Prior-Based Enhanced ASD-POCS for Artifact Suppression and Structural Preservation in Sparse-View CBCT

D M Bappy
Incheon National University
Dept. of Computer Science and Engineering
Incheon, Republic of Korea
dewan8511@gmail.com

Donghwa Kang
Incheon National University
Dept. of Computer Science and Engineering
Incheon, Republic of Korea
anima0729@inu.ac.kr

Jinkyu Lee
Sungkyunkwan University
Dept. of Computer Science and Engineering
Suwon, Republic of Korea
jinkyu.lee@skku.edu

Hyeongboo Baek*
Incheon National University
Dept. of Computer Science and Engineering
Incheon, Republic of Korea
hbbaek@inu.ac.kr

ABSTRACT

Among CT (Computed Tomography) techniques that produce cross-section images by acquiring X-ray data from multiple angles around the individual, CBCT (Cone Beam CT) that collects the data through cone beam is popular due to its potential to reduce radiation risks. Although employing low-dose and sparse-view CBCT is a cornerstone approach to reducing radiation risks, reconstructing CBCT from noisy and sparsely-acquired cone beam scans often result in significant artifacts due to ill-posed inverse problem.

In this paper, we aim to suppress the artifacts derived from reconstructing CBCT. To this end, we target the well-known existing approach ASD-POCS (Adaptive Steepest Descent with Projections Onto Convex Sets), focus on its downside of potentially introducing over-smoothing near edges and removing intricate fine structures, and develop an advanced ASD-POCS strategy that addresses the downside by leveraging prior image information. Our approach to integrating the prior information not only maintains structural integrity (without compromising edge detail or erasing fine structures) but also diminishes artifacts, thereby counteracting noise and blur to a significant extent. Our solution underwent stringent testing using both simulated data and publicly available datasets. The testing results demonstrate that the proposed technique is robust to the prevailing challenges in sparse-view CBCT reconstruction, skillfully mitigating artifacts (compared to contemporary state-of-the-art methods) while safeguarding intricate structures.

CCS CONCEPTS

• **Applied computing** → **Imaging**; **Bioinformatics**; **Health informatics**.

*Hyeongboo Baek is the corresponding author of this paper.

SAC '24, April 8–12, 2024, Avila, Spain

ACM ISBN 979-8-4007-0243-3/24/04…$15.00
https://doi.org/10.1145/3605098.3635910

KEYWORDS

Computed Tomography, Sparse-view, Low-dose, Artifacts, FISTA

ACM Reference Format:
D M Bappy, Donghwa Kang, Jinkyu Lee, and Hyeongboo Baek. 2024. Prior-Based Enhanced ASD-POCS for Artifact Suppression and Structural Preservation in Sparse-View CBCT. In *The 39th ACM/SIGAPP Symposium on Applied Computing (SAC '24), April 8–12, 2024, Avila, Spain.* ACM, New York, NY, USA, Article 4, 10 pages. https://doi.org/10.1145/3605098.3635910

1 INTRODUCTION

Tomographic methodologies are differentiated by their data acquisition methods. If radiation sources are positioned within the patient, emission data is recorded. In contrast, when the sources are external, the system collects transmission data. Our study primarily centers on transmission data. Computed tomography (CT) uses this data type to generate cross-sectional imagery of a patient's body, necessitating a substantial radiation amount [18]. Of the prevalent X-ray CT techniques, Cone Beam Computed Tomography (CBCT) is increasingly recognized across various fields, spanning from medical imaging to material science. The capability to reconstruct comprehensive cross-section images with a minimized X-ray radiation dose is a pivotal advancement for CBCT in the medical realm, leading to enhanced volume reconstruction in micro-CT [21]. This has found applications in areas such as maxillofacial imaging [9], radiation therapy guidance for cancer treatment [7], insect visualization [13], and material science investigations [3]. Regardless of the specific CBCT application, the fundamental process remains consistent: capturing 2D X-ray images from varying angles of the object, and then employing a tomographic reconstruction algorithm to form a cross-section image from the aggregated data.

CBCT has become a versatile tool utilized across various domains, including industrial and medical sectors. In the scope of CT applications, the reconstruction process is significantly influenced by critical factors such as the specific scanning environment [31], the nuances of imaging geometry [34], and the exact dosage of X-ray radiation deployed [33]. Despite the accolades received for its fast scanning capabilities and remarkable spatial resolution, CBCT technology faces critical examination over its utilization of X-rays, associated with potential health risks including cancer and genetic disorders [30]. Consequently, there is an imperative drive among

both medical professionals and patients to discover strategies to reduce the radiation dose involved in CBCT procedures. This endeavor chiefly pursues two directions: lowering the product of the tube current and exposure time (mAs) or strategically decreasing the number of projection views [12]. The former pathway successfully reduces the dose but introduces direct noise into the projection data. In contrast, the latter strategy, while curtailing the dose, brings forth a unique set of challenges; notably, the limited observational data culminates in the appearance of artifacts in the reconstructed model [17]. Our study navigates this intricate landscape, focusing on the nuanced task of reconstructing CBCT images from sparsely acquired projections.

Early advancements in the domain saw the introduction of analytical techniques, encompassing the seminal FDK method [11] along with its evolved iterations [28], for a comprehensive view of CBCT reconstruction. These methods, albeit pioneering, grapple with significant artifacts when tasked with sparse view reconstructions, largely stemming from inadequate projection data. To counter the shortcomings intrinsic to these analytical solutions, the scholarly community has pivoted towards iterative algorithms. This shift perceives CBCT reconstruction not as a straightforward analytical problem but as an ill-posed linear issue, necessitating a more nuanced and iterative approach for resolution.

The task of reconstructing images from partial projection data is predominantly tackled through model-based iterative reconstruction strategies, including renowned algorithms such as the Algebraic Reconstruction Technique (ART) [14] and its counterpart, the Simultaneous Algebraic Reconstruction Technique (SART) [2]. These iterative avenues often infuse additional image details into the objective function, nurturing a solution that stands both convergent and robust. A case in point is the recent advent of the Compressed Sensing (CS) paradigm [8], which argues that a signal—provided it maintains a sparse or compressible representation—can be faithfully recovered from a limited set of measurements with high likelihood.

In the domain of CS reconstruction, the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [6] is notable due to its streamlined complexity and enhanced efficiency. Despite its merits, applying FISTA to low-dose CT projections unveils a deficiency in reconstruction efficacy, a problem that chiefly originates from the pronounced noise levels native to low-dose CT projections. During FISTA's iterative cycle, the potential for information loss is high, catalyzing substantial reconstruction inaccuracies. Where images bear piecewise constant features, Total Variation (TV) minimization [26] has carved a name for itself, becoming a widely adopted solution for grappling with the hurdles posed by partial data in tomographic image construction. While TV minimization marks a stride forward in diminishing streak artifacts, it carries with it a set of shortcomings, including an over-smoothing effect on fine structures and introduction of blocky and caricatured artifacts — repercussions of TV's isotropic operations coupled with the intrinsic limitations set by incomplete projection datasets [29].

The Adaptive Steepest Descent with Projections Onto Convex Sets (ASD-POCS) algorithm as introduced by Sidky and Pan [27], operates through an optimization strategy focused on minimizing the Total Variation (TV) of the estimated image while adhering to data conditions and various constraints. This algorithm integrates ART iterations with gradient descent steps targeting the TV penalty, thus aiming to solve the intricate problem of constrained minimization. However, it's worth noting that the ASD-POCS algorithm doesn't inherently include a prior image as part of the reconstruction process. Moreover, this technique can lead to excessive smoothing near edges when dealing with noisy sparse projections. Prior methodologies often necessitated a trade-off between these qualities, requiring compromises that either compromised edge clarity for artifact reduction or vice versa. However, in the context of accurate diagnosis, all these attributes—sharp edges, fine details, and minimal artifacts—carry the utmost significance. This underscores the pivotal importance of attaining a balanced reconstruction approach that successfully upholds these diverse qualities.

We introduce a novel approach that utilizes prior information to achieve enhanced reconstruction, drawing inspiration from the promising attributes of FISTA reconstruction. This choice is anchored in the FISTA technique's capacity to provide a robust approximation for our enhanced ASD-POCS algorithm. Such an approximation is a beneficial starting point for our algorithm, given its capability to minimize artifacts, even amid noise from sparse projections. In our proposed method, we integrate a small-sized median filter into the ART estimate beforehand to eliminate noise originating from low-dose projections and prior FISTA reconstruction. Following this, we employ unsharp masking, using the smoothed reconstruction from Filtered ART, to counteract the blur introduced during the reconstruction process. The synergy of these modifications leads to a superior-quality reconstruction that skillfully maintains sharp edges and detailed structures while effectively diminishing pronounced artifacts and noise.

We evaluated our method against two prominent techniques: the state-of-the-art iterative approach by [5] and the recent self-supervised Convolutional Neural Network (CNN) technique from [20]. The former, tested on a digital brain phantom [1], aims to reconstruct CT from sparse views. It acquires interpolated projection views through optical flow-based frame interpolation and then combines these for the final reconstruction. However, it operates on 4-degree intervals for interpolation, while we utilize a wider 11-degree gap, which caused inaccuracies during frame interpolation and deteriorate reconstruction quality. The latter technique, tested on a clinical dataset [22], employs a self-supervised CNN to mitigate streak artifacts in sparse-view CT images. They train their CNN on downsampled views with pronounced artifacts, generating prior images to rectify original artifacts. Notably, their approach chiefly targets streak artifacts, leaving CT noise unaddressed. Contrarily, our method simultaneously suppresses both noise and artifacts, leading to a notably superior reconstruction.

Our contributions in this paper encompass the following:

- We introduce a pioneering approach that relies on prior image-based enhancements within the context of the enhanced ASD-POCS technique, yielding superior reconstructions in low-dose sparse-view scenarios.
- We establish the value of utilizing the prior FISTA reconstruction as potent prior information, enriching the enhanced ASD-POCS methodology.

- We introduce filtered ART, which incorporates a small-sized filter, effectively reducing the initial noise typically found in low-dose CT acquisitions.
- Our innovation involves introducing unsharp masking. By utilizing the filtered reconstruction from the filtered ART step, we enhance the reconstruction quality and reduce blur.

In sum, our approach represents a meaningful stride towards achieving high-quality reconstruction, noise, and artifact suppression within the intricate domain of sparse-view reconstructions.

2 METHOD

2.1 X-ray CT measurement

The CBCT's geometric configuration can be visualized as depicted in Figure 1. An X-ray source, is positioned at a *DSO* distance from a rotation center, denoted as *O*. This point *O* also marks the origin of a specific coordinate system. When the X-ray source is activated, it emits a cone of rays that encompasses the image volume. As these rays penetrate this region, a detector captures the photon intensity that reaches it. The photons' intensity diminishes based on the Beer-Lambert law. The detector is situated at a *DSD* distance from the X-ray source, and its central point is marked as C. A separate projection coordinate system, referred to as *st* is based on the detector's midpoint. As the data collection process unfolds, both the source and the detector orbit the y-axis, deviating at an angle from their starting positions.

The X-ray CT measurement can be represented approximately as a discrete linear system [15]:

$$\boldsymbol{p} = A\boldsymbol{\mu} \tag{1}$$

In the context of X-ray CT, we can express the measurement process using a discrete linear framework [15]. Here, the vector of attenuation coefficients is represented by $\boldsymbol{\mu}$, such that $\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_N)^T$, where the T superscript indicates the matrix transpose operation. Additionally, the obtained sinogram data, which corresponds to the calibrated and logarithmically transformed projections, is denoted by $\boldsymbol{p} = (p_1, p_2, ..., p_M)^T$, where the T superscript indicates the matrix transpose operation. The operator A signifies the system or projection matrix with dimensions $M \times N$, where M and N represents the number of elements in $\boldsymbol{p}$ and $\boldsymbol{\mu}$ respectively. The element A_{ij} specifies the length of intersection between projection ray i and pixel j. In our implementation, we compute this element using an efficient ray-tracing technique [25]. The core objective of CT image reconstruction revolves around estimating the attenuation coefficients $\boldsymbol{\mu}$ from the measurements $\boldsymbol{p}$. In an ideal scenario, the image reconstruction task is equivalent to solving the inversion problem presented by the linear system Equation (1), aimed at determining $\boldsymbol{\mu}$ based on the dataset $\boldsymbol{p}$. However, even when conditions are ideal, the solution to this linear system can be complicated by the ill-conditioned nature of the system matrix A, to be detailed now.

The ill-conditioned nature of the system matrix can arise due to two primary factors: inadequate coverage within the scanning configuration, which encompasses scenarios like projection data truncation or incomplete X-ray source trajectories; and under-sampling, which includes situations involving an insufficient number of projection views or low-resolution detectors. In the context of circular cone-beam scanning configurations, both of these contributors to ill-conditioning come into play. Specifically, when examining circular scanning trajectories under the constraint of continuous sampling, achieving a stable one-shot reconstruction for the volume image becomes unattainable. This is compounded by the fact that most points within the image fail to satisfy Tuy's condition, further complicating the reconstruction process.

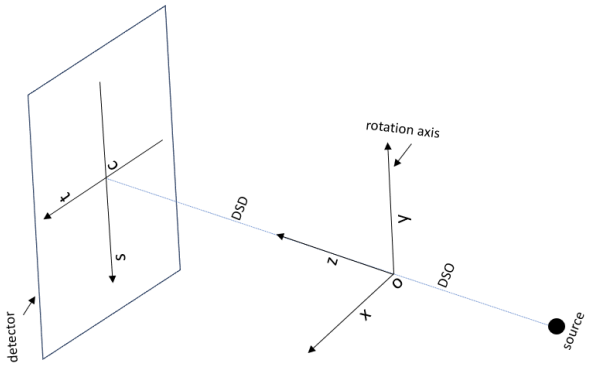

Figure 1: CBCT geometry

2.2 ASD-POCS

The ASD-POCS technique offers a solution to the aforementioned reconstruction challenge through the resolution of the subsequent constrained optimization problem:

$$\boldsymbol{\mu}^* = argmin \parallel \boldsymbol{\mu} \parallel_{TV} \tag{2}$$

subject to inequality constraints: data fidelity

$$\mid A\boldsymbol{\mu} - \boldsymbol{p} \mid \leq \epsilon \tag{3}$$

and non-negativity

$$\boldsymbol{\mu} \geq 0 \tag{4}$$

$\parallel \boldsymbol{\mu} \parallel_{TV}$ is defined as:

$$\begin{aligned} \parallel \boldsymbol{\mu} \parallel_{TV} = \sum_{s,t,v} \Big(& (\boldsymbol{\mu}_{s,t,v} - \boldsymbol{\mu}_{s-1,t,v})^2 \\ & + (\boldsymbol{\mu}_{s,t,v} - \boldsymbol{\mu}_{s,t-1,v})^2 \\ & + (\boldsymbol{\mu}_{s,t,v} - \boldsymbol{\mu}_{s,t,v-1})^2 \Big)^{1/2} \end{aligned} \tag{5}$$

Here, the indices s, t, and v denote the positions within the discrete image.

The ASD-POCS algorithm integrates a crucial step designed to enhance data consistency, where the core projection operation enforces non-negativity. Notably, ASD-POCS undertakes an iterative approach to minimize the Total Variation (TV) norm, signifying that image reconstruction precedes the reduction of the TV norm in each iteration. To guide the image toward a solution with minimal TV, the algorithm interleaves POCS steps with TV-steepest descent. Solely executing the TV-minimization step throughout the algorithm would result in a uniformly flat image. In contrast, the

ROF model ensures that significant changes are not imposed on the image. The importance of selecting optimal parameters for image quality has been emphasized in previous studies [10]. However, the performance of ASD-POCS deteriorates due to the presence of noise originating from low-dose projections and the amplification of artifacts stemming from extremely sparse views. Additionally, as a result of increased noise and artifacts, this approach tends to eliminate fine details through an excessive smoothing process during the TV minimization steps.

2.3 Reconstruction Algorithm

In this section, we present our proposed technique. Our method is designed to reconstruct CT images from low-dose and sparse projections, challenges that often introduce significant noise and artifacts into the reconstruction. To address these issues, we employ a multi-step approach. First, we utilize prior reconstruction information from FISTA to initially suppress artifacts. This initial reconstruction is then subjected to a filtered ART step to further reduce noise. Following this, we apply unsharp masking to enhance the reconstruction by reducing blur. Finally, we employ a steepest descent TV step to minimize the TV norm and further reduce artifacts. This iterative process continues until we achieve the minimum TV, resulting in a high-quality CT.

Figure 2 illustrates key stages of our proposed technique and visualizes the enhancement in reconstruction at each step. Our proposed approach commences with an initial estimate reconstructed through FISTA by solving Equation (6).

$$\boldsymbol{\mu}_{FISTA} = argmin_{\boldsymbol{\mu}} \frac{1}{2} \parallel \boldsymbol{p} - A\boldsymbol{\mu} \parallel_2^2 + \lambda_F \parallel \boldsymbol{\mu} \parallel_1 \tag{6}$$

In this expression, λ_F stands for the regularization parameter, and $\parallel . \parallel_1$ symbolizes the l_1 norm. This initial estimate offers a representation with a resemblance to the underlying structure while also reducing artifacts, albeit retaining some noise.

To mitigate this noise, we introduce filtered ART in which the noise from low-dose projections is suppressed using a small-sized filter. This provides an improved residual estimate for the steepest descent TV.

$$\boldsymbol{\mu} = \boldsymbol{\mu} + \lambda A_i \frac{\boldsymbol{p} - A_i\boldsymbol{\mu}}{A_i.A_i} \tag{7}$$

$$\boldsymbol{\mu}_{fil}(s,t,v) = median\{\boldsymbol{\mu}(s+l, t+m, v+n) | (l,m,n) \epsilon R\} \tag{8}$$

During, TV minimization and filtering of the noise in the main loop the reconstruction was affected by blurring. We perform unsharp masking to enhance the reconstruction quality.

$$\boldsymbol{\mu} = \boldsymbol{\mu} + (\boldsymbol{\mu} - \boldsymbol{\mu}_{fil}) \tag{9}$$

The optimization problem is defined by the projection data $\boldsymbol{p}$ and the data-inconsistency tolerance parameter ϵ. There are six parameters and two filters that govern the proposed algorithm.

The Prior-Based Enhanced ASD-POCS technique is presented in the form of pseudocode in Algorithm 1. Initial parameters are set during the parameter initialization phase (in lines 1-2). Then, we reconstruct and initiate with FISTA (in lines 3-4) due to its pertinent initial structure and, crucially, its minimal artifacts.

Algorithm 1 Prior-Based Enhanced ASD-POCS

```
   Parameter initialization
1: λ = 1, λ_red = 0.99, η = 15, γ_max = 0.94;
2: β = 0.002, β_red = 0.95; k = 10;
   Prior reconstruction
3: μ_FISTA = initial reconstruction using FISTA
4: μ = μ_FISTA
   Repeat main loop
5: μ_0 = μ
   Filtered ART
6: for i = 1 : N_d do
7:    μ = μ + λA_i (p − A_i μ)/(A_i.A_i)
8:    μ_fil = median_filter(μ)
   Enforce positivity
9:    if μ_i < 0 then
10:      μ_i = 0
11:   end if
12: end for
   Unsharp masking
13: μ = μ + (μ − μ_fil)
14: μ_res = μ
15: p = Aμ
16: dd = |p − p_0|
17: dp = |μ − μ_0|
18: if iter == 1 then
19:    dtvg = β * dp
20: end if
21: μ_0 = μ
   TV steepest descent
22: for i = 1 : η do
23:    dμ = || μ ||_TV
24:    d̂μ = dμ/|dμ|
25:    μ = μ − dtvg * d̂μ
26: end for
   Residual
27: dg = |μ − μ_0|
28: if dg ≥ γ_max and dd > ϵ then
29:    dtvg = dtvg * β_red
30: end if
31: λ = λ * λ_red
   until {stopping criteria}
32: return μ_res
```

The primary loop (in lines 5-31) comprises two fundamental elements: steps that ensure data consistency within the loop, and a steepest descent towards images with lower TV. The essence of the algorithm lies in managing the step sizes for both the main loop and the TV steepest descent. The reconstructed image vector $\boldsymbol{\mu}_0$ (in lines 5 and 21) serves as a temporary variable for the computation of alterations in the current estimate. During the filtered ART step (in lines 6-12), low-dose CT noise is effectively suppressed, while ensuring data consistency and maintaining positivity. This control over noise reduction and data fidelity is governed by the relaxation parameter λ. The current image reconstruction (in line 14), denoted as $\boldsymbol{\mu}_{res}$, is attained by applying unsharp masking (in line 13) to mitigate the blurring introduced during the TV steepest descent step. The residual data is recalculated (in lines 15-16), while the alteration in the image attributed to the filtered ART is determined

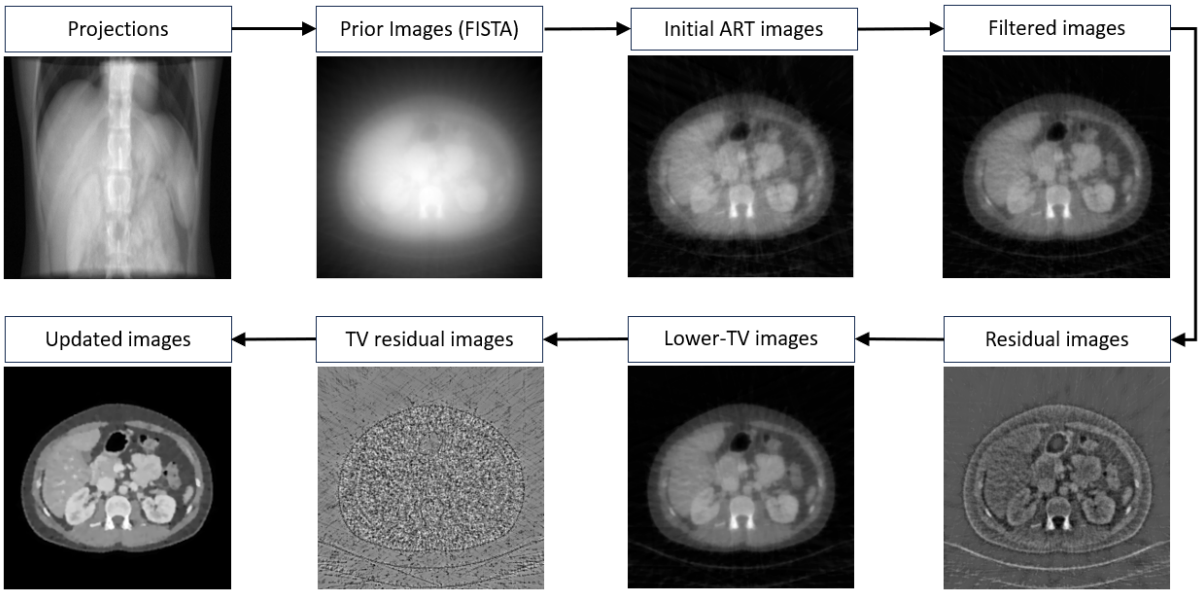

Figure 2: Schematic of the proposed technique

(in line 17). Line 19 serves the purpose of converting the initial steepest-descent step size, originally expressed as a fraction of a step size, into an absolute image distance during the first iteration. The implementation of TV steepest descent can be observed within lines 22-26. The step size for the TV steepest descent evolves from a fraction of the initial step size to an absolute image distance as the iteration progresses. The TV steepest descent-induced changes in the reconstructed image are assessed (in line 27). In cases where the ratio of the change in the reconstructed image due to the steepest descent to the change in the image caused by the main loop surpasses the threshold γ_{max}, the gradient descent step size is adjusted downward by a factor of β_{red} (in line 29). By regulating the steepest-descent step size in this manner, the current image estimate meets the data-tolerance condition. In the end, the filtered ART-relaxation parameter is decreased by a consistent fraction represented by λ_{red} (in line 31). This specific parameter governs the overall iteration count, as it influences the step size of data consistency. Consequently, this also impacts the size of the steepest-descent step.

The stopping criteria encompass several checks. Initially, the current reconstruction estimate undergoes scrutiny to determine whether it complies with the constraints outlined in Equations (3) and (4). Subsequently, the parameter c is evaluated, as defined in Equation (10), to ascertain its proximity to the value of -1.0. Finally, the process is halted if the λ value is found to be excessively small.

$$c = \frac{dg \cdot dp}{|dg||dp|} \tag{10}$$

In an ideal scenario, the value of c should be exactly -1.0. However, in practical applications, achieving this precise value is challenging due to the substantial number of iterations required. Based on our extensive numerical studies, we have observed that c is an exceptionally sensitive parameter. It is common for the image to exhibit only barely perceptible changes once c falls below -0.5.

The filtered ART operator relies on the relaxation parameter λ. The ART-relaxation parameter undergoes reduction by a constant fraction λ_{red}. The TV hyperparameter is denoted as β. Variables β_{red} and γ_{max} are utilized to steer the evolution of β. The variable η represents the iteration number for the TV-steepest descent and k represents the number of iterations for the main algorithm. Image-space variables are signified by the bold notation (e.g., $\boldsymbol{\mu}$), while data-space variables are denoted with a tilde (e.g., $\tilde{\boldsymbol{p}}$). The vector A corresponds to the row of the system matrix responsible for producing the i-th data element. $\tilde{\boldsymbol{p}}_0$ defines the optimization problem specified by the projection. Lastly, ϵ represents the data-inconsistency-tolerance parameter.

3 ANALYSIS AND FINDINGS FROM EXPERIMENTAL INVESTIGATIONS

3.1 Implementation and Data

In our investigation, we assess the efficacy of the proposed method for low-dose and sparse-view CT reconstruction. Our experimentation is conducted on a Windows 10 x64 platform, utilizing MATLAB's image processing and computer vision toolbox. The hardware configuration includes an Intel i5-9400K processor running at 2.90 GHz with dual cores, accompanied by 16 GB of RAM. The algorithm is implemented using MATLAB 2022 for the evaluation process.

The validation procedure utilizes the digital brain phantom, constructed from authentic MRI data of the human brain, and serves as a widespread benchmark for evaluating CBCT reconstruction techniques [1]. This phantom consists of 256×256×256 voxels, with each voxel possessing a resolution of 1×1×1 mm^3. For assessing performance under conditions of both low-dose and sparse-view scenarios, we employ the actual clinical dataset from the 2016 AAPM

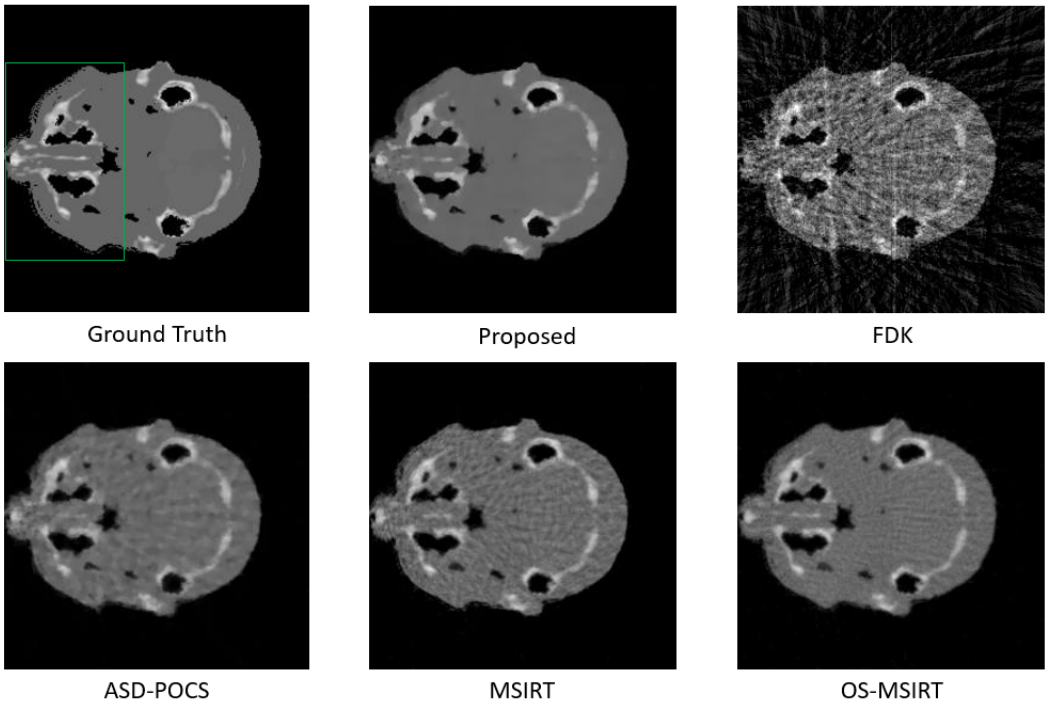

Figure 3: Brain phantom reconstruction using recent techniques

Low-Dose CT Grand Challenge [22], provided by the Mayo Clinic. We create the projection data within a cone-beam geometry system using Siddon's ray-driven algorithm [25]. This dataset is already infused with realistic noise.

Figure 1 illustrates the CBCT geometry employed for object scanning. The detailed descriptions of the CBCT geometry system for image generation are explained as follows. The distances from the source to the iso-center and from the source to the detector were 1000 and 1536 mm, respectively. For sparse-view CT, the number of projection views was 256, 128, and 64. The rotation angle for each set of projection views covered a complete 360 degrees.

3.2 Assessment of Image Quality

To assess the reconstruction outcomes, we employ four qualitative metrics: the root mean square error (RMSE), normalized root mean square error (NRMSE), structural similarity index (SSIM), and peak signal-to-noise ratio (PSNR).

The RMSE is expressed as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{\mu} - \boldsymbol{\mu}_{ref})^2} \tag{11}$$

The NRMSE is expressed as

$$NRMSE(\boldsymbol{\mu}, \boldsymbol{\mu}_{ref}) = \frac{RMSE}{max(\boldsymbol{\mu}_{ref}) - min(\boldsymbol{\mu}_{ref})} \tag{12}$$

In this context, $\boldsymbol{\mu}$ represents the reconstructed image, while $\boldsymbol{\mu}_{ref}$ stands for the reference image.

SSIM evaluates the visual elements of an image in terms of luminance, contrast, and structure, and it is defined as:

$$SSIM(\boldsymbol{\mu}, \boldsymbol{\mu}_{ref}) = \frac{2\boldsymbol{\mu}\boldsymbol{\mu}_{ref} + c_1}{\boldsymbol{\mu}^2 + \boldsymbol{\mu}_{ref}^2 + c_1} \cdot \frac{2\sigma_{\boldsymbol{\mu}\boldsymbol{\mu}_{ref}} + c_2}{\sigma_{\boldsymbol{\mu}}^2 + \sigma_{\boldsymbol{\mu}_{ref}}^2 + c_2} \tag{13}$$

Here, $\delta_{\boldsymbol{\mu}}$ represents the average value of $\boldsymbol{\mu}$, and $\sigma_{\boldsymbol{\mu}}$ is its variance. The terms corresponding to $\boldsymbol{\mu}$ are defined analogously. The term $\sigma_{\boldsymbol{\mu}\boldsymbol{\mu}_{ref}}$ denotes the covariance between $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_{ref}$. The constants c_1 and c_2 are defined as $(0.01L)^2$ and $(0.03L)^2$ respectively, where L indicates the maximum pixel value in $\boldsymbol{\mu}_{ref}$.

The PSNR measures the ratio in decibels between two images. This ratio is commonly utilized to compare the quality of the original image to the resulting image. A higher PSNR value indicates the superior quality of the output image. To determine the PSNR, we employ the MSE. The PSNR is computed as

$$PSNR(\boldsymbol{\mu}, \boldsymbol{\mu}_{ref}) = 20 * log(\frac{max(\boldsymbol{\mu}_{ref})}{RMSE}) \tag{14}$$

In this context, $max(\boldsymbol{\mu}_{ref})$ represents the highest potential value of the reference image, while MSE denotes the mean squared error between the reference and the reconstructed images.

3.3 Digital Brain Phantom Reconstruction

Siddon's ray-driven algorithm [25] is employed to simulate projection data from 32 sparse views. For sparse CT reconstruction, recent iterative techniques like MSIRT [4] and OS-MSIRT [5] are utilized. In addition, two classical methods, namely FDK and ASD-POCS, are

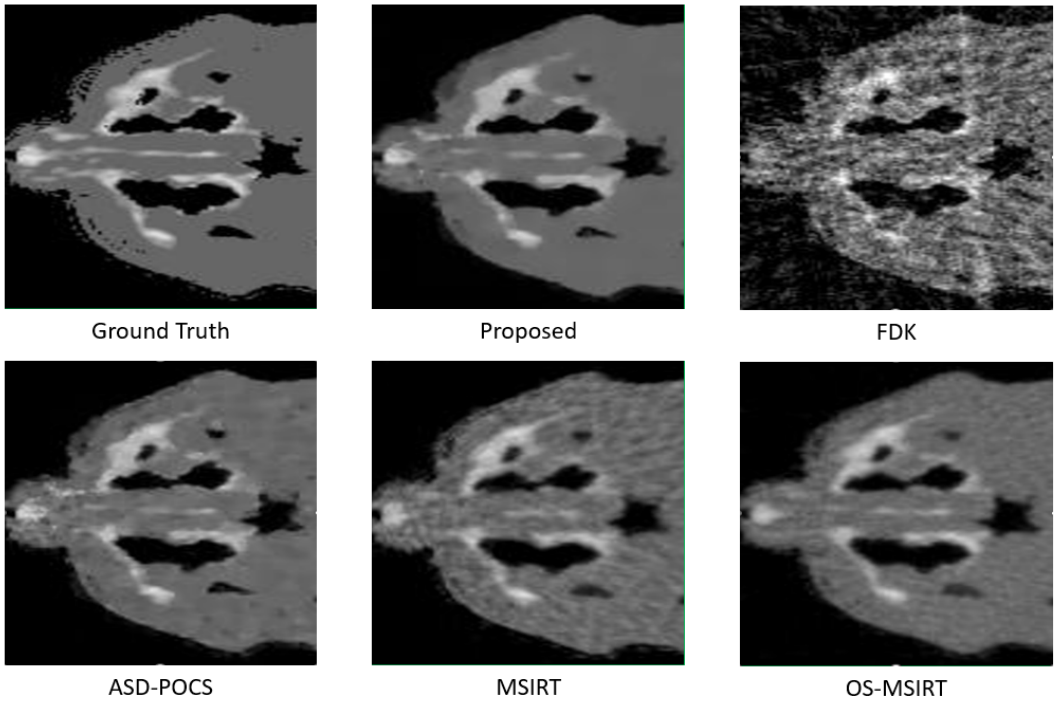

Figure 4: Zoom in on ROI in Figure 3

included for comparative analysis, as discussed in the Introduction section.

The reconstruction results, employing both classical and recent techniques, are depicted in Figure 3. We specifically selected the 66th slice in the horizontal direction for comparison due to its rich structural information. In all cases, 32 projections were used for reconstruction. The FDK method, designed for full-view reconstruction, fails to produce a clinically acceptable image from the limited 32 projection data, resulting in significant streak artifacts. On the other hand, the ASD-POCS method effectively removes artifacts and noise, but it tends to blur edges and omit critical structural details. The MSIRT maintains the structural integrity of the reconstructed slice but introduces noticeable artifacts, obstructing fine information. The OS-MSIRT slightly improves artifact reduction but still struggles to recover small structures while exhibiting blurry edges. Our proposed technique performs reconstruction using 32 projections. The reconstruction quality excels in noise and artifact suppression and accurately preserves fine structures.

For a more detailed comparison, we zoomed in on the region of interest (ROI) marked by the green box in the ground truth image in Figure 3, as shown in Figure 4. Upon magnification, it becomes evident that the proposed technique, restores nearly perfect edges. In contrast, the edges reconstructed by ASD-POCS appear smoother but noticeably lose structural details and may contain artifacts when compared to the ground truth in Figure 4.

3.3.1 Quantitative Analysis. The results of the quantitative evaluation are presented in Table 1. This table compares three qualitative metrics to assess the proposed reconstruction . In alignment with the visual inspection depicted in Figures 3 and 4, the proposed algorithm attains the highest scores for all three indices when utilizing 32 views.

3.4 Clinical Data Reconstruction

We utilize Siddon's ray-driven algorithm to generate projection data from 64, 128, and 254 sparse views. Our proposed method's reconstruction quality is then benchmarked against recent techniques such as SAR-SSNR [19] and Sparsier2Sparse [20].

In Figure 5, we showcase images derived from the reference, SAR-SSNR, Sparsier2Sparse, and our proposed method, all of which uses the clinical dataset. For a more detailed assessment of fine structures, artifacts, and noise, zoomed-in views of the Regions of Interest (ROIs) are highlighted with green boxes within Figures 5. For the 64 views, it is evident that while both SAR-SSNR and Sparsier2Sparse can reconstruct structures akin to the reference image, they are not devoid of noise and artifacts. On the other hand, our proposed method demonstrates a superior ability to mitigate

Table 1: Quantitative analysis for digital brain phantom

Methods	RMSE	PSNR	SSIM
FDK	0.1263	17.9711	0.1479
ASD-POCS	0.0439	27.1528	0.8423
MSIRT	0.0396	28.0414	0.8154
OS-MSIRT	0.0383	28.3311	0.8265
Proposed	**0.0315**	**30.0436**	**0.9270**

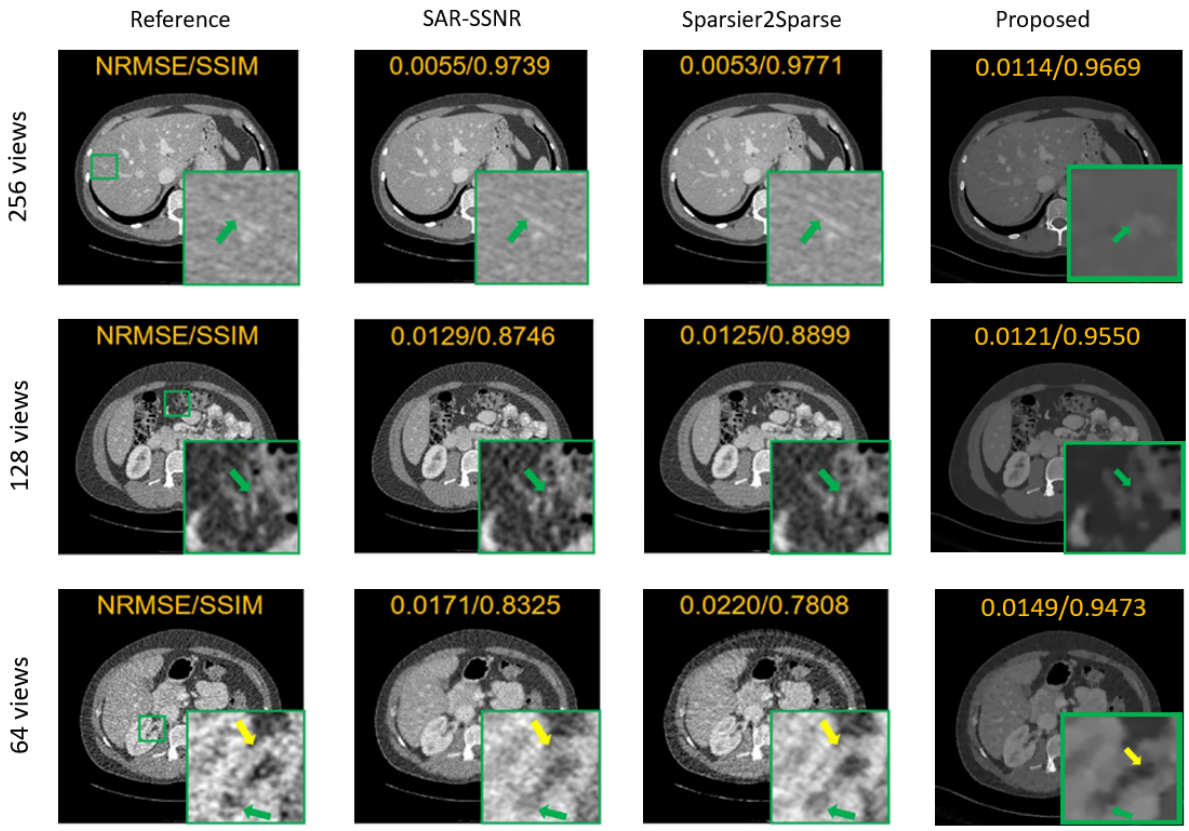

Figure 5: Clinical dataset reconstruction

both noise and artifacts, offering a crisper structural representation. However, when evaluating the 128 and 256 views, the proficiency of SAR-SSNR and Sparsier2Sparse in artifact reduction improves. Despite this enhancement, our method still surpasses both in terms of reducing noise and artifacts, as well as preserving fine details.

3.4.1 Quantitative Analysis. In Figure 5, we have also presented quantitative results together with specific reconstruction techniques, specifically the NRMSE and SSIM scores. Upon analyzing these scores, it becomes evident that our proposed technique excels in both 64 and 128 views. However, when it comes to 256 views, there is a slight dip in the NRMSE and SSIM. In general, our evaluation shows that our proposed method significantly outperforms recent techniques in quantitative assessments, especially when the number of projections are very sparse.

4 RELATED WORK

Techniques aimed at sparse-view reconstruction, including API-TV, as suggested in [16], leverage a prior image to serve as the inaugural estimate for the ASD-POCS algorithm, significantly aiding in the reduction of artifacts in sparse-view data contexts. However, this methodology harbors a substantial downside: it necessitates a prior scan from the identical patient to furnish the requisite prior information, a stipulation that potentially curtails its practical utility. As a parallel development, the streamlined ASD-POCS method highlighted in [24] employs a simplified matrix-form depiction of the TV gradient, maintaining the reconstruction precision synonymous with the original ASD-POCS initiative. A standout feature of this strategy is its capacity to bolster computational velocity, accelerating processes by a factor ranging between 1.8 and 2.7 times. Despite these forward strides, the overarching academic discourse still grapples with an unresolved critical issue: orchestrating a reconstruction approach that not only safeguards sharp edges and delicate fine structures but also actively diminishes the severe artifacts born from sparse projection datasets.

In the realm of sparse-view CT reconstruction, a diverse array of deep learning (DL) methodologies has been explored. Among them, the research delineated in [32] presented an innovative technique intertwining deep embedding-attention-refinement modules with a residual error feedback mechanism, operational in both sinogram and image domains. Concurrently, other scholars have embarked on crafting the MIST-net, spotlighted in [23], which addresses the issue of streak artifacts through the dual training of a convolutional neural network (CNN) and a transformer-based network, enhancing the capture of long-range features. It is worth noting that CNNs, with their restricted receptive fields, predominantly concentrate on local image details. These strides in technology have indeed showcased encouraging outcomes; however, they largely hinge on a fully supervised learning paradigm. This approach demands pairs of sparse-view and full-view CT images showcasing identical

anatomical structures to facilitate network training. This prerequisite poses a significant roadblock in genuine clinical scenarios, given that executing two distinct scans on patients — one embracing full views and another accommodating sparse views — fundamentally violates the ALARA (as low as reasonably achievable) principle that governs radiological safety. Moreover, discrepancies in patients' positioning across separate scans can potentially induce movement, culminating in subpar training data. Offering a divergent pathway, some propose the synthesis of sparse-view CT images via computer simulations of CT geometry, an approach viable when full-view CT images are at hand. Nevertheless, this strategy hits a bottleneck owing to the occasional unavailability of high-resolution full-view CT images, thus throwing a spanner in the works of a seamless reconstruction process.

5 CONCLUSION

In this research, we introduced an enhanced ASD-POCS technique that leverages prior images for the purpose of reconstructing CT scans from sparsely available low-dose projections. Our primary objective was to effectively mitigate noise and artifacts arising from the reduced dose and a limited number of projections while ensuring the accurate recovery of intricate details.

To assess the efficacy of our proposed technique, we conducted reconstructions using both digital phantom data and real clinical data. The clinical data, in particular, posed a greater challenge due to its inherent noise stemming from the lower dose levels. Our technique exhibited impressive performance in terms of noise and artifact suppression, alongside the accurate reconstruction of fine structures when compared to recent methods.

For a more comprehensive evaluation, we employed quantitative metrics such as the NRMSE and SSIM for the clinical dataset. These measurements clearly indicate that our technique surpasses recent approaches by a considerable margin, especially when considering 64 and 128 views. From the reconstruction results using 256 views, it's evident that our method performs slightly below the learning-based techniques. This is understandable, as learning-based approaches benefit considerably from robust training with 256 views. The training data at this view count exhibits minimal artifacts, which offers a more precise prior approximation for the learning-based method.

Our technique demonstrates the ability to achieve good-quality reconstruction, characterized by effective artifact reduction and noise suppression when using sparse projections. This suggests the potential application of our method in transmission X-ray microscopy (TXM) systems. In TXM, the rotation of a scanned sample may be constrained within a limited angular range to prevent collisions with other system components or to avoid high attenuation at specific tilting angles. However, it's important to note that image reconstruction from such limited angle data often leads to artifacts due to the absence of complete data.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2022R1A4A3018824, 2021R1F1A1059277, RS-2023-00250742).

REFERENCES

[1] André Aichert, Michael T Manhart, Bharath K Navalpakkam, Robert Grimm, Jana Hutter, Andreas Maier, Joachim Hornegger, and Arnd Doerfler. 2013. A realistic digital phantom for perfusion C-arm CT based on MRI data. In *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*. IEEE, 1–2.

[2] Anders H Andersen and Avinash C Kak. 1984. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm. *Ultrasonic imaging* 6, 1 (1984), 81–94.

[3] RC Atwood, JR Jones, PD Lee, and LL Hench. 2004. Analysis of pore interconnectivity in bioactive glass foams using X-ray microtomography. *Scripta Materialia* 51, 11 (2004), 1029–1033.

[4] DM Bappy and Insu Jeon. 2017. Modified simultaneous iterative reconstruction technique for fast, high-quality CT reconstruction. *IET Image Processing* 11, 9 (2017), 701–708.

[5] DM Bappy and Insu Jeon. 2019. High-quality X-ray computed tomography reconstruction using projected and interpolated images. *IET Image Processing* 13, 7 (2019), 1074–1080.

[6] Amir Beck and Marc Teboulle. 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences* 2, 1 (2009), 183–202.

[7] Judit Boda-Heggemann, Frank Lohr, Frederik Wenz, Michael Flentje, and Matthias Guckenberger. 2011. kV-Cone-beam-CT-basierte bildgeführte Strahlentherapie–ein klinischer Überblick: A Clinical Review. *Strahlentherapie und Onkologie* 187 (2011), 284–291.

[8] Emmanuel J Candès, Justin Romberg, and Terence Tao. 2006. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory* 52, 2 (2006), 489–509.

[9] W De Vos, Jan Casselman, and GRJ19464146 Swennen. 2009. Cone-beam computerized tomography (CBCT) imaging of the oral and maxillofacial region: a systematic review of the literature. *International journal of oral and maxillofacial surgery* 38, 6 (2009), 609–625.

[10] Yi Du, Xiangang Wang, Xincheng Xiang, and Zhouping Wei. 2016. Evaluation of hybrid SART+ OS+ TV iterative reconstruction algorithm for optical-CT gel dosimeter imaging. *Physics in Medicine & Biology* 61, 24 (2016), 8425.

[11] Lee A Feldkamp, Lloyd C Davis, and James W Kress. 1984. Practical cone-beam algorithm. *Josa a* 1, 6 (1984), 612–619.

[12] Yang Gao, Zhaoying Bian, Jing Huang, Yunwan Zhang, Shanzhou Niu, Qianjin Feng, Wufan Chen, Zhengrong Liang, and Jianhua Ma. 2014. Low-dose X-ray computed tomography image reconstruction with a combined low-mAs and sparse-view protocol. *Optics express* 22, 12 (2014), 15190–15210.

[13] Russell Garwood, Andrew Ross, Daniel Sotty, Dominique Chabard, Sylvain Charbonnier, Mark Sutton, and Philip J Withers. 2012. Tomographic reconstruction of neopterous Carboniferous insect nymphs. (2012).

[14] Richard Gordon, Robert Bender, and Gabor T Herman. 1970. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of theoretical Biology* 29, 3 (1970), 471–481.

[15] Jiang Hsieh. 2003. Computed tomography: principles, design, artifacts, and recent advances. (2003).

[16] Zhanli Hu and Hairong Zheng. 2014. Improved total variation minimization method for few-view computed tomography image reconstruction. *BioMedical Engineering OnLine* 13 (2014), 1–10.

[17] Marc Kachelriess, Oliver Watzke, and Willi A Kalender. 2001. Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT. *Medical physics* 28, 4 (2001), 475–490.

[18] Avinash C Kak and Malcolm Slaney. 2001. *Principles of computerized tomographic imaging*. SIAM.

[19] Byeongjoon Kim, Hyunjung Shim, and Jongduk Baek. 2022. A streak artifact reduction algorithm in sparse-view CT using a self-supervised neural representation. *Medical physics* 49, 12 (2022), 7497–7515.

[20] Seongjun Kim, Byeongjoon Kim, Jooho Lee, and Jongduk Baek. 2023. Sparsier2Sparse: Self-supervised convolutional neural network-based streak artifacts reduction in sparse-view CT images. *Medical Physics* (2023).

[21] K Machin and S Webb. 1994. Cone-beam X-ray microtomography of small specimens. *Physics in Medicine & Biology* 39, 10 (1994), 1639.

[22] Cynthia H McCollough, Adam C Bartley, Rickey E Carter, Baiyu Chen, Tammy A Drees, Phillip Edwards, David R Holmes III, Alice E Huang, Farhana Khan, Shuai Leng, et al. 2017. Low-dose CT for the detection and classification of metastatic liver lesions: results of the 2016 low dose CT grand challenge. *Medical physics* 44, 10 (2017), e339–e352.

[23] Jiayi Pan, Heye Zhang, Weifei Wu, Zhifan Gao, and Weiwen Wu. 2022. Multi-domain integrative swin transformer network for sparse-view tomographic reconstruction. *Patterns* 3, 6 (2022).

[24] Zhiwei Qiao. 2021. A simple and fast ASD-POCS algorithm for image reconstruction. *Journal of X-ray Science and Technology* 29, 3 (2021), 491–506.

[25] Robert L Siddon. 1985. Fast calculation of the exact radiological path for a three-dimensional CT array. *Medical physics* 12, 2 (1985), 252–255.

[26] Emil Y Sidky and Xiaochuan Pan. 2006. Accurate image reconstruction in circular cone-beam computed tomography by total variation minimization: a preliminary investigation. In *2006 IEEE Nuclear Science Symposium Conference Record*, Vol. 5. IEEE, 2904–2907.

[27] Emil Y Sidky and Xiaochuan Pan. 2008. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in Medicine & Biology* 53, 17 (2008), 4777.

[28] Shaojie Tang, Kuidong Huang, Yunyong Cheng, Tianye Niu, and Xiangyang Tang. 2017. Three-dimensional weighting in cone beam FBP reconstruction and its transformation over geometries. *IEEE Transactions on Biomedical Engineering* 65, 6 (2017), 1235–1244.

[29] Zhen Tian, Xun Jia, Kehong Yuan, Tinsu Pan, and Steve B Jiang. 2011. Low-dose CT reconstruction via edge-preserving total variation regularization. *Physics in Medicine & Biology* 56, 18 (2011), 5949.

[30] Ning Wen, Huaiqun Guan, Rabih Hammoud, Deepak Pradhan, T Nurushev, Shidong Li, and Benjamin Movsas. 2007. Dose delivered from Varian's CBCT to patients receiving IMRT for prostate cancer. *Physics in Medicine & Biology* 52, 8 (2007), 2267.

[31] Dong Wu and Li Zeng. 2014. Limited-angle reverse helical cone-beam CT for pipeline with low rank decomposition. *Optics Communications* 328 (2014), 109–115.

[32] Weiwen Wu, Xiaodong Guo, Yang Chen, Shaoyu Wang, and Jun Chen. 2022. Deep embedding-attention-refinement for sparse-view CT reconstruction. *IEEE Transactions on Instrumentation and Measurement* 72 (2022), 1–11.

[33] Weiwen Wu, Yanbo Zhang, Qian Wang, Fenglin Liu, Peijun Chen, and Hengyong Yu. 2018. Low-dose spectral CT reconstruction using image gradient L0–norm and tensor dictionary. *Applied Mathematical Modelling* 63 (2018), 538–557.

[34] Wei Yu and Li Zeng. 2016. Iterative image reconstruction for limited-angle inverse helical cone-beam computed tomography. *Scanning* 38, 1 (2016), 4–13.