



OPEN

## Low-dose CBCT reconstruction via joint non-local total variation denoising and cubic B-spline interpolation

Ho Lee<sup>1</sup>, Jiyoung Park<sup>1</sup>, Yeonho Choi<sup>1</sup>, Kyung Ran Park<sup>2</sup>, Byung Jun Min<sup>3</sup>✉ & Ik Jae Lee<sup>1</sup>✉

This study develops an improved Feldkamp–Davis–Kress (FDK) reconstruction algorithm using non-local total variation (NLTV) denoising and a cubic B-spline interpolation-based backprojector to enhance the image quality of low-dose cone-beam computed tomography (CBCT). The NLTV objective function is minimized on all log-transformed projections using steepest gradient descent optimization with an adaptive control of the step size to augment the difference between a real structure and noise. The proposed algorithm was evaluated using a phantom data set acquired from a low-dose protocol with lower milliamperes-seconds (mAs). The combination of NLTV minimization and cubic B-spline interpolation rendered the enhanced reconstruction images with significantly reduced noise compared to conventional FDK and local total variation with anisotropic penalty. The artifacts were remarkably suppressed in the reconstructed images. Quantitative analysis of reconstruction images using low-dose projections acquired from low mAs showed a contrast-to-noise ratio with spatial resolution comparable to images reconstructed using projections acquired from high mAs. The proposed approach produced the lowest RMSE and the highest correlation. These results indicate that the proposed algorithm enables application of the conventional FDK algorithm for low mAs image reconstruction in low-dose CBCT imaging, thereby eliminating the need for more computationally demanding algorithms. The substantial reductions in radiation exposure associated with the low mAs projection acquisition may facilitate wider practical applications of daily online CBCT imaging.

Volumetric imaging, such as cone-beam computed tomography (CBCT) or megavoltage CT, is being rapidly adopted as a common modality in image guided radiation therapy (IGRT)<sup>1–4</sup>. In particular, CBCT mounted on the gantry of a linear accelerator is scanned to improve the target accuracy and precision by minimizing anatomic deviations prior to the beginning of each fraction<sup>5</sup>. The reduction of these setup uncertainties permits the safe implementation of a noticeable dose falloff between the boundaries of the planned treatment volume and the surrounding organs at risk. Moreover, several studies have utilized CBCT images on every treatment day for dose calculation to verify if the actual delivered dose deviates from the planned dose distributions<sup>6,7</sup>. In current protocols, the dose amount from each CBCT scan is approximately 3 cGy for parts of the skin and as high as 5 cGy for inside the body. When a patient is imaged daily during the course of radiotherapy that lasts four to six weeks<sup>8</sup>, the accumulated dose exceeds 100 cGy at the region inside the field of view (FOV). Although this amount may seem adequate in the context of radiotherapy, it cannot be ignored. If a judicious image quality can be obtained with a low dose, it is clear to comply the directions. Consequently, it is vital to develop an efficient reconstruction algorithm for daily CBCT imaging with dose reduction.

The most straightforward way to reduce dose of a CBCT system is to use either a lower mAs of the projection data or a shorter exposure time per projection<sup>9</sup>. Utilization of these approaches leads to excessive noise in the reconstructed images. The iterative and analytical reconstruction algorithms can be considered via these strategies. Compressed sensing (CS) theory in iterative reconstruction methods has been commonly utilized to generate a reconstructed image with satisfactory quality from a highly undersampled number of projections below the Nyquist rate. The CS method involves minimizing a total variation objective function subject to data fidelity and physical constraints through constrained or unconstrained optimization methods. Its efficacy has

<sup>1</sup>Department of Radiation Oncology, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Korea. <sup>2</sup>Department of Radiation Oncology, Kosin University College of Medicine, Busan, Korea. <sup>3</sup>Department of Radiation Oncology, Chungbuk National University Hospital, Cheongju, Korea. ✉email: bjmin@cnuh.or.kr; ikjae412@yuhs.ac

been improved by incorporating prior images into an objective function<sup>10,11</sup>. However, the use of such approaches is primarily limited by their high computational demands in practical applications, because the reprojection and backprojection steps are iteratively performed to calculate the fidelity between measured and estimated projections. Meanwhile, analytical reconstruction algorithms such as Feldkamp–Davis–Kress (FDK)<sup>12</sup> enable high speed performance by using only one filtering process and one backprojection process. These approaches are required when the projection data are complete (with the number of projections fully available) because its performance is greatly influenced by the filtering process. However, the excess noise caused by the low mAs protocol leads to significantly degraded CBCT image qualities when they are attempted using FDK with conventional filtering. An efficient denoising technique<sup>13,14</sup> is thus required to enhance the difference in the signals between real structures and unwanted noise for better CBCT image quality. Recently, several deep learning-based reconstruction approaches have been reported. Deep learning is a new concept with respect to image quality optimization without hand-engineering mapping functions. One way to use deep learning methods for image-to-image conversion is supervised training based on paired images<sup>15–18</sup>. These approaches employ deep neural networks in the image or projection domain<sup>19</sup>. Another way is unsupervised training using unpaired images, such as cycle-generative adversarial networks<sup>20</sup>. For supervised learning, it is difficult to obtain anatomically exact matching paired images. Even for unsupervised learning, sufficient training data are required to achieve good results. However, the lack of injectivity and stability of deep learning approaches for image reconstruction poses potential issues<sup>21</sup>. Therefore this type of deep learning algorithm should be further explored.

In this study, an improved FDK-based CBCT reconstruction algorithm is proposed based on non-local total variation (NLTV) filtration and cubic B-spline interpolation for low mAs of projection data. The proposed approach employs the framework of the standard FDK algorithm. The NLTV denoising process is applied on the log-transformed projection data. Minimizing the NLTV objective function using the steepest gradient descent optimization with an adaptive control of the step size indicates that edges having high contrast relative to the surroundings are preserved and noisy pixels having a low contrast are smoothed. With filtered projections that apply a ramp filtering on NLTV minimized projections, a voxel-driven backprojector based on a cubic B-spline interpolation is utilized to generate the reconstructed images. The effectiveness of the proposed method is demonstrated using phantom studies.

## Methods

Figure 1 shows a flowchart of the proposed algorithm using joint NLTV denoising-based filtration and a cubic B-spline interpolation-based backprojector to enhance the image quality of low-dose CBCT from projection data acquired from a low-dose protocol with lower mAs. The NLTV denoising is used to increase the difference between the actual structure and the noise of the projection data. The cubic B-spline interpolation is applied to update the voxels in the reconstructed image by resampling 16 adjacent pixels in the projection domain during backprojection.

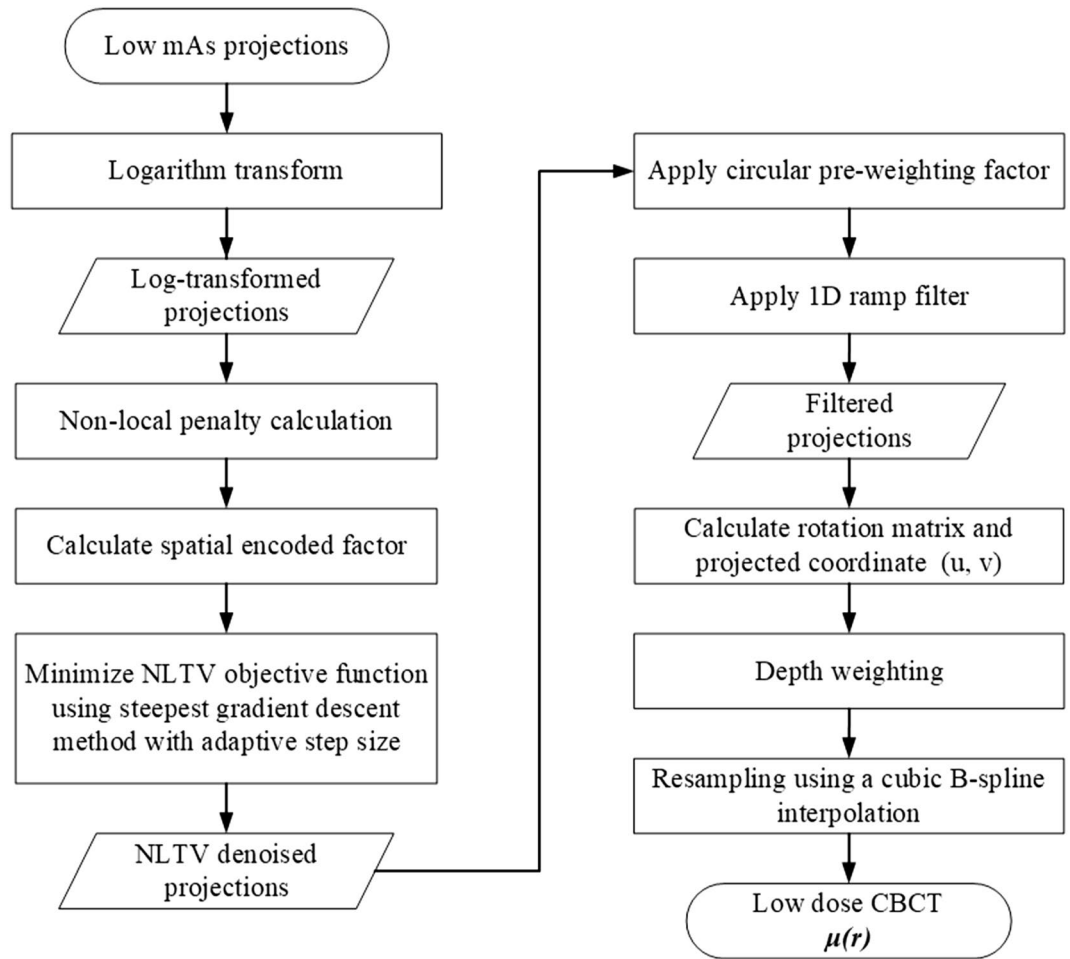
**Filtration using non-local total variation.** A total variation<sup>22</sup> denoising process on log-transformed projections is applied to enhance the difference in the signals between striking features and unwanted noise by similarities between non-local patches<sup>23–25</sup>. Minimizing the NLTV objective function indicates that edges having high contrast relative to the surroundings are preserved and noisy voxels having a low contrast are smoothed. The non-local penalty with different weights for neighbors of the same distance in the TV object function is written as follows:

$$R(P) = \sum_j R(P_j) = \sum_j w_j D(P_j), \quad (1)$$

$$w_j = \sum_{i \in \Omega} \exp \left( - (P_j / \tau)^\varepsilon \frac{\sum_{k=-a}^{k=a} G(k) (P_{(j+k)} - P_{(i+k)})^2}{2h_0^2} \right), \quad (2)$$

$$D(P_j) = \sqrt{(P_{(u,v)} - P_{(u-1,v)})^2 + (P_{(u,v)} - P_{(u,v-1)})^2}, \quad (3)$$

where index  $j$  identifies the index of the pixel element in the projection and  $P_{(u,v)}$  is the pixel element at the 2D position  $(u, v)$ . The search set for the non-local areas is defined as  $i \in \Omega$  and  $w_j$  represents the weights between the current voxel  $j$  and the voxels in  $\Omega$ .  $G(k)$  is the Gaussian kernel with a patch size of  $((2a + 1) \times (2a + 1))$ , and  $k$  identifies the index of the patch. In this work, the patch of the Gaussian kernel is defined to be  $5 \times 5$  with unit variance; the non-local search area is  $21 \times 21$  with unit variance, and the filtering parameter  $h_0$  is set to 90% of the cumulative distribution function (CDF) histogram created by accumulating the gradient magnitude at each pixel of the log-transformed projection data.  $(P_j / \tau)^\varepsilon$  is the spatially encoded factor for the  $j$ -th pixel, which has two hyperparameters  $\tau$  and  $\varepsilon$ <sup>24</sup>. Its role is to reduce the weighted averaging effect in high intensity regions to maintain the contrast. The parameter  $\tau$  is the normalization factor to make the  $P_j / \tau$  ratio exceed 1. In this study,



**Figure 1.** Overview of the proposed method for low-dose CBCT imaging.

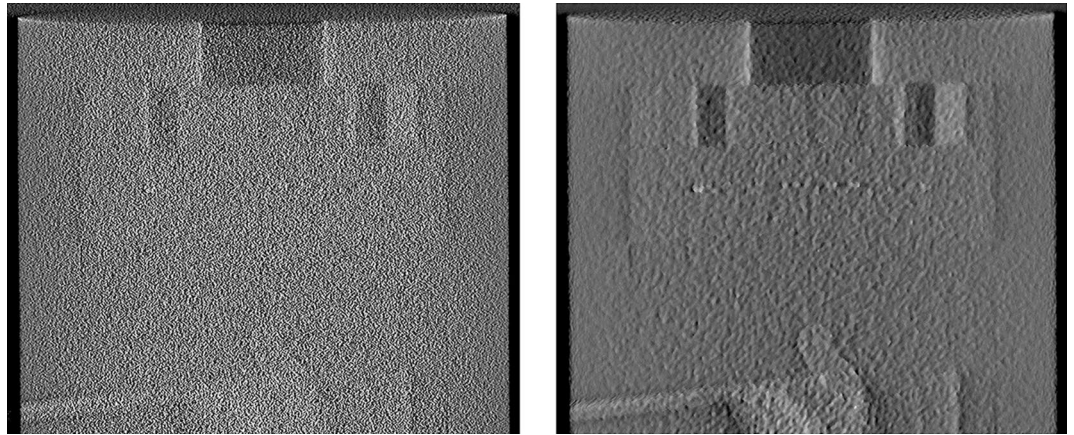
the parameter  $\tau$  was set to 90% of the CDF histogram created by accumulating the intensity at each pixel of the log-transformed projection data. The parameter  $\varepsilon$  (in this study, we used  $\varepsilon = 3$ ) is the scaling factor to make smaller weights for higher intensities, thereby reducing the blurring effect in high intensity regions.

The NLTV objective function of Eq. (2) is minimized by the steepest gradient descent method with an adaptive step size, which is expressed as follows:

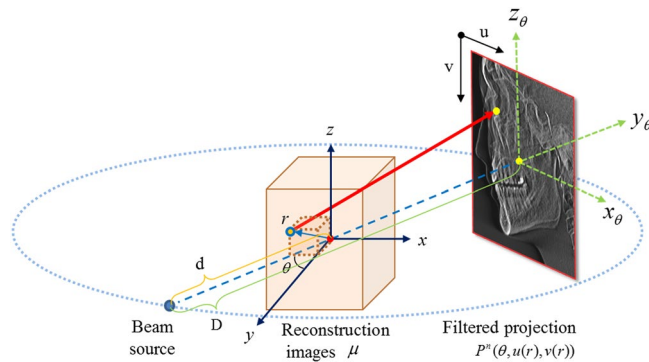
$$P_j^{t+1} = P_j^t - \lambda \frac{\nabla R(P_j)}{|\nabla R(P)|}, \tag{4}$$

$$\lambda = \gamma \sqrt{\sum_j P_j^2}, \tag{5}$$

where  $\lambda$  is an adaptive parameter that controls the step size so that the smoothing degree is decreased according to an advancement in the iteration steps. By using the square root of all pixel elements updated in each steepest gradient descent step, this parameter is forced to progressively smaller values with an increase in the number of iterations. To avoid a local minimization due to an abrupt change, we use a scaling parameter  $\gamma$  and set it to an initial value of 1.0. If  $R(P)$  calculated at the current iteration step is larger than that at the previous one, this value is linearly decreased by multiplying a constant value ( $r_{red} = 0.8$ ).  $\nabla R(P_j)$  is the gradient of the objective function  $R(P)$  at the  $j$ -th indexed pixel. The root-sum-square of the gradient calculated at all pixels,  $|\nabla R(P)|$ , is required for the normalized gradient calculation. For clarity, we provide the following expression:



**Figure 2.** Comparison of filtered projection without (left) and with (right) the NLTV denoising algorithm.



**Figure 3.** Geometry of the projection data and notations for the voxel-driven backprojection step.

$$\nabla R(P_j) = \frac{\partial R(P)}{\partial P_j} = \frac{\partial R(P)}{\partial P_{(u,v)}} = \begin{pmatrix} w_{(u,v)} \frac{2P_{(u,v)} - P_{(u-1,v)} - P_{(u,v-1)}}{\sqrt{(P_{(u,v)} - P_{(u-1,v)})^2 + (P_{(u,v)} - P_{(u,v-1)})^2}} \\ + w_{(u+1,v)} \frac{P_{(u,v)} - P_{(u+1,v)}}{\sqrt{(P_{(u+1,v)} - P_{(u,v)})^2 + (P_{(u+1,v)} - P_{(u+1,v-1)})^2}} \\ + w_{(u,v+1)} \frac{P_{(u,v)} - P_{(u,v+1)}}{\sqrt{(P_{(u,v+1)} - P_{(u-1,v+1)})^2 + (P_{(u,v+1)} - P_{(u,v)})^2}} \end{pmatrix}, \quad (6)$$

$$|\nabla R(P)| = \sqrt{\sum_j (\nabla R(P_j))^2}, \quad (7)$$

The optimal number of iterations for the steepest gradient descent step was fine-tuned so that the noisy pixels were minimized while preserving the striking features. In this study, the number of iterations for the steepest gradient descent step was set to ten. In considering all the components mentioned above, we present the pseudo-code of NLTV-based filtration as in Algorithm 1.

A circular pre-weighting factor was then applied on each NLTV denoised projection data to avoid the intensity drop due to the cone angle effect. A 1D ramp filter was then applied to suppress the highest spatial frequency. For the ramp filter, we used the product of the SL filter and raised the cosine window function in the frequency domain. Each line parallel to the *u*-direction of the preweighted projection data was converted into the frequency domain via 1D Fourier transformation. The Fourier transformed values were multiplied by the ramp filter. By applying an inverse Fourier transformation, we could obtain the filtered projection data with the reduced impulse noise. Figure 2 depicts the differences between filtered projections before and after applying the NLTV denoising algorithm.

**Algorithm 1**


---

```

For n ← 1 to all projections do
  r ← 1,   rred ← 0.8,   ε ← 3,   a ← 2,   Ω ← 10
  Calculate Gaussian kernel with a patch
  For j ← 1 to all pixels do
    Calculate D(Pj) using Eq. (3)
    Create gradient magnitude CDF histogram using D(Pj)
    Create intensity CDF histogram using Pj
  End For
  h0 ← gradient magnitude at 90% of gradient magnitude CDF
  τ ← Intensity at 90% of intensity CDF
  R(P) ← 0
  For j ← 1 to all pixels do
    wj ← 0
    For i ← j - Ω to j + Ω do
      S ← 0
      For k ← a to a do
        S ← S + G(k)(Pj+k - Pi+k)
      End For
      wj ← wj + exp(-Pj/τ) * (S / (2h02))
    End For
    R(Pj) ← wjD(Pj)
    R(P) ← R(P) + R(Pj)
  End For
  For t ← 1 to 10 do
    For j ← 1 to all pixels do
      λ ← √(∑j Pj2)
      λ ← λ × r
      ∂Pj ← ∇R(Pj) calculated by Eq. (6)
      |∇R(P)| ← √(∑j (∂Pj)2) as in Eq. (7)
    End For
    For j ← 1 to all pixels do
      Pj' ← Pj + λ∂Pj/|∇R(P)|
    End For
    While R(Pj') > R(P) do
      r ← r × rred
      λ ← λ × r
      For j ← 1 to all pixels do
        Pj' ← Pj + (λ∂Pj) / |∇R(P)|
      End For
    End While
    Update Pj' to Pj
  End For
End For

```

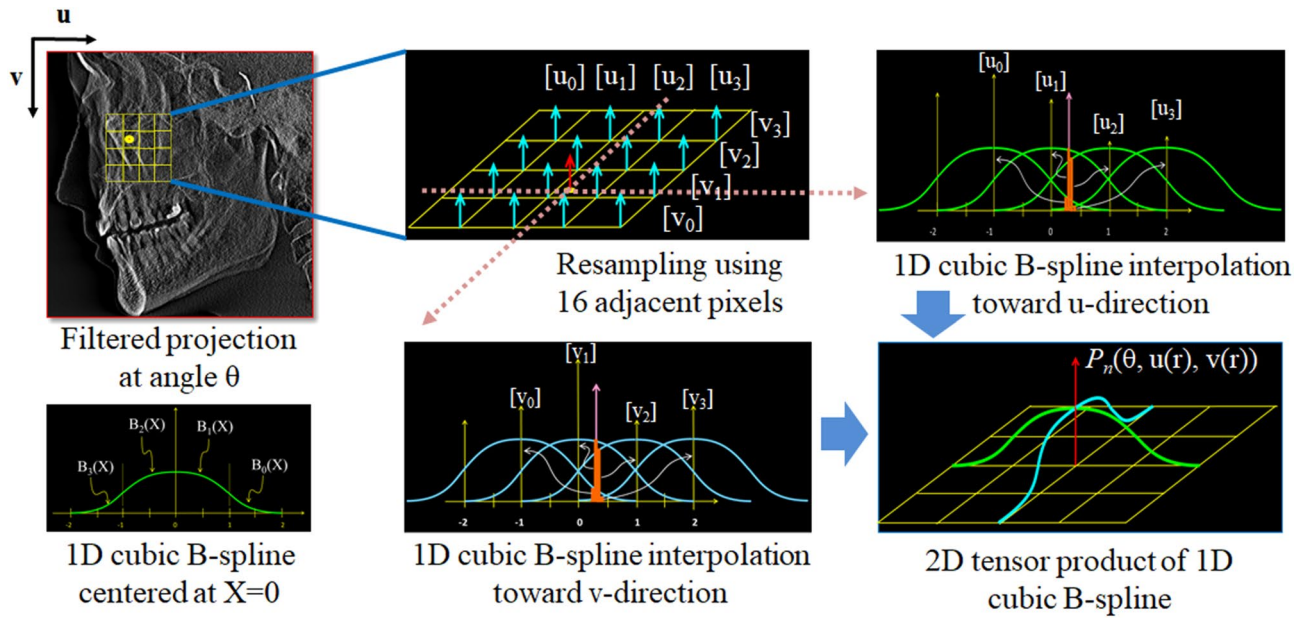
---

**Voxel-driven backprojector based on cubic B-spline interpolation.** Figure 3 shows the geometric information of the projection data and notations for voxel-driven backprojection calculation using the filtered projection data.

Given a position vector  $\mathbf{r}$  (or  $r_x, r_y, r_z$ ) of any voxel, corresponding positions ( $u(r)$  and  $v(r)$ ) in the  $n$ -th projection data at a given angle  $\theta$  are obtained as follows:

$$P(\theta, r) = P^n(\theta, u(r), v(r)), \quad (8)$$

where



**Figure 4.** Resampling process using 2D tensor product of 1D cubic B-spline interpolation from n-th filtered projection.

$$u(r) = \frac{r_{x\theta}}{d + r_{y\theta}}, \tag{9}$$

$$v(r) = \frac{r_{z\theta}}{d + r_{y\theta}}, \tag{10}$$

Here,  $d$  is the source-to-axis distance, and the rotation vectors  $[r_{x\theta}, r_{y\theta}, r_{z\theta}]^T$  are calculated by applying the rotation matrix through a given angle  $\theta$  from the origin of the reconstruction images in the following manner:

$$\begin{pmatrix} r_{x\theta} \\ r_{y\theta} \\ r_{z\theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix}, \tag{11}$$

The final attenuation coefficient value,  $\mu(r)$ , at a position vector  $r$  is weighted and accumulated as

$$\mu(r) = \mu(r_x, r_y, r_z) = \frac{\pi}{N} \sum_{n=1}^N (w(r)P^n(\theta, u(r), v(r))), \tag{12}$$

$$w(r) = d^2 / (d + r_{y\theta})^2, \tag{13}$$

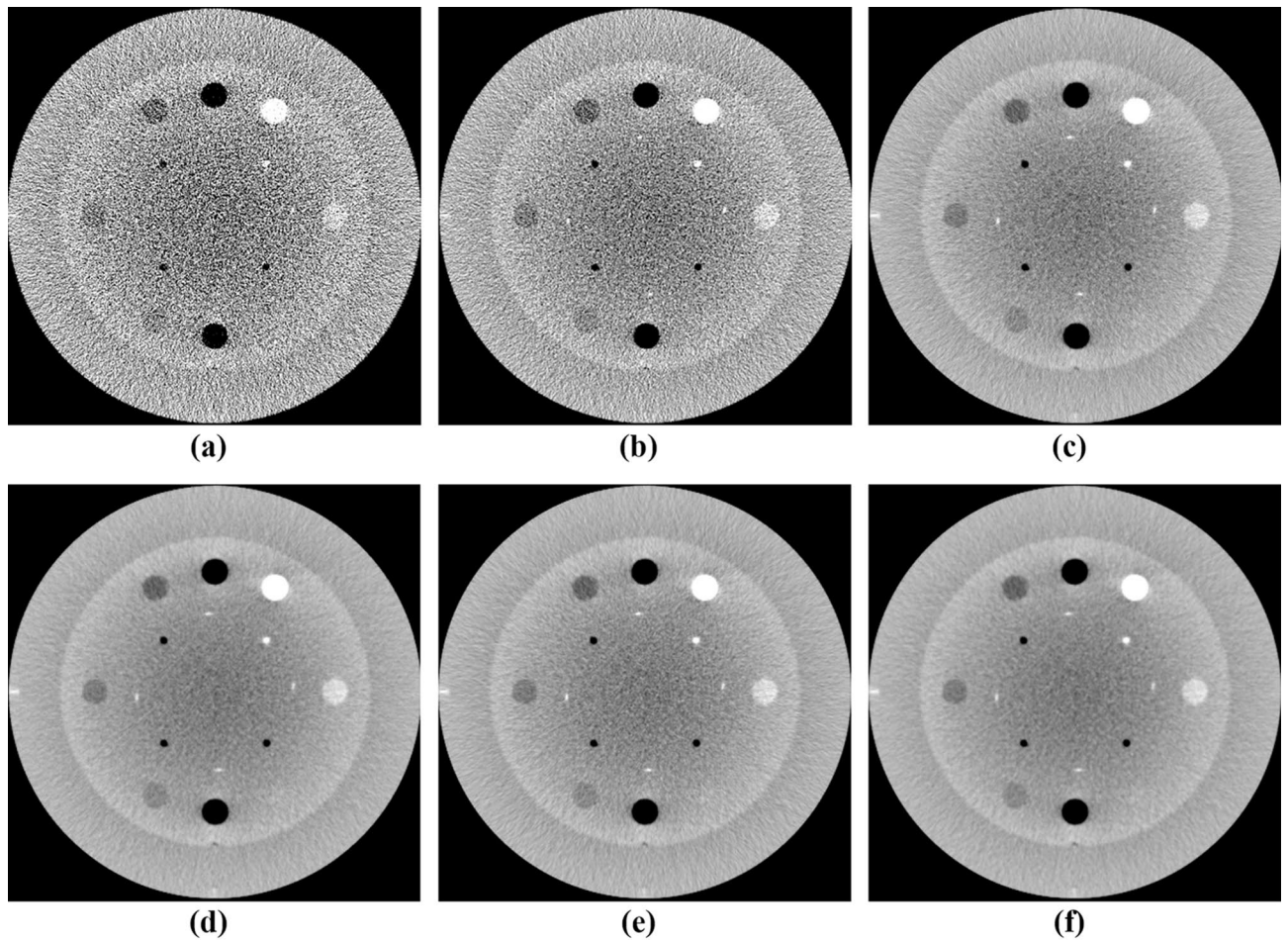
where  $w(r)$  is the depth weighting and  $N$  is the total number of projection data. Every voxel has measurements in all filtered projection data and is normalized by a uniform factor  $N$  in the backprojection step. As shown in Fig. 4,  $P^n(\theta, u(r), v(r))$  is represented as the 2D tensor product of a 1D cubic B-spline:

$$P^n(\theta, u(r), v(r)) = \sum_{p=0}^3 \sum_{q=0}^3 B_p(u - [u_p])B_q(v - [v_q])P^n(\theta, [u_p], [v_q]), \tag{14}$$

where  $u_p = u + p - 1, v_q = v + q - 1$ , and  $[\cdot]$  is the floor function that maps a real number to the largest previous integer.  $B_p$  and  $B_q$  correspond to the following cubic B-spline basis functions centered at  $X=0$ :

$$\begin{aligned} B_0(X) &= (2 - X)^3/6 && \text{if } 1 \leq X \leq 2 \\ B_1(X) &= (3X^3 - 6X^2 + 4)/6 && \text{if } 0 \leq X < 1 \\ B_2(X) &= (-3X^3 - 6X^2 + 4)/6 && \text{if } -1 \leq X < 0 \\ B_3(X) &= (2 + X)^3/6 && \text{if } -2 \leq X < -1. \end{aligned} \tag{15}$$

In considering all the components mentioned above, we present the pseudo-code of cubic B-spline interpolation-based backprojector as in Algorithm 2.



**Figure 5.** Comparison between the same views of the high resolution reconstruction image generated by applying FDK with (a) SL, (b) ATV, (c) NLTv, (d) SL + cubic B-spline, (e) ATV + cubic B-spline, and (f) NLTv + cubic B-spline using the Catphan503 phantom. These images are displayed at window (width and level) settings of (975, 0) HU.

---

### Algorithm 2

---

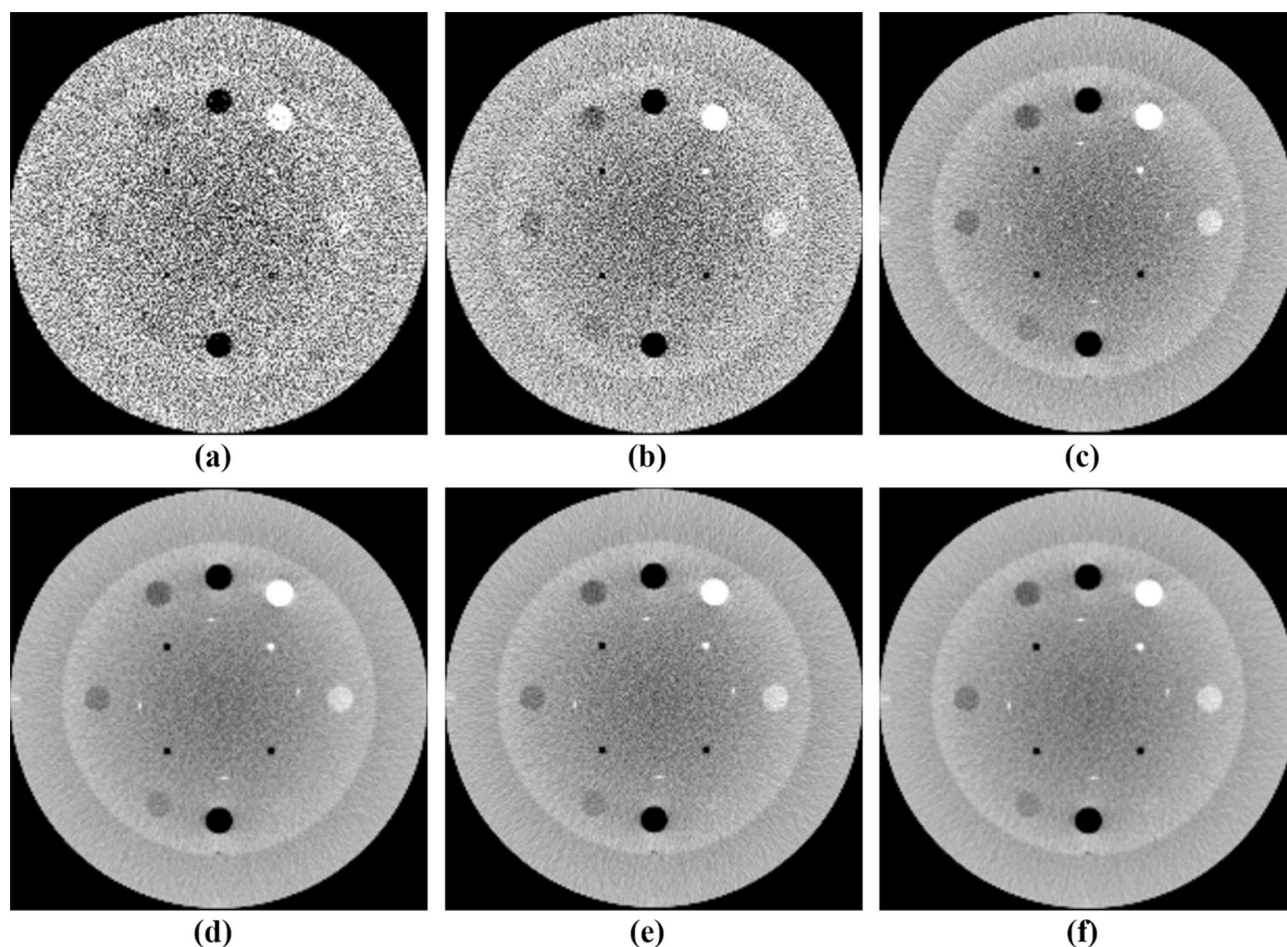
```

For  $n \leftarrow 1$  to all projections do
  Calculate rotation matrix  $M$  which maps voxel coordinate to projection coordinate
  For  $r \leftarrow 1$  to all voxels do
    Calculate projected coordinate  $(u, v) \leftarrow Mr$ 
    For  $p \leftarrow 0$  to 3 do
       $[u_p] \leftarrow \text{floor}(u + p - 1)$ 
    End For
    For  $q \leftarrow 0$  to 3 do
       $[v_q] \leftarrow \text{floor}(v + q - 1)$ 
    End For
    Sampling  $P^n(\theta, u(r), v(r))$  by cubic B-spline interpolation using Eqs. (14) and (15)
     $r \leftarrow r + w(r) \times P^n(\theta, u(r), v(r))$ 
  End For
End For

```

---

**Experimental studies.** The Catphan503 (Phantom Laboratory, Salem, NY) was scanned using an Infinity™ linear accelerator system with XVI R5.0 (Elekta Limited, Stockholm, Sweden). The phantom includes various modules for assessing the quality of 3D images. As the gantry rotates around the phantom, a series of CBCT projections were acquired with 360° rotation under a small FOV protocol using a S20 collimator and an F0 filter.



**Figure 6.** Comparison between the same views of the low resolution reconstruction image generated by applying FDK with (a) SL, (b) ATV, (c) NLTV, (d) SL + cubic B-spline, (e) ATV + cubic B-spline, and (f) NLTV + cubic B-spline using the Catphan503 phantom. These images are displayed at window (width and level) settings of (975, 0) HU.

ROI	Material of insert	SL	ATV	NLTV	SL + cubic B-spline	ATV + cubic B-spline	NLTV + cubic B-spline
1	Delrin™	2.0	8.7	14.4	3.5	11.6	16.4
2	Teflon	4.2	14.3	23.6	7.0	18.2	25.9
3	Air	4.2	20.5	34.8	8.0	28.1	39.8
4	PMP	0.4	1.8	3.2	0.7	2.5	3.8
5	LDPE	0.1	0.2	0.4	0.1	0.3	0.5
6	Polystyrene	0.1	0.8	1.4	0.3	1.1	1.6
7	Air	4.3	21.0	35.9	8.3	29.1	41.9

**Table 1.** Comparison of CNRs at seven ROIs in the high resolution reconstructed image generated by six FDK algorithms with low-dose projection data of the Catphan503 phantom.

The number of full scan projections was approximately 665. The size of each projection acquired on the image receptor was  $409.6 \times 409.6 \text{ mm}^2$ , containing  $1024 \times 1024$  pixels. The distance from the source to the detector and the distance from the source to the axis were 1536 mm and 1000 mm, respectively. The low-dose acquisition protocol had 100kVp with 0.1 mAs during each projection. Projections with different mAs settings were also acquired for comparison. The reconstructed image of the phantom was generated with two resolution settings: high resolution CBCT ( $512 \times 512 \times 200$  voxels as well as  $0.5 \times 0.5 \times 0.5 \text{ mm}^3$  per each voxel) and low resolution CBCT ( $256 \times 256 \times 100$  voxels as well as  $1.0 \times 1.0 \times 1.0 \text{ mm}^3$  per each voxel). For comparison, a high-dose CBCT was created as a benchmark image by applying a conventional FDK algorithm with projections acquired using increased mAs setting (100 kVp and 1.6 mAs). The Hounsfield unit (HU) conversion was performed for all



ROI	Material of insert	SL	ATV	NLTV	SL + cubic B-spline	ATV + cubic B-spline	NLTV + cubic B-spline
1	Delrin™	1.7	7.6	11.3	3.4	10.1	13.7
2	Teflon	4.2	13.2	20.6	6.5	17.8	24.2
3	Air	4.7	19.3	29.9	8.3	25.5	33.7
4	PMP	0.5	2.1	3.4	1.0	2.9	4.1
5	LDPE	0.1	0.3	0.4	0.1	0.3	0.5
6	Polystyrene	0.1	0.4	0.9	0.2	0.6	1.0
7	Air	4.0	19.7	31.6	7.8	27.3	37.1

**Table 2.** Comparison of CNRs at seven ROIs in the low resolution reconstructed image generated by six FDK algorithms with low-dose projection data of the Catphan503 phantom.

	SL	ATV	NLTV	SL + cubic B-spline	ATV + cubic B-spline	NLTV + cubic B-spline
RMSE	181.0	69.0	62.4	112.6	64.3	61.4
Correlation	0.38	0.76	0.82	0.57	0.80	0.83

**Table 3.** Quantitative comparisons using two metrics in the high resolution reconstruction image generated by six FDK algorithms with low-dose projection data of the Catphan503 phantom.

	SL	ATV	NLTV	SL + cubic B-spline	ATV + cubic B-spline	NLTV + cubic B-spline
RMSE	207.0	136.8	133.1	155.9	135.3	132.2
Correlation	0.43	0.77	0.82	0.57	0.80	0.82

**Table 4.** Quantitative comparisons using two metrics in the low resolution reconstruction image generated by six FDK algorithms with low-dose projection data of the Catphan503 phantom.

	SL	ATV	NLTV	SL + cubic B-spline	ATV + cubic B-spline	NLTV + cubic B-spline
High resolution CBCT	95.8	269.4	1398.6	192.0	355.2	1485.8
Low resolution CBCT	79.7	251.4	1379.4	91.1	261.9	1390.0

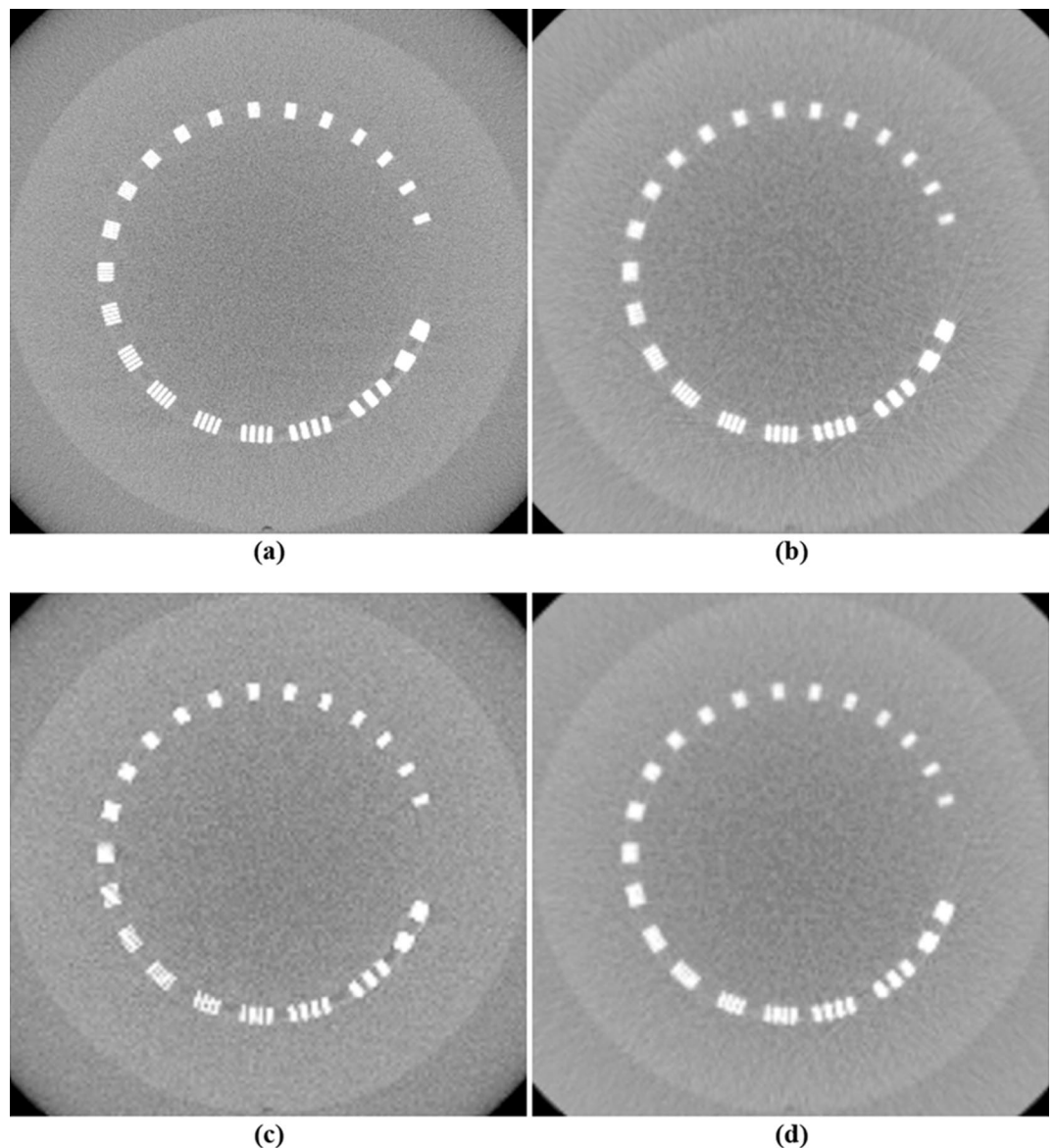
**Table 5.** Computation time (seconds) of six FDK algorithms when creating high resolution and low resolution CBCTs of the Catphan503 phantom.

reconstructed images. The CBCT image quality was quantified using three metrics (CNR, HU accuracy, and correlation).

The accuracy of the HU number was assessed using the entire area of the CTP404 module of the phantom. It has seven sensitometry inserts and the RMSE is calculated. The correlation and the CNR were also measured using the same module. The correlation was calculated to evaluate the overall differences of the reconstructed CBCT image versus benchmark CBCT images. For the RMSE and correlation measurement, a circular area was selected to be measured only inside the phantom excluding the air area outside the phantom. For the CNR measurement, the mean HU value and standard deviation were recorded by selecting seven ROIs within the density insert and a central ROI.

## Results

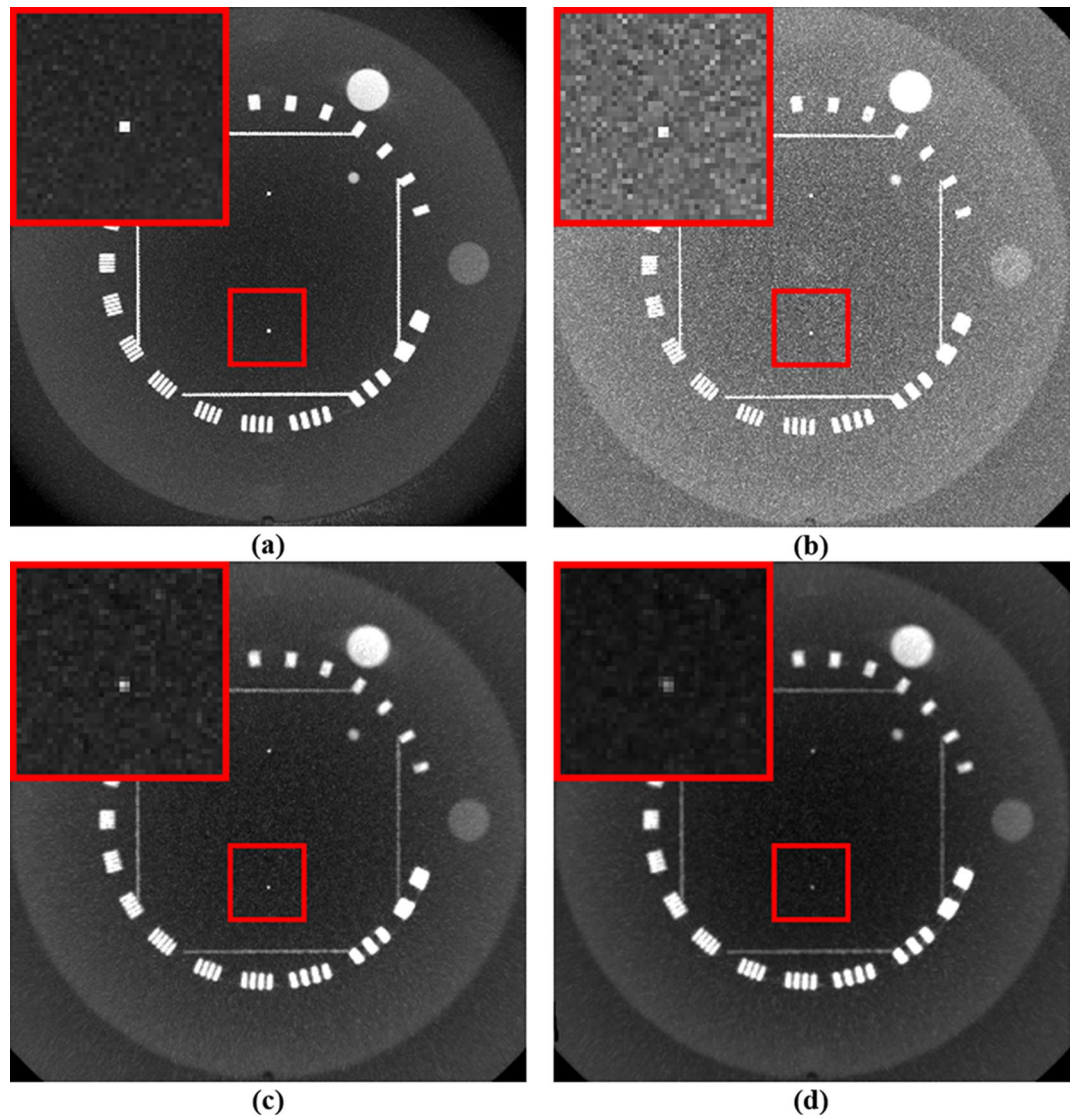
The performance of the NLTV denoised FDK reconstruction algorithm was evaluated with the Catphan503 phantom and compared with the Shepp–Logan (SL)-filtered FDK and anisotropic total variation (ATV)-denoised FDK<sup>26</sup>. The ATV combines the edge stopping function used locally in an anisotropic diffusion filter<sup>27</sup> for both image enhancement and noise reduction. Figures 5 and 6 show a representative slice of the high resolution CBCT (512 × 512 × 200 voxels) and low resolution CBCT (256 × 256 × 100 voxels) reconstructed from FDKs with SL, ATV, and NLTV, as well as when adding the voxel-driven backprojector based on cubic B-spline interpolation on filtered projections with SL, ATV, and NLTV, respectively. The artifacts were significantly suppressed in images reconstructed by the proposed NLTV denoising compared to the other two algorithms. The CBCT images resampled with the cubic B-spline interpolation using 16 pixels (4 × 4) were smoother and had fewer interpolation artifacts, resulting in slightly improved reconstruction accuracy for SL, ATV, and NLTV. The effect of this image



**Figure 7.** Comparison of a representative slice of the reconstruction image generated by applying FDK using high mAs projections and FDK with NLTV + cubic B-spline interpolation using low mAs projections. (a) high resolution benchmark image, (b) high resolution CBCT produced by the proposed method, (c) low resolution benchmark image, and (d) low resolution CBCT produced by the proposed method. These images are displayed at window (width and level) settings of (2000, 0) HU.

quality improvement was the same for low resolution CBCT. In particular, the cubic B-spline interpolation still visibly achieved substantial gains on all SL, ATV, and NLTV.

In addition, we calculated the contrast-to-noise ratio (CNR) at a selected region of interest (ROI) in the reconstructed image. It was possible to conduct a comparison of the relative image contrast between the corresponding regions. Tables 1 and 2 provide a comparison of CNRs at seven ROIs in the high resolution and low resolution reconstructed CBCT images obtained by six FDK algorithms. The proposed NLTV denoising shows a better CNR in all ROIs compared to other two algorithms. Adding the cubic B-spline interpolation process slightly improved the CNR to higher values for SL, ATV, and NLTV. Tables 3 and 4 list two quantitative measures (root-mean-square error (RMSE) and correlation) obtained from the high resolution and low resolution reconstructed images based on the six FDK methods. The errors were computed against full projections with a high mAs acquisition protocol (40 mA and 40 ms). The NLTV elicited a lower RMSE and a higher correlation compared to the other algorithms. Adding the cubic B-spline interpolation produced the best results for the Catphan503 phantom with a low noise scenario. Even in the SL and ATV, the cubic B-spline interpolation could yield a lower RMSE and a higher correlation. Our proposed NLTV with the cubic B-spline interpolation produced the lowest RMSE and the highest correlation. This relative superiority of the proposed method was the same for the low resolution CBCT. The low resolution CBCT doubled the voxel size compared to the high

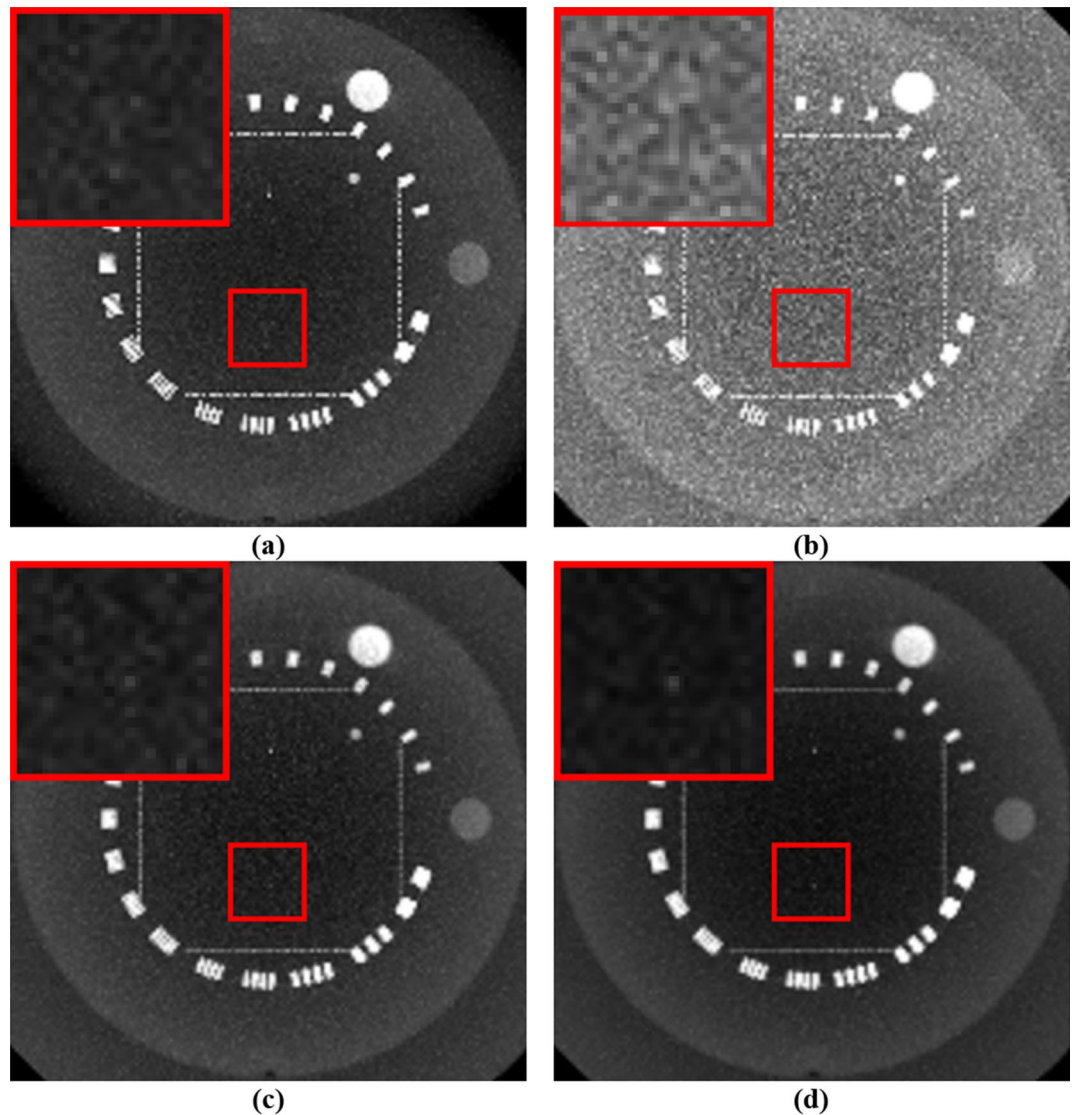


**Figure 8.** Comparison of MIP of (a) high resolution benchmark CBCT images with a high mAs and high resolution CBCT reconstruction images generated by applying FDK algorithms with (b) SL + cubic B-spline, (c) ATV + cubic B-spline, and (d) NLTV + cubic B-spline using Catphan503 phantom acquired under a low mAs. These images are displayed at window (width and level) settings of (1100, 500) HU.

resolution CBCT, leading to inferior feature details. This resulted in larger RMSE values and smaller CNR values in the quantitative analysis.

Table 5 compares the computation times between the proposed method and other methods used in Figs. 5 and 6. Compared to SL and ATV, the computation time using the proposed NLTV denoising was approximately 14.5 times and 5.1 times longer for the high resolution CBCT generation and 17.3 times and 5.4 times longer for the low resolution CBCT. It was influenced by the size and number of projections because NLTV was applied to the projection data. Thus, the difference in calculation time due to high and low resolutions was not large. On the other hand, the calculation time before and after applying the cubic B-spline interpolation required approximately 90 s longer for high resolution CBCT and 10 s longer for the low resolution CBCT. Since the cubic B-spline interpolation was applied after estimating the projection pixel for each voxel in the reconstructed volume, it was affected by the reconstruction volume size. This resulted in a large difference in the calculation time between the high and low resolution CBCTs.

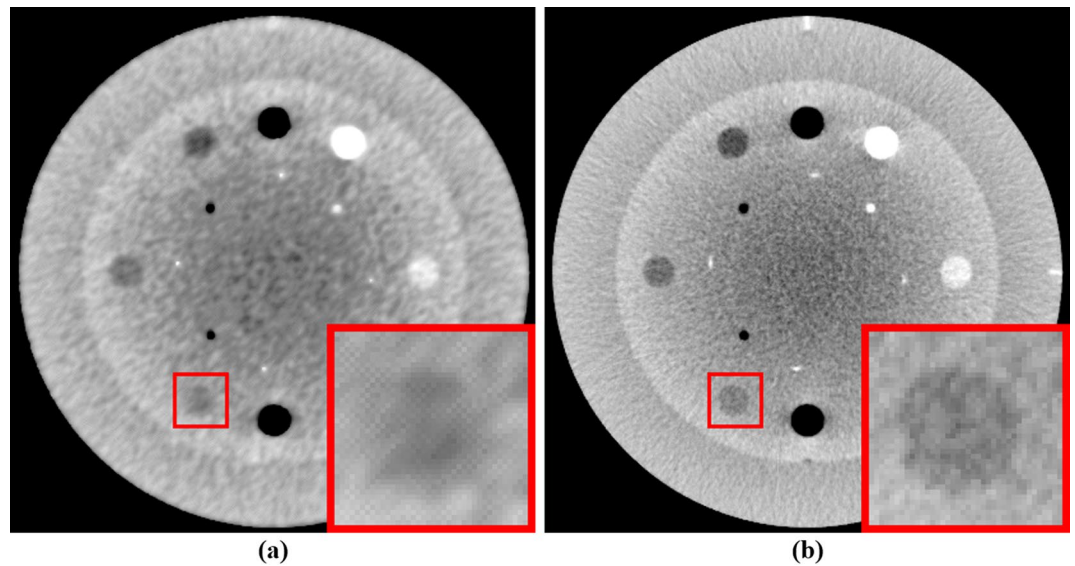
Figure 7 shows the proposed method with projections acquired at low mAs levels and the conventional FDK algorithm with projections acquired at high mAs levels (40 mA and 40 ms). These images enable us to qualitatively judge the correspondence of line bars between the two reconstruction images. It was observed that the proposed method shows a spatial resolution comparable to that of high-dose CBCT imaging through visual inspection. Even with a high voxel size and low resolution, the visual effect was better than a low resolution benchmark image with false artificial traces.



**Figure 9.** Comparison of MIP of (a) low resolution benchmark CBCT images with a high mAs and low resolution CBCT reconstruction images generated by applying FDK algorithms with (b) SL + cubic B-spline, (c) ATV + cubic B-spline, and (d) NLTV + cubic B-spline using Catphan503 phantom acquired under a low mAs. These images are displayed at window (width and level) settings of (1100, 500) HU.

Figures 8 and 9 compare the maximum intensity projections (MIP) of the reconstructed images. The MIP image facilitates extraction of the highest value encountered along the corresponding viewing ray at each pixel of the reconstructed images so that one can determine whether some regions with a high contrast are preserved while reducing the noise. The proposed method showed that regions with remarkable features are well preserved while reducing artifacts in the central region corresponding to the benchmark image compared to the other two algorithms. After adding NLTV, the changes in the homogeneous region were smoother, whereas those in the inhomogeneous regions were almost preserved. For the low resolution CBCT, a thin point feature showing high contrast in the red ROI disappeared, even in the case of the benchmark image. However, it was found that this thin feature was preserved in the reconstructed image when the proposed NLTV denoising was performed.

Figure 10 shows a visual inspection of the CBCTs generated by the post-processing approach of applying NLTV to the FDK reconstructed image. In addition, it depicts the proposed method, i.e., the pre-processing approach of applying FDK to the NLTV-denoised projections. We compared the low contrast resolution properties between the two approaches, as in the magnified area. The pre-processing strategy further revealed a higher low contrast detection and detail-preserving effect compared to the post-processing strategy.



**Figure 10.** Comparison between high resolution CBCT images generated by applying (a) NLTV on FDK reconstructed images as a post-processing strategy and (b) FDK with NLTV-denoised projection data as a pre-processing strategy. These Catphan503 phantom images are displayed at window (width and level) settings of (750, 0) HU.

## Discussion

CBCT reconstruction from low mAs has been challenging. X-ray tubes produce photons via electron bombardment of a metallic anode. The number of photons is determined by the current across the tube (mA) and the exposure time (s), which are sometimes multiplied together as mAs. Decreasing the mAs decreases the number of electrons hitting the anode, and consequently, the number of photons generated by the tube decreases linearly. This decreases the number of photons reaching the detector, reducing the signal-to-noise ratio (SNR) of the projection data.

The total variation (TV) transform was considered to mitigate the noise imprinted on the projection data, but it tended to keep the edge information too smooth by uniformly penalizing the local image gradient. To overcome this limitation, ATV was proposed to reduce the blurriness at edge regions by deriving the anisotropic edges expressed as the exponential function of the adaptively weighted local image gradient of the projection data. The limitation of ATV is that the capability of deriving local edge information is restricted by the quality of projection data. The projection data acquired at low mAs settings suffer from severe noise. Thus, it is difficult to develop a local edge detection operator that reliably finds features with smaller details and is noise-insensitive. Conversely, the NLTV can help overcome this shortcoming by utilizing more global searches and non-uniform weights depending on the similarities between non-local patches. Combining the cubic B-spline interpolation-based backprojector allows the NLTV denoising-based FDK algorithm to be further enhanced and obtain more precise low-dose CBCT reconstruction images. The reason for such a performance improvement was that the CBCT reconstructed from conventional FDK with SL, ATV, and NLTV were based on nearest neighbor (NN) interpolation-based backprojector. NN is the simplest and fastest interpolation method, but it can introduce significant distortion. The cubic B-spline interpolation is more complex than the NN method and is thus more computationally expensive. For an unknown pixel in the projection data, the influence range is extended to 16 adjacent pixels. Then the pixel value is calculated as 16 pixels based on the distance. More complex improved interpolation methods such as cubic B-spline make the interpolation curve smoother. Thus, there are no discrete defects and the high-frequency components are faded. This characteristic provides better visual effects without false artificial traces even at high voxel sizes and low resolutions.

We applied our proposed approach, the combination of NLTV and cubic B-spline interpolation, to phantom data and compared its performance with the results from SL and ATV filtrations in terms of visual inspection, RMSE, CNR, and correlation. Reconstructions performed by the proposed method indicated that a higher image quality could be achieved with projection data acquired under a lower mAs protocol. We also determined that the proposed method shows a CNR with a spatial resolution comparable to those of images reconstructed using projections acquired from a high-dose protocol. Moreover, NLTV denoising yields filtered projection data with a greatly reduced noise using non-local patches. The cubic B-spline interpolation process slightly improved the reconstruction accuracy on SL, ATV, and NLTV, as well as the image clarification. All these results indicate that the proposed algorithm not only achieves a potential radiation dose reduction in the repeated CBCT scans, but also eliminates noise effectively without heavily compromising the striking features.

Technically, compared with the ATV transform, the proposed technique might increase the computational cost due to the larger search area when computing the weight function. In this work, to balance the reconstructed image quality with the computational burden, the search area was limited to  $21 \times 21$  pixels around each pixel. The extra computational burden is related to the time elapsed for the backprojector with cubic B-spline interpolation.

In our implementation, OpenMP was used to parallelize the proposed algorithm on an Intel Xeon CPU with 24 physical cores and 48 logical processors. Therefore, the proposed algorithm can be extended and accelerated by using a GPU<sup>28</sup>. The number of iterations for the NLTV minimization was determined to be ten, however, this value may not be optimal. Among the datasets used, we observed better reconstructed images with suppressed noise and preserved edges when the number of iterations was ten.

The same magnitude of imaging dose reduction can be achieved by reducing the number of projections while maintaining the mAs setting. However, this approach assumes fast gantry rotation. The maximum gantry speed of clinically used CBCT scanners is still limited to 1 rpm or 6°/s. Thus, low mAs acquisition may be a better option for imaging dose reduction than sparse view acquisition in practical aspects.

The NLTV denoising method can be extended to various applications. It can be used as an image fidelity term in compressed sensing-based approaches<sup>23</sup>. NLTV-based FDK reconstruction images can be used as an initial guess image for iterative reconstruction<sup>10,29</sup>. In general, it has been observed that the better the initial guess image when using an iterative reconstruction algorithm, the faster the convergence and CBCT images that can be created with a high contrast. In addition, a ray-driven backprojector can be combined with NLTV<sup>28</sup>.

In conclusion, an improved FDK reconstruction algorithm for low-dose CBCT imaging has been developed by adding non-local total variation (NLTV) denoising in the filtering process and a voxel-driven backprojector based on cubic-B spline interpolation. The proposed approach allows the conventional FDK algorithm to be used for low mAs image reconstruction without requiring iterative reconstruction algorithms, which are computationally demanding. Moreover, substantial reductions in radiation exposure associated with the low mAs projection acquisition can facilitate wider practical applications of IGRT using daily online CBCT imaging with corrections for setup errors.

## Data availability

All data generated or analyzed during this study are included in the article.

Received: 12 August 2020; Accepted: 28 January 2021

Published online: 11 February 2021

## References

- Nabavizadeh, N. *et al.* Image guided radiation therapy (IGRT) practice patterns and IGRT's impact on workflow and treatment planning: results from a national survey of American Society for Radiation Oncology members. *Int. J. Radiat. Oncol. Biol. Phys.* **94**, 850–857 (2016).
- Fuchs, F. *et al.* Interfraction variation and dosimetric changes during image-guided radiation therapy in prostate cancer patients. *Radiat. Oncol. J.* **37**, 127 (2019).
- Yoon, J., Park, K., Kim, J. S., Kim, Y. B. & Lee, H. Skin dose comparison of CyberKnife and helical tomotherapy for head-and-neck stereotactic body radiotherapy. *Prog. Med. Phys.* **30**, 1–6 (2019).
- Han, Y. Current status of proton therapy techniques for lung cancer. *Radiat. Oncol. J.* **37**, 232 (2019).
- Lu, L. *et al.* Intra- and inter-fractional liver and lung tumor motions treated with SBRT under active breathing control. *J. Appl. Clin. Med. Phys.* **19**, 39–45 (2018).
- Yang, Y., Schreiber, E., Li, T., Wang, C. & Xing, L. Evaluation of on-board kV cone beam CT (CBCT)-based dose calculation. *Phys. Med. Biol.* **52**, 685 (2007).
- Yoganathan, S. *et al.* Evaluating the image quality of cone beam CT acquired during rotational delivery. *Br. J. Radiol.* **88**, 20150425 (2015).
- Alaei, P. & Spezi, E. Imaging dose from cone beam computed tomography in radiation therapy. *Phys. Med.* **31**, 647–658 (2015).
- Wang, J., Li, T., Liang, Z. & Xing, L. Dose reduction for kilovoltage cone-beam computed tomography in radiation therapy. *Phys. Med. Biol.* **53**, 2897 (2008).
- Lee, H. *et al.* Improved compressed sensing-based cone-beam CT reconstruction using adaptive prior image constraints. *Phys. Med. Biol.* **57**, 2287 (2012).
- Chen, G. H., Tang, J. & Leng, S. Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med. Phys.* **35**, 660–663 (2008).
- Feldkamp, L., Davis, L. & Kress, J. Practical cone-beam algorithm. *J. Opt. Soc. Am. A* **1**, 612–619 (1984).
- Ahmadi, R., Farahani, J. K., Sotudeh, F., Zhaleh, A. & Garshasbi, S. Survey of image denoising techniques. *Life Sci. J.* **10**, 753–755 (2013).
- Motwani, M. C., Gadiya, M. C., Motwani, R. C. & Harris, F. C. Survey of image denoising techniques in *Proceedings of GSPX* 27–302004.
- Yuan, N. *et al.* Convolutional neural network enhancement of fast-scan low-dose cone-beam CT images for head and neck radiotherapy. *Phys. Med. Biol.* **65**, 035003 (2020).
- Chen, L., Liang, X., Shen, C., Jiang, S. & Wang, J. Synthetic CT generation from CBCT images via deep learning. *Med. Phys.* **47**, 1115–1125 (2020).
- Kida, S. *et al.* Cone beam computed tomography image quality improvement using a deep convolutional neural network. *Cureus* **10**, e2548 (2018). <https://doi.org/10.7759/cureus.2548>.
- Chen, H. *et al.* Low-dose CT via convolutional neural network. *Biomed. Opt. Express* **8**, 679–694 (2017).
- Lee, H., Lee, J., Kim, H., Cho, B. & Cho, S. Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction. *IEEE Trans. Radiat. Plasma Med. Sci.* **3**, 109–119 (2018).
- Liang, X. *et al.* Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. *Phys. Med. Biol.* **64**, 125002 (2019).
- Sidky, E. Y., Lorente, I., Brankov, J. G. & Pan, X. Do CNNs solve the CT inverse problem. *IEEE Trans. Biomed. Eng.* (2020). <https://doi.org/10.1109/TBME.2020.3020741>.
- Sidky, E. Y. & Pan, X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys. Med. Biol.* **53**, 4777 (2008).
- Kim, H. *et al.* Non-local total-variation (NLTV) minimization combined with reweighted L1-norm for compressed sensing CT reconstruction. *Phys. Med. Biol.* **61**, 6878 (2016).
- Kim, K., El Fakhri, G. & Li, Q. Low-dose CT reconstruction using spatially encoded nonlocal penalty. *Med. Phys.* **44**, e376–e390 (2017).

25. Sohn, J. J. *et al.* Analytical low-dose CBCT reconstruction using non-local total variation regularization for image guided radiation therapy. *Front. Oncol.* **10**, 242 (2020).
26. Lee, H., Yoon, J. & Lee, E. Anisotropic total variation denoising technique for low-dose cone-beam computed tomography imaging. *Prog. Med. Phys.* **29**, 150–156 (2018).
27. Perona, P. & Malik, J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 629–639 (1990).
28. Park, H.-G., Shin, Y.-G. & Lee, H. A fully GPU-based ray-driven backprojector via a ray-culling scheme with voxel-level parallelization for cone-beam CT reconstruction. *Technol. Cancer Res. Treat.* **14**, 709–720 (2015).
29. Lee, H., Fahimian, B. P. & Xing, L. Binary moving-blocker-based scatter correction in cone-beam computed tomography with width-truncated projections: proof of concept. *Phys. Med. Biol.* **62**, 2176 (2017).

### Author contributions

Conception, design, and drafting of the manuscript were performed by H.L., B.J.M., and I.J.L. Data collection and interpretation were performed by J.P. and Y.C., K.R.P. reviewed the contents. All authors read and approved the final manuscript.

### Funding

This study was supported by the 2020 Research Grant of Gangnam Severance Hospital Research Committee (DHHD000067), by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (2019R111A1A01062157), and by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health & Welfare, Republic of Korea (HI19C1330).

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to B.J.M. or I.J.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021