, A. Marquina, and P. Mulet, "High-order total variation-based image restoration," *SIAM Journal on Scientific Computing*, vol. 22, pp. 503-516, 2000.
- [14] J. Yuan, C. Schnörr, and G. Steidl, "Total-variation based piecewise affine regularization," in *Scale Space and Variational Methods in Computer Vision*, ed: Springer, 2009, pp. 552-564.
- [15] T. Sun, N. Sun, J. Wang, and S. Tan, "Hessian-Based Norm Penalty for Weighted Least-Square CBCT Reconstruction," presented at the the 56th Annual Meeting of the American Association of Physicists in Medicine (AAPM), Austin, TX, USA, 2014.
- [16] J. Wang, T. Li, and L. Xing, "Iterative image reconstruction for CBCT using edge-preserving prior," *Medical physics*, vol. 36, pp. 252-260, 2008.
- [17] J. Wang, H. Lu, Z. Liang, D. Eremina, G. Zhang, S. Wang, *et al.*, "An experimental study on the noise properties of x-ray CT sinogram data in Radon space," *Physics in medicine and biology*, vol. 53, p. 3327, 2008.
- [18] J. Wang, T. Li, Z. Liang, and L. Xing, "Dose reduction for kilovoltage cone-beam computed tomography in radiation therapy," *Physics in medicine and biology*, vol. 53, p. 2897, 2008.
- [19] D. R. Hunter and K. Lange, "A tutorial on MM algorithms," *The American Statistician*, vol. 58, pp. 30-37, 2004.
- [20] C. J. Wu, "On the convergence properties of the EM algorithm," *The Annals of statistics*, pp. 95-103, 1983.
- [21] R. L. Siddon, "Fast calculation of the exact radiological path for a three - dimensional CT array," *Medical physics*, vol. 12, pp. 252-255, 1985.
- [22] X. Jia, B. Dong, Y. Lou, and S. B. Jiang, "GPU-based iterative cone beam CT reconstruction using tight frame regularization," *arXiv preprint arXiv:1008.2042*, 2010.
- [23] Y. Long, J. A. Fessler, and J. M. Balter, "3D forward and back-projection for X-ray CT using separable footprints," *Medical Imaging, IEEE Transactions on*, vol. 29, pp. 1839-1850, 2010.
- [24] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, pp. 600-612, 2004.