An interprojection sensor fusion approach to estimate blocked projection signal in synchronized moving grid-based CBCT system

Hong Zhang

Department of Radiation Oncology, Georgia Regents University, Augusta, Georgia 30912

Lei Ren

Department of Radiation Oncology, Duke University, Durham, North Carolina 27710

Vic Kong

Department of Radiation Oncology, Georgia Regents University, Augusta, Georgia 30912

William Giles and You Zhang

Department of Radiation Oncology, Duke University, Durham, North Carolina 27710

Jian-Yue Jin^{a)}

Department of Radiation Oncology, Georgia Regents University, Augusta, Georgia 30912

and Department of Radiology, Georgia Regents University, Augusta, Georgia 30912

(Received 13 June 2015; revised 14 September 2015; accepted for publication 16 November 2015; published 28 December 2015)

Purpose: A preobject grid can reduce and correct scatter in cone beam computed tomography (CBCT). However, half of the signal in each projection is blocked by the grid. A synchronized moving grid (SMOG) has been proposed to acquire two complimentary projections at each gantry position and merge them into one complete projection. That approach, however, suffers from increased scanning time and the technical difficulty of accurately merging the two projections per gantry angle. Herein, the authors present a new SMOG approach which acquires a single projection per gantry angle, with complimentary grid patterns for any two adjacent projections, and use an interprojection sensor fusion (IPSF) technique to estimate the blocked signal in each projection. The method may have the additional benefit of reduced imaging dose due to the grid blocking half of the incident radiation.

Methods: The IPSF considers multiple paired observations from two adjacent gantry angles as approximations of the blocked signal and uses a weighted least square regression of these observations to finally determine the blocked signal. The method was first tested with a simulated SMOG on a head phantom. The signal to noise ratio (SNR), which represents the difference of the recovered CBCT image to the original image without the SMOG, was used to evaluate the ability of the IPSF in recovering the missing signal. The IPSF approach was then tested using a Catphan phantom on a prototype SMOG assembly installed in a bench top CBCT system.

Results: In the simulated SMOG experiment, the SNRs were increased from 15.1 and 12.7 dB to 35.6 and 28.9 dB comparing with a conventional interpolation method (inpainting method) for a projection and the reconstructed 3D image, respectively, suggesting that IPSF successfully recovered most of blocked signal. In the prototype SMOG experiment, the authors have successfully reconstructed a CBCT image using the IPSF-SMOG approach. The detailed geometric features in the Catphan phantom were mostly recovered according to visual evaluation. The scatter related artifacts, such as cupping artifacts, were almost completely removed.

Conclusions: The IPSF-SMOG is promising in reducing scatter artifacts and improving image quality while reducing radiation dose. © 2016 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4937934>]

Key words: scatter correction, dose reduction, CBCT, SMOG, moving grids, interpolation, geometric model, sensor fusion

1. INTRODUCTION

On-board cone beam computed tomography (CBCT) has been developed as an important imaging modality in image guided radiation therapy.¹ CBCT is usually performed before radiation delivery with a patient in a treatment position, often with the aid of various immobilization devices.² The CBCT image is then registered with the planning computed tomography (CT). The patient's treatment position is finally

adjusted according to the registration to ensure a precise radiation dose delivery.³ As a result, the image quality of CBCT is critical, especially for treatments requiring highly accurate dose delivery. The imaging dose is another important factor if CBCT scans are performed multiple times through the treatment course.

Unfortunately, CBCT image quality is largely degraded by scatter artifacts because a large volume of the imaging object is irradiated for each projection.⁴ Many approaches have been

proposed to reduce the scatter artifacts.⁵⁻⁸ One approach is to utilize a preobject grid to divide the large cone beam field into multiple small narrow strips, and thus physically reduce the amount of scatter. In addition, the remaining scatter can be measured behind the blocked regions of the grid and be corrected after the scan. However, the imaging signal in these shadow regions is also blocked by the grid. We have previously proposed two approaches to resolve this issue:^{5,7} (1) deriving the blocked signal directly with interpolation in each projection⁵ and (2) using a synchronized moving grid (SMOG) to acquire two complimentary projections at each gantry angle with the grid off-set by certain distance, and merging the two projections to obtain a complete signal.^{5,7} The SMOG approach may also potentially reduce the image lag induced artifacts.^{5,7}

The interpolation approach has the advantage of reducing imaging dose by half. Many inpainting techniques, including texture synthesis, partial differential equation, and exemplar based techniques as well as their hybrids⁹⁻¹² can be used for interpolation. However, these approaches usually cannot recover sufficient information in the blocked regions, especially when wider grids are used.⁵ On the other hand, the SMOG approach can recover full information in principle. However, it is technically challenging. It requires a fast moving grid, especially for fast gantry rotation. The merge of the two complimentary projections requires very accurate system design, and special care has to be given to the penumbral regions to avoid introducing new artifacts. More importantly, the CBCT control system needs to be redesigned completely to enable acquiring at least two projections at the same gantry angle. This may also increase the scan time due to the limitation of the frame rate of the flat panel detector. In addition, the approach removes any dose reduction benefits.

In this study, we propose an interprojection sensor fusion (IPSF) method, which utilizes the information of two adjacent projections as references to estimate the blocked image information. The new method still uses a SMOG approach. However, instead of acquiring two projections at each gantry angle, it acquires only one projection and moves the grid to the complementary position for the next gantry angle. Thus, any two adjacent projections have complimentary grid patterns. That is, the blocked regions in each projection correspond to the open regions in the two adjacent projections. Unlike the inpainting approach, which estimates the blocked signal in the same projection, the IPSF method estimates blocked signals alternatively from multiple paired pixels of the two adjacent projections with a weighted least square regression (WLSR).

2. MATERIALS AND METHODS

In this section, we first describe the basic concept of IPSF. We then present the methodology to verify our approach, including testing the IPSF with a simulated SMOG on projections of a head phantom, and a prototype SMOG system on a Catphan phantom.

2.A. IPSF concept

The IPSF concept is based on the sensor fusion theory, which combines the information for the same object detected from different sensors to improve accuracy.¹³ We borrow the idea and assume that the projection observations from three consecutive gantry angles to different sampling voxels can function as sensors to detect the blocked signal, which is represented by an intensity (or pixel) value for the projection ray at the central gantry angle. Although the intensity is the result of line integration along the projection ray rather than over a single sampling voxel, the assumption is valid under certain conditions. As shown in Fig. 1, we consider m sampling voxels equal-distantly along the ray of the central gantry angle, at which the true intensity is to be estimated. Each of the three consecutive gantry angles (S_{g+1} , S_g , S_{g-1}) provides m observations through the m sampling voxels, with the observation values being the intensities corresponding to the projection rays from the source to the sampling voxels, respectively. The similarity between these observations decreases with the increasing gantry-angle step. Under a strict condition that the gantry-angle step is sufficiently small (such as 0.2°), the rectangle in Fig. 1 becomes sufficiently narrow (~ 1 mm) so that the lateral variation of density is negligible within the entire rectangle. Consequently, the $3m$ line integrals corresponding to the $3m$ observations are very similar to each other, and the $3m$ observations can almost accurately detect the true intensity to be recovered. If the gantry-angle step increases to a not-too-large value (such as 1°), the lateral variation is only likely negligible due to lateral continuity. Under this condition, the $3m$ observations may statistically reflect the true intensity similar to a sensor fusion approach. In addition, comparisons of m paired observations from two adjacent gantry angles provide the likelihood of lateral variation for each observation pair, so that a weighted statistical model can be used to estimate the true intensity from the total of $3m$ observations.¹³

It should be noted that we consider $3m$ observations from the three consecutive gantry angles (S_{g+1} , S_g , S_{g-1}) here,

FIG. 1. Illustration of m sampling voxels in an object, and $3m$ observations through the sampling voxels from three consecutive gantry angles. Note that the m observations from the central gantry angle are the same.

because a pixel to be recovered at the central gantry angle (S_g) could be in the penumbral region, and the observation from this gantry angle also contributes to the final estimation. A pixel could be in a nonblocked, a blocked, or a penumbral region. For consistency, we treat the three regions with a unified algorithm by weighting the contribution of S_g differently according to the pixel location. The weight is 1 for the nonblocked region, is 0 for the completely blocked region, and varies from 0 to 1 according to the location in the penumbral region (or the blockage ratio). It should also be noted that the m observations at S_g have the exact same value.

Therefore, a main part of the IPSF algorithm is to estimate the blocked signal from the $2m$ observations from the two adjacent gantry angles (S_{g+1}, S_{g-1}). Figure 2 shows a flow chart that illustrates the principle and process of this part of algorithm. It includes the following steps: (1) For a pixel to be recovered in a projection, the locations of projecting points corresponding to the m observations for each of the adjacent projection pair are identified using a CBCT geometric model;^{14,15} (2) the intensity values of the m observations are sampled for the pair of projections; (3) the weight of each observation pair is calculated according to the intensity difference of the pair; and (4) the voxel intensity is estimated by a weighted least square regression model for the $2m$ observations and the m observations from the central gantry angle.

A key for this IPSF algorithm is Step 3: to determine the weight of each observation. In this study, we used the difference of paired observations to determine the weight. If a pair of observations from the two adjacent gantry angles

to a same sampling voxel has similar observation values, we consider them having a higher probability to reflect the true intensity value. Therefore, a high weight is given to the pair of observation. For simplicity, the detailed derivations and formalisms for the geometric model, the weight of each paired observations, the weight of the contribution of the central gantry angle in the penumbral region, and the weighted regression model are given in Appendices A and B.

2.B. Verification of the IPSF

To verify the recovery ability of IPSF, we tested it first on a simulated SMOG applied to CBCT projections. A head phantom was scanned using the onboard CBCT system on a Varian® Trilogy™ Linac with the x-ray tube voltage of 120 kVp. In each scan, approximately 650 projections were acquired and each projection dimensions are 1024×768 with resolutions of 0.388×0.388 mm. The reconstruction image contains $384 \times 384 \times 64$ voxels with resolutions of $0.651 \times 0.651 \times 2.5$ mm. The SMOG was simulated by erasing the information (filling with “0”) of the regions in each projection corresponding to the grid which is designed as 1:1 space ratio and 21 pixels width. The artificial grid moved back and forth in each projection at the step of the grid width without considering a penumbra. In this way, a blocked projection pixel in one gantry angle was guaranteed to be “seen” in the two adjacent gantry angles. The blocked image information was derived from the information in the two adjacent projections immediately before and after the projection using the IPSF algorithm. Scatter information was not able to be measured and removed using the simulated

FIG. 2. Illustration of the IPSF principle and process. (1) Calculate the locations of m observation points for the pair of adjacent projections using a geometric model; (2) sample the m observation values; (3) calculate the weight of each observation pair according to the intensity difference of the pair; and (4) estimate the pixel intensity by WLSR.

FIG. 3. Comparison of recovery abilities between inpainting and IPSF. (a) Original projection, (b) projection blocked by the simulated SMOG, (c) projection recovered by inpainting (SNR = 15.1 dB), and (d) projection recovered by IPSF (SNR = 35.6 dB).

SMOG. A complete projection was derived, and the CBCT image was reconstructed using the Feldkamp-Davis-Kress (FDK) reconstruction algorithm in the vendor system.¹⁶ The original nonblocked CBCT image was used as the gold standard and the difference between the recovered image to the original image was calculated to quantify the blocked

signal recovery accuracy. This difference is usually defined as signal-to-noise ratio (SNR) in image processing and the signal and noise here represent specifically the original image and the difference between the recovered image to the original image, respectively.¹⁷ We adapt the definition and thus the SNR is expressed as

$$\text{SNR} = 10 \log_{10} \left\{ \frac{\left[\sum_{i=1}^M \sum_{j=1}^N \text{Im}_O(i,j)^2 \right]}{\left[\sum_{i=1}^M \sum_{j=1}^N (\text{Im}_O(i,j) - \text{Im}_R(i,j))^2 \right]} \right\} \text{ dB}, \quad (1)$$

where $\text{Im}_O(i,j)$ and $\text{Im}_R(i,j)$ represent a 2D cut of the original and recovered CBCT image, respectively; M and N represent the size of the 2D image in terms of number of pixels.¹⁷ SNR is very easy to be extended and defined in 3D after simply adding a dimension. The IPSF method was then compared with an exemplar based inpainting interpolation method¹⁸ in terms of the SNR.

We then tested the framework on a prototype SMOG system to demonstrate the feasibility. The system includes the following components: (a) x-ray tube [Varian G-242, Varian Medical Systems (VMS), CA], (b) flat panel detector (4030CB, VMS, CA), (c) x-ray generator (EPS 50RF, EMD, Canada), and (d) in-house SMOG (assembled with QCI-S2-IG, QCI, CA, and LAS16-25, BEI, CA). Controlling system is written in LabVIEW to synchronize the grid position in each gantry angle to make it move at each acquirement. The study was performed on a Catphan phantom with the gantry angle step of 1° . The scatter information was measured in

the blocked area, estimated using interpolation and removed before IPSF processing for each projection. An in-house FDK algorithm was used to reconstruct the CBCT image.

3. RESULTS

3.A. IPSF on the simulated SMOG

Figure 3 illustrates the recovery ability of IPSF comparing with inpainting. Figure 3(a) shows an original projection of a head phantom; Fig. 3(b) shows the corresponding projection blocked by the simulated SMOG; Fig. 3(c) shows the projection recovered by the inpainting, with a SNR of 15.1 dB to the original projection; Fig. 3(d) shows the projection recovered by the IPSF, with a SNR of 35.6 dB to the original projection.

Figure 4 shows the comparison of the intensity profiles among inpainting, IPSF, and actual image for two lines in

FIG. 4. Comparison of intensity profiles among inpainting, IPSF, and actual image. (a) Profiles for a line in the orbits region in Fig. 3(b); (b) profiles for a line in the teeth region in Fig. 3(b).

Fig. 5. Comparison of the head phantom reconstruction image between inpainting and IPSF: (a)–(c) axial (with zoom-in view), coronal, and sagittal cut of the original CBCT image; (d)–(f) axial (with zoom-in view) (SNR = 14.1 dB), coronal (SNR = 17.2 dB), and sagittal (SNR = 13.8 dB) cuts for inpainting; (g)–(i) axial (with zoom-in view) (SNR = 27.7 dB), coronal (SNR = 26.9 dB), and sagittal (SNR = 23.3 dB) cuts for IPSF.

Fig. 3(b). Figure 4(a) corresponds to the line at the eye-ball region, while Fig. 4(b) corresponds to the line at the teeth region. The IPSF profiles consist with the original profiles much better than the inpainting profiles. However, the high frequency components in the original profiles are apparently filtered in the IPSF profiles.

Figure 5 shows the comparison of the reconstruction CBCT image between inpainting and IPSF. Figures 5(a)–5(c) show axial (with zoom-in view), coronal, and sagittal cuts, respectively, for the original CBCT image; Figs. 5(d)–5(f) show the same axial (with zoom-in view), coronal, and sagittal cuts, respectively, for the CBCT image by inpainting; and Figs. 5(g)–5(i) show the same axial (with zoom-in view), coronal, and sagittal cuts, respectively, for the CBCT image

by IPSF. The SNR improvements from inpainting to IPSF are summarized in Table I. The SNR for the 2D cut in axial, coronal, and sagittal directions for the CBCT image using IPSF was 27.7, 26.9, and 23.3 dB, respectively, in comparison to 14.1, 17.2, and 13.8 dB correspondingly for the image

TABLE I. Summary of the SNR improvement from inpainting to IPSF.

Methods	Images			
	Axial	Coronal	Sagittal	3D
Inpainting	14.1 dB	17.2 dB	13.8 dB	12.7 dB
IPSF	27.7 dB	26.9 dB	23.3 dB	28.9 dB

FIG. 6. An example of recovered projections for the Catphan phantom on the prototype SMOG: (a) a conventional CBCT projection, (b) the corresponding blocked projection, (c) the projection recovered by IPSF.

using the inpainting method. The SNR for the 3D CBCT image using IPSF was 28.9 dB comparing with 12.7 dB using the inpainting.

3.B. IPSF on the prototype SMOG

This experiment demonstrated the scatter reduction ability of IPSF-SMOG comparing with conventional CBCT. Figure 6 shows an example of recovered projections for the Catphan phantom on the prototype SMOG. Figure 6(a) shows a conventional CBCT Catphan phantom projection; Fig. 6(b) shows the corresponding projection blocked by the prototype SMOG; Fig. 6(c) shows the projection recovered by the IPSF.

Figure 7 shows the comparison of the Catphan phantom reconstruction image between conventional CBCT and IPSF-SMOG on the prototype SMOG. Figures 7(a) and 7(b) show

two axial cuts for the conventional CBCT reconstruction image; Figs. 7(c) and 7(d) show the same two axial cuts for IPSF-SMOG image reconstructed. The cupping artifacts due to scatter are apparent in Fig. 7(a), and it is almost invisible in Fig. 7(c).

The cupping artifacts are quantitatively evaluated by plotting the intensity profile over a line through the region of artifacts. Figure 8 shows the comparison of the cupping artifacts between the conventional CBCT and IPSF-SMOG by comparing intensity profiles. Figure 8(a) shows the intensity profile over the line in Fig. 7(a). The profile is well fitted with a simple two-order polynomial function ($y = 0.0042x^2 - 0.43x + 160$). Figure 8(b) shows the intensity profile over the line in Fig. 7(c). It is fitted by a much flatter two-order polynomial function ($y = 0.0005x^2 - 0.038x + 180$).

FIG. 7. The comparison of the Catphan phantom reconstruction image between conventional CBCT and IPSF-SMOG on the prototype SMOG: (a) and (b) two axial cuts for conventional CBCT, (c) and (d) corresponding axial cuts for IPSF-SMOG.

FIG. 8. Comparison of cupping artifacts between the conventional CBCT and IPSF-SMOG by plotting the intensity profiles for a line in Figs. 7(a) and 7(c). (a) The intensity profile over the line in Fig. 7(a) and (b) the intensity profile over line in Fig. 7(c).

The IPSF-SMOG approach may degrade the spatial resolution, which can be evaluated by the modulation transfer function (MTF). Figure 9 shows the comparison of MTF between the conventional CBCT and IPSF-SMOG approach. The MTFs for the two CBCT images were derived from the images of the line-pair pattern in Figs. 7(b) and 7(d). Indeed, the spatial resolution was slightly degraded for the IPSF-SMOG approach.

4. DISCUSSION

Based on Fig. 5, we found that IPSF-based FDK significantly improved the reconstruction image quality comparing with inpainting-based FDK. In the simulated experiment, the SNR was increased from 15.1 to 35.6 dB for a head phantom projection. In the reconstruction image, the SNR was increased from 14.1, 17.2, and 13.8 dB to 27.7, 26.9, and 23.3 dB in axial, coronal, and sagittal directions, respectively. In 3D image, the SNR was increased from 12.7 to 28.9 dB. The SNR of IPSF is about twice as much as the SNR of inpainting and it means that IPSF improves the recovery accuracy by a factor of 10 comparing with inpainting. The improvement is significant and matches the observation on zoom-in view of Figs. 5(a), 5(d), and 5(g).

FIG. 9. Comparison of MTF between conventional CBCT and the IPSF-SMOG approach.

These results demonstrate that the IPSF technique can provide superior results to the inpainting based interpolation. The recovery accuracy highly depends on the gantry-angle steps. The recovery will be accurate when the gantry-angle step is very small. The technique also differs from interpolation because it provides multiple observations to approximate the missing data.

In the prototype SMOG, the cupping artifacts caused by scatters were clearly removed as shown in Fig. 7. We fitted the intensity profiles over a sampling line in Fig. 7(a) and the same sampling line in Fig. 7(c) with a two-order polynomial curve. The mean curvature of all points on the fitted curve was calculated to measure the bending of the curve.¹⁹ Curvature is an efficient tool to measure bending of a curve and the smaller the curvature is, the straight the curve is. The mean curvatures of the fitted curve of Figs. 8(a) and 8(b) are 0.0078 and 0.0010, respectively, and it implies that the cupping artifacts have been reduced significantly with IPSF-SMOG.

Though we recovered most of blocked signal, some degradation in the spatial resolution is unavoidable since we only utilize half of the whole signal comparing with conventional CBCT to reconstruct image. The reduction of spatial resolution is mainly the loss of high frequency information, as demonstrated in Figs. 4(a) and 4(b). It shows the blurry of the edge, but not loss of the edge information, or deformation of the structures, as demonstrated in Figs. 5(g)–5(i) and Figs. 7(c) and 7(d). Similar behavior is also reflected in the degradation of MTF as shown in Fig. 9. Some very fine structures in the line pair patterns in Figs. 7(d) and 5(g) are still visible, suggesting that the clinical impact of the reduction of spatial resolution is small. This loss may also associate with the visible horizontal stripes and lines in the recovered projection as shown in Fig. 4(c). We can see that the grid divides each projection into two regions: blocked and nonblocked regions (in addition, a penumbral region for actual SMOG). In the nonblocked regions, the recovered signals are mainly from the original projection directly. Thus, high frequent signals are kept. However, in the blocked regions, the recovered signals are mainly estimated from the two adjacent projections. The estimation uses a regression approach, which inherits an averaging effect and thus loses some high frequent signal and serves as a low-pass filter. Therefore, the recovered projection demonstrates two

types of regions interleaving with each other and forms the visible horizontal stripes/lines. The horizontal stripes/lines are not as obvious in Fig. 3(d) because (1) there are no penumbra regions for the artificial grid and (2) the head phantom has more intensity variation so that the variation of the regions is not as outstanding.

The SMOG system may increase noise level due to scatter subtraction^{5,7} and this phenomenon could be observed in Fig. 7(c). We utilized the fitted curves, $f_1(x)$ and $f_2(x)$, in Figs. 6(a) and 6(b), respectively, to estimate the noise level excluding cupping artifacts. The noise level is quantified as the variance between the intensity profile and the fitted curve. Thus, the noise levels for the conventional CBCT and IPSF-SMOG were calculated as $\text{mean}\{[f_1(x) - g_1(x)]^2\}$ and $\text{mean}\{[f_2(x) - g_2(x)]^2\}$, respectively, where $g_1(x)$ and $g_2(x)$ are intensity profiles, respectively. The results were 532.3 and 755.3, respectively. The noise level was increased after IPSF-SMOG. In the future work, we plan to apply total variation based CBCT reconstruction method in combination with IPSF to reduce the noise level and improve image quality.^{20,21}

5. CONCLUSIONS

The IPSF technique applied a novel method utilizing the adjacent projections to estimate the blocked projection following a geometric model. It showed much better reconstruction results than the inpainting algorithm to recover the blocked information in simulated SMOG. In combination with a prototype SMOG, it successfully reduced the scatter artifacts and may reduce the image dose. It may also be used for sparse-view CBCT reconstruction to further reduce imaging dose.

ACKNOWLEDGMENTS

This study was supported by Award No. 1R01CA166948 from the NIH and National Cancer Institute.

APPENDIX A: GEOMETRIC MODEL

The geometric model^{14,15} is shown in Fig. 10 and the nomenclature is listed after it. In Fig. 10, the whole system is placed in a standard 3D global Eulerian coordinate system with origin of O and X , Y , Z axis and viewed from top to bottom (Z axis). S_{g-1} , S_g , and S_{g+1} are positions of the source in the three consecutive gantry angles. E_j is the j th sampling voxel inside the scanned object along a ray casting from S_g . Three rays are casted from S_{g-1} , S_g , and S_{g+1} to E_j , respectively, and project on the detector at $PS_{g-1}(E_j)$, $PS_g(E_j)$, and $PS_{g+1}(E_j)$, respectively. On the detector, a local coordinate system with origin $PS_{g-1}(O)$ [or $PS_g(O)$ or $PS_{g+1}(O)$ depending on the gantry angle], i.e., projection of O on the detector, V axis (parallel to Z axis), and U axis (perpendicular to the plane that rotation orbit lies on) is built. The known inputs of the model are S_{g-1} , S_g , S_{g+1} , E_j , $PS_g(E_j)$, O , SAD, and SDD and the outputs are the local coordinates of $PS_{g-1}(E_j)$ and $PS_{g+1}(E_j)$.

The calculations of the two outputs from the geometric model are elaborated here. Since rotation orbit is perpendicular to Z axis, $PS_g(E_j)$, $PS_{g-1}(E_j)$, and $PS_{g+1}(E_j)$ have the same known Z and local V coordinate v_g and we only need to calculate the local U coordinate. Taking $PS_{g-1}(E_j)$ as an example, the U coordinate can be derived from the angle, θ , between vectors $\overrightarrow{S_{g-1}O}$ and $\overrightarrow{S_{g-1}E_j}$.

The angle θ is calculated as

FIG. 10. The geometric model. SMOG: in-house SMOG, SDD: source-to-detector distance, SAD: source-to-axis distance, S_{g-1} , S_g , S_{g+1} : the locations of source in consecutive gantry angle, O : origin of global coordinate, E_j : the j th sampling voxel on the ray between S_g and $PS_g(E_j)$, $PS_{g-1}(E_j)$, $PS_g(E_j)$, $PS_{g+1}(E_j)$: the projection pixel of E_j under source S_{g-1} , S_g , and S_{g+1} on the detector, respectively, $PS_{g-1}(O)$: the local coordinate origin under source S_{g-1} on the detector, U , $V(Z)$: the local coordinate of the detector projection plane and the global axis Z is parallel to the axis V , and X , Y , Z : the global coordinate.

$$\theta = \cos^{-1} \left(\frac{\overrightarrow{S_{g-1}E_j} \cdot \overrightarrow{S_{g-1}O}}{\|\overrightarrow{S_{g-1}E_j}\| \cdot \|\overrightarrow{S_{g-1}O}\|} \right) \quad (A1)$$

and the distance between $PS_{g-1}(E_j)$ and $PS_{g-1}(O)$ is derived as

$$\|\overrightarrow{PS_{g-1}(E_j)PS_{g-1}(O)}\| = SDD \cdot \tan(\theta). \quad (A2)$$

Finally, the absolute value of local U coordinate is

$$|PS_{g-1}(E_j)_U| = \sqrt{(SDD \cdot \tan(\theta))^2 - v_g^2}, \quad (A3)$$

where $PS_{g-1}(E_j)_U$ is the local U coordinate of $PS_{g-1}(E_j)$. The sign of the U coordinate is the same as the Z coordinate of $\overrightarrow{S_{g-1}E_j} \times \overrightarrow{S_{g-1}O}$. In the same way, the local U coordinate of $PS_{g+1}(E_j)$ can be calculated.

$$\begin{aligned} I[PS_{\text{view}}(E_j)] &= b[PS_{\text{view}}(E_j)] \cdot (\bar{I}[PS_{\text{view}}(E_j)] + \varepsilon) + \text{scatter} \\ &= b[PS_{\text{view}}(E_j)] \cdot (\bar{I}[PS_g(E_j)] + \bar{I}[PS_{\text{view}}(E_j)] - \bar{I}[PS_g(E_j)] + \varepsilon) + \text{scatter} \\ &= b[PS_{\text{view}}(E_j)] \cdot (\bar{I}[PS_g(E_j)] + \text{Diff}\{\bar{I}[PS_g(E_j)], \bar{I}[PS_{\text{view}}(E_j)]\} + \varepsilon) + \text{scatter}, \end{aligned} \quad (B1)$$

where $\text{view} \in \{g-1, g, g+1\}$ and $j = 1 \dots m$.

In this model, $b[PS_{\text{view}}(E_j)]$ is the blockage ratio between 0 and 1 at the pixel $PS_{\text{view}}(E_j)$ caused by the grid blockage. Lower values correspond to more blocking with 0 indicated 100% blockage. We obtain $b[PS_{\text{view}}(E_j)]$ in this way: 20 sets of air projections $\text{Proj}_{\text{view}}^{\text{air}}$ and 20 sets of air projections with the grid mounted $\text{Proj}_{\text{view}}^{\text{grid}}$ are scanned and $b[PS_{\text{view}}(E_j)]$ is calculated by $\text{mean}(\text{Proj}_{\text{view}}^{\text{grid}}) / \text{mean}(\text{Proj}_{\text{view}}^{\text{air}})$.

ε is a Gaussian variable with mean of 0 and variance of $\bar{I}[PS_{\text{view}}(E_j)]$. *Scatter* is the scatter intensity at $PS_{\text{view}}(E_j)$.

$\text{Diff}\{\bar{I}[PS_g(E_j)], \bar{I}[PS_{\text{view}}(E_j)]\}$ is a difference term representing the change of value between $\bar{I}[PS_g(E_j)]$ and $\bar{I}[PS_{\text{view}}(E_j)]$ caused by the gantry rotation.

$$I^{\text{no scatter}}[PS_{\text{view}}(E_j)] = b[PS_{\text{view}}(E_j)] \cdot (\bar{I}[PS_g(E_j)] + \text{Diff}\{\bar{I}[PS_g(E_j)], \bar{I}[PS_{\text{view}}(E_j)]\} + \varepsilon), \quad (B2)$$

where $I^{\text{no scatter}}[PS_{\text{view}}(E_j)]$ is the observed intensity after scatter reduction.

The second unknown term is the difference term and it is the main source of bias for this model since it is not measurable and uncertain. We simplify Eq. (B2) by dropping the difference term with the assumption of lateral structure continuity due to a small gantry angle step. Thus, we have

$$I^{\text{no scatter}}[PS_{\text{view}}(E_j)] = b[PS_{\text{view}}(E_j)] \cdot (\bar{I}[PS_g(E_j)] + \varepsilon). \quad (B3)$$

In this way, a paired of projection pixels, $PS_{g-1}(E_j)$ and $PS_{g+1}(E_j)$, can be localized analytically for any sampling voxel E_j under the geometric restriction.

APPENDIX B: SENSOR FUSION ALGORITHM

We fuse the projection observations obtained above at pixels of $PS_{g-1}(E_j)$, $PS_g(E_j)$, and $PS_{g+1}(E_j)$ from three consecutive gantry angles (S_{g-1}, S_g, S_{g+1}) intersecting at the same sampling voxel E_j to estimate the true intensity. Based on the scan field size, m sampling voxels (i.e., $j = 1 \dots m$) are placed at an equal interval along the ray connecting the source and the pixel $PS_g(E_j)$. We build a model to describe the relationship between the true intensity $\bar{I}[PS_{g-1}(E_j)]$, $\bar{I}[PS_g(E_j)]$ and $\bar{I}[PS_{g+1}(E_j)]$ and the observed intensity $I[PS_{g-1}(E_j)]$, $I[PS_g(E_j)]$, and $I[PS_{g+1}(E_j)]$ as

Equation (B1) is a complete model considering all factors to describe the true intensity on each pixel. In the model, the observed intensity or the number of photons mainly is composed by two parts: one is a Poisson variable decayed by scanned object and SMOG from the x-ray source directly and the other is the scatter signal. However, it is unsolvable due to several unknown terms in this equation. So some degeneration is required to solve it.

The first unknown term is the scatter information and it is the main source of artifacts in CBCT. Fortunately, it is measurable and estimable in SMOG.^{5,7} It could be estimated simply by interoperating and propagating the blockage intensity to the unblocked regions to obtain the whole scatter in projections. After scatter reduction, the model becomes

Equation (B3) could be easily solved by linear regression. However, it could bring in severe biased errors after dropping the difference term. We use WLSR, especially a weight $\text{Weight}_{S_{\text{view}}}(E_j)$ representing the unbiased level of Eq. (B3) to compensate the dropped difference term in this way:

1. When view is g , $\text{Diff}\{\bar{I}[PS_g(E_j)], \bar{I}[PS_{\text{view}}(E_j)]\}$ is zero and $\text{Weight}_{S_g}(E_j)$ is $b[PS_{\text{view}}(E_j)]$ since the blockage dominates the bias.

2. When view is $g-1$ or $g+1$, the weight could be estimated based on the lateral continuity: it could be derived based on Eq. (B2) that

$$\begin{aligned} I^{\text{no scatter}}[PS_{g-1}(E_j)] \\ = b[PS_{g-1}(E_j)] \cdot \{\bar{I}[PS_g(E_j)] \\ + \text{Diff}\{\bar{I}[PS_g(E_j)], \bar{I}[PS_{g-1}(E_j)]\} + \varepsilon\} \end{aligned} \quad (\text{B4})$$

$$\begin{aligned} I^{\text{no scatter}}[PS_{g+1}(E_j)] \\ = b[PS_{g+1}(E_j)] \cdot \{\bar{I}[PS_g(E_j)] \\ + \text{Diff}\{\bar{I}[PS_g(E_j)], \bar{I}[PS_{g+1}(E_j)]\} + \varepsilon\} \end{aligned} \quad (\text{B5})$$

After eliminating $\bar{I}[PS_g(E_j)]$ from Eqs. (B4) and (B5), this can be derived that

$$\begin{aligned} \text{abs}\left(\text{Diff}\{\bar{I}[PS_g(E_j)], \bar{I}[PS_{g-1}(E_j)]\} - \text{Diff}\{\bar{I}[PS_g(E_j)], \bar{I}[PS_{g+1}(E_j)]\}\right) \\ = \text{abs}(I^{\text{no scatter}}[PS_{g-1}(E_j)]/b[PS_{g-1}(E_j)] - I^{\text{no scatter}}[PS_{g+1}(E_j)]/b[PS_{g+1}(E_j)] + \varepsilon'). \end{aligned} \quad (\text{B6})$$

Here, ε' is a Gaussian variable with mean of 0 and variance of $\bar{I}[PS_{g-1}(E_j)] + \bar{I}[PS_{g+1}(E_j)]$ since $I^{\text{no scatter}}[PS_{g-1}(E_j)]$ and $I^{\text{no scatter}}[PS_{g+1}(E_j)]$ are observed independently. From Eq. (B6), we can observe that if $\text{abs}(\text{Diff}\{\bar{I}[PS_g(E_j)], \bar{I}[PS_{g-1}(E_j)]\} - \text{Diff}\{\bar{I}[PS_g(E_j)], \bar{I}[PS_{g+1}(E_j)]\})$, i.e., left term, is too large, this implies that $I^{\text{no scatter}}[PS_{g-1}(E_j)]$ and $I^{\text{no scatter}}[PS_{g+1}(E_j)]$ are both biased or at least one is biased. If it is small, this implies that they are unbiased with high probability since they present high continuity from two source positions $g-1$ and $g+1$. As a result, we define the weights as below,

$$\begin{aligned} \text{Weight}_{S_{g-1}}(E_j) = 2 \cdot b[PS_{g-1}(E_j)] \\ \cdot \{1 - \text{CDF}[\text{abs}(I^{\text{no scatter}}[PS_{g-1}(E_j)]/b[PS_{g-1}(E_j)] - I^{\text{no scatter}}[PS_{g+1}(E_j)]/b[PS_{g+1}(E_j)])]\}, \end{aligned} \quad (\text{B7})$$

$$\begin{aligned} \text{Weight}_{S_{g+1}}(E_j) = 2 \cdot b[PS_{g+1}(E_j)] \\ \cdot \{1 - \text{CDF}[\text{abs}(I^{\text{no scatter}}[PS_{g+1}(E_j)]/b[PS_{g+1}(E_j)] - I^{\text{no scatter}}[PS_{g-1}(E_j)]/b[PS_{g-1}(E_j)])]\}. \end{aligned} \quad (\text{B8})$$

Here, CDF means cumulative distribution function of a Gaussian distribution with mean of 0 and variance of $I^{\text{no scatter}}[PS_{g-1}(E_j)]/b[PS_{g-1}(E_j)] + I^{\text{no scatter}}[PS_{g+1}(E_j)]/b[PS_{g+1}(E_j)]$. Here, we use the estimated variance not unknown true variance without inducing large bias. From the definition, we form the weight with two parts: one is the blockage and the other is the continuity. The first part has been explained above and the second part is calculated from CDF of the right term of Eq. (B6). When the right term of Eq. (B6) is close to 0, CDF is close to 0.5 and the second part is close to 1. In this case, high continuity is observed. On the contrary, the second part is close to 0 when the right term of Eq. (B6) is much larger than 0.

Currently, three equations could be built up following Eq. (B3) for E_j and each equation is assigned a weight following the above steps. As we mentioned above, we place m sampling voxels along the ray and totally $3m$ equations with $3m$ weights form $3m$ observations of $\bar{I}[PS_g(E_j)]$. The $3m$ observations form an equation set and $\bar{I}[PS_g(E_j)]$ is the only unknown variable. Finally, with WLSR the estimation is¹³

$$\widehat{\bar{I}[PS_g(E_j)]} = (B^T W B)^{-1} B^T W S, \quad (\text{B9})$$

where

$$\begin{aligned} B &= [b[PS_{g-1}(E_1)] \dots b[PS_{g-1}(E_m)] b[PS_g(E_1)] \dots b[PS_g(E_m)] b[PS_{g+1}(E_1)] \dots b[PS_{g+1}(E_m)]]^T, \\ W &= [\text{Weight}_{S_{g-1}}(E_1) \dots \text{Weight}_{S_{g-1}}(E_m) \text{Weight}_{S_g}(E_1) \dots \text{Weight}_{S_g}(E_m) \text{Weight}_{S_{g+1}}(E_1) \dots \text{Weight}_{S_{g+1}}(E_m)]^T, \end{aligned}$$

and

$$\begin{aligned} S &= [I^{\text{no scatter}}[PS_{g-1}(E_1)] \dots I^{\text{no scatter}}[PS_{g-1}(E_m)] I^{\text{no scatter}}[PS_g(E_1)] \dots I^{\text{no scatter}}[PS_g(E_m)] \\ &\quad \times I^{\text{no scatter}}[PS_{g+1}(E_1)] \dots I^{\text{no scatter}}[PS_{g+1}(E_m)]]^T. \end{aligned}$$

^{a1}Author to whom correspondence should be addressed. Electronic mail: jjin@gru.edu; Telephone: 706-721-2971.

¹L. Xing, B. Thorndyke, E. Schreibmann, Y. Yang, T.-F. Li, G.-Y. Kim, G. Luxton, and A. Koong, "Overview of image-guided radiation therapy," *Med. Dosim.* **31**(2), 91–112 (2006).

²B. Schaly, J. Kempe, G. Bauman, J. Battista, and J. Van dyk, "Tracking the dose distribution in radiation therapy by accounting for variable anatomy," *Phys. Med. Biol.* **49**(5), 791–805 (2004).

³L. A. Dawson and M. B. Sharpe, "Image-guided radiotherapy: Rationale, benefits, and limitations," *Lancet Oncol.* **7**(10), 848–858 (2006).

⁴L. Zhu, Y. Xie, J. Wang, and L. Xing, "Scatter correction for cone-beam CT in radiation therapy," *Med. Phys.* **36**(6), 2258–2268 (2009).

⁵J.-Y. Jin, L. Ren, Q. Liu, J. Kim, N. Wen, H. Guan, B. Movsas, and I. J. Chetty, "Combining scatter reduction and correction to improve image quality in cone-beam computed tomography (CBCT)," *Med. Phys.* **37**(11), 5634–5644 (2010).

⁶R. Ning, X. Tang, and D. Conover, "X-ray scatter correction algorithm for cone beam CT imaging," *Med. Phys.* **31**(5), 1195–1202 (2004).

⁷L. Ren, F.-F. Yin, I. J. Chetty, D. A. Jaffray, and J.-Y. Jin, "Feasibility study of a synchronized-moving-grid (SMOG) system to improve image quality in

- cone-beam computed tomography (CBCT)," *Med. Phys.* **39**(8), 5099–5110 (2012).
- ⁸J. Siewerdsen, M. Daly, B. Bakhtiar, D. Moseley, S. Richard, H. Keller, and D. Jaffray, "A simple, direct method for x-ray scatter estimation and correction in digital radiography and cone-beam CT," *Med. Phys.* **33**(1), 187–197 (2006).
- ⁹D. J. Heeger and J. R. Bergen, "Pyramid-based texture analysis/synthesis," in *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques* (ACM, New York, NY, 1995), pp. 229–238.
- ¹⁰J. Yang, K. Hua, Y. Wang, W. Wang, H. Wang, and J. Shen, "Automatic objects removal for scene completion," in *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (IEEE, Piscataway, NJ, 2014), pp. 553–558.
- ¹¹J. Wu and Q. Ruan, "A novel hybrid image inpainting model," in *ICALIP 2008. International Conference on Audio, Language and Image Processing* (IEEE, 2008), pp. 138–142.
- ¹²M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (ACM, New York, NY, 2000), pp. 417–424.
- ¹³L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks* (IEEE, Piscataway, NJ, 2005), pp. 63–70.
- ¹⁴H. Zhang, V. Kong, L. Ren, and J. Jin, "An inter-projection interpolation (IPI) approach for the synchronized moving grid (SMOG) to reduce dose in cone beam CT," *Med. Phys.* **41**(6), 123–124 (2014).
- ¹⁵H. Zhang, F. Kong, L. Ren, and J.-Y. Jin, "An inter-projection interpolation (IPI) approach with geometric model restriction to reduce image dose in cone beam CT (CBCT)," in *Computational Modeling of Objects Presented in Images. Fundamentals, Methods, and Applications* (Springer International Publishing, Switzerland, 2014), pp. 12–23.
- ¹⁶L. Feldkamp, L. Davis, and J. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Am. A* **1**(6), 612–619 (1984).
- ¹⁷R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. (Prentice-Hall, Inc., Upper Saddle River, NJ, 2006).
- ¹⁸V. Kong, H. Zhang, L. Ren, and J. Jin, "An inpaint-based interpolation technique to recover blocked information for cone beam CT with a synchronized moving grid (SMOG)," *Med. Phys.* **41**(6), 131 (2014).
- ¹⁹J. Casey, *Exploring Curvature* (Vieweg, Wiesbaden, 1996).
- ²⁰S. J. LaRoque, E. Y. Sidky, and X. Pan, "Accurate image reconstruction from few-view and limited-angle data in diffraction tomography," *JOSA A* **25**(7), 1772–1782 (2008).
- ²¹E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.* **53**(17), 4777–4807 (2008).