Fast Monte Carlo calculation of scatter corrections for CBCT images

Ernesto Mainegra-Hing and Iwan Kawrakow
Ionizing Radiation Standards, National Research Council of Canada, Ottawa K1A 0R6, Canada
E-mail: mainegra@irs.phy.nrc.ca

Abstract. A very fast Monte Carlo algorithm for the calculation of the scatter contribution in cone beam computed tomography, implemented within the EGSnrc framework, is presented. Based on the combination of several variance reduction techniques, an efficiency improvement of three orders of magnitude over an analog simulation is obtained. A denoising algorithm applied to the computed scatter distribution is shown to further increase the efficiency of the calculation by about a factor of 10. An iterative scatter correction algorithm is proposed and its feasibility is demonstrated on three different phantoms.

1. Introduction
In cone beam computed tomography (CBCT) scatter corrections are commonly performed experimentally by increasing the object-detector air-gap, using anti-scatter grids, or combining both. These methods can reduce the scatter in some cases, but also reduce the primary signal intensity. Beam stop arrays (BSA) techniques have been shown to effectively reduce the scatter-primary ratio by a factor of 14 [1]. However, BSA techniques are not practical since two acquisitions per angle are required increasing the patient dose and acquisition time. Algorithmic methods using analytical models have been also used which are fast, but are of limited use when complex geometries or heterogenous media are involved.

An attractive alternative is to use Monte Carlo (MC) simulations to obtain a more accurate estimate of the scatter distribution. However, previous attempts in the literature report extremely long calculation times [2]. The aim of the algorithm introduced in this paper is to correct a CBCT scan taken under full scatter conditions without previous correction. Extensive use of variance reduction techniques (VRT) for efficient photon transport, denoising the scatter distribution, and a new approach for removing the scatter from the measured scan, bring this type of approach closer to on-line clinical applicability.

2. Methods
2.1. Definitions
A CBCT scan consists of a set of projections onto a 2D detector array around the volume of interest in a 180° or 360° orbit. If \( R_i \) denotes the reading in pixel \( i \) for a given projection and \( S_i \) the corresponding scatter contribution to the signal, then the quantity

\[
a_i = \ln \frac{b_i}{A_i} = \ln \frac{b_i}{R_i - S_i} = \int \mu(l) \, dl
\]
is the attenuation along a line connecting the x-ray source with the detecting element \( i \). The blank scan \( b_i \) is the signal at pixel \( i \) without the phantom and \( \mu \) is the attenuation coefficient at position \( l \) along the line. If the scatter contribution \( S_i \) is known, then the scatter-free scan \( a_i \) (or \( A_i \)) can be computed according to Eq. (1) and given to an image reconstruction algorithm to obtain the 3D distribution of attenuation coefficients in the phantom. The quantity actually measured by the scanner is

\[
r_i = \ln \frac{b_i}{R_i}.
\] (2)

In what follows \( s_i = S_i/b_i \) is a short hand notation and symbols with a tilde denote MC computed quantities, e.g. \( \tilde{r}_i \) is the real scan obtained from the MC simulation, \( \tilde{S}_i \) the scatter, etc. Because the main purpose of this investigation is to assess VRT in the MC simulation and the iterative scatter correction algorithm described in the next section, only scans generated by a MC simulation are used in this paper, a very simple x-ray source is employed (a monoenergetic 60 keV point source) and the detector response is assumed to be given by air-kerma. Use of measured scans, a realistic photon source, and a realistic detector response is left for future investigations.

### 2.2. Iterative scatter correction algorithm

The proposed iterative scatter correction algorithm works as follows:

(i) Set \( a_i = r_i \)

(ii) Reconstruct the phantom using the current estimate of \( a_i \). A filtered back-projection algorithm due to Feldkamp, Davis and Kress[3] is used in this step.

(iii) Compute the scatter-free scan \( \tilde{A}_i \) and the scatter \( \tilde{S}_i \) corresponding to this phantom

(iv) Perform a scatter correction. The first approach considered is

\[
a_i = (1 + \alpha)r_i - \alpha \tilde{r}_i - \ln \frac{\tilde{A}_i}{\tilde{A}_i + \tilde{S}_i}
\] (3)

with \( \alpha \) a relaxation parameter. The second approach is defined via

\[
a_i = r_i - \ln \left[ 1 - \tilde{s}_i \exp \left( \beta r_i + (1 - \beta) \tilde{r}_i \right) \right]
\] (4)

with \( \beta \) a free parameter.

(v) Go to step (ii).

The iteration is stopped when a convergence has been reached, with a suitable termination criterion currently under investigation. In all examples presented in this paper a manual inspection after each iteration is used.

Eq. (3) is easily obtained from Eq. (1) by replacing \( A_i \) with \( \tilde{A}_i \cdot R_i / \tilde{R}_i \), assuming that major differences in \( R_i \) and \( \tilde{R}_i \) are given by differences in their attenuation properties. Finally a relaxation term \( \alpha \cdot (r_i - \tilde{r}_i) \) is introduced to compensate for errors introduced by this assumption. Eq. (4) is Eq. (1) rewritten as a function of \( r_i \), where \( a_i \) has been replaced by \( \tilde{s}_i \) and a relaxation term \( \beta \cdot (r_i - \tilde{r}_i) \) added to the exponent. It is easy to see that when the computed quantities approach the real ones, both methods recover the actual scatter-free scan given in Eq. (1).

The advantage of Eq. (3) is that it is numerically stable, its disadvantage is a relatively slow convergence. The advantage of Eq. (4) is a very fast convergence, its main disadvantage is the fact that small errors in the scatter estimate lead to large errors in \( a_i \) in regions with large attenuation (i.e. large \( r_i \)) as seen from the following equation (\( \beta = 1 \) is used for the sake of simplicity):

\[
\Delta a_i = \frac{x}{1 - x} \frac{\Delta \tilde{S}_i}{\tilde{S}_i}, \quad x = \tilde{s}_i \exp(r_i)
\] (5)

and from the fact that \( x \) approaches unity for large \( r_i \).
2.3. Fast Monte Carlo simulation of photon transport

A new EGSnrc [4] user code, *egs_cbct*, is developed for the purposes of this investigation. *egs_cbct* is written in C++ and uses the EGSnrc C++ class library [5]. Air-kerma at the detector plane is scored efficiently using a track-length estimation, whereby all photons crossing the scoring plane contribute to the kerma. This has been shown elsewhere [6] to be about 20 times more efficient than computing kerma by switching off electron transport and collecting the energy deposited by electrons on the spot. To further improve the efficiency of the kerma estimator, use is made of a VRT known as forced detection. Every time a photon direction is aimed at the scoring field, its contribution from the current position is scored, including the effect of attenuation through the geometry. Interaction splitting combined with Russian Roulette (RR) is employed to increase the number of photons directed towards the detector array while reducing the time spent following photons moving away from the scoring field. Woodcock tracing (a.k.a. delta transport) is also implemented to speed up the transport of those few photons aimed away from the scoring plane that survived RR. Since the forced detection technique introduces large fluctuations in the weights of the photons reaching the scoring plane, a sophisticated scheme is devised, which consist of path-length biasing combined with position dependent splitting numbers to ensure nearly constant weights of the photons contributing to the score. It is worth
noting that all these techniques, to be described in more detail in a forthcoming publication, are true VRT, i.e. they don’t change the result. To ensure their correct implementation, comparisons with analog simulations were performed and no statistically significant differences were found. The combined efficiency gain is about a factor of 200 compared to track-length estimation.

To further decrease the simulation time, the scatter scan $S_i$ is subjected to denoising using a 2D version of the locally-adaptive Savitzky-Goley (LASG) filter presented in Ref. [7]. Due to the fact that in general the scatter distribution is a “well behaved” function without sharp discontinuities, the denoising algorithm performs very well as illustrated in Fig. 1.

![Figure 3.](image_url) A 30 cm$^3$ water phantom with 5 bone inserts as modeled with the EGSnrc C++ class library.

**Figure 4.** Iterative correction shown for a profile across middle slice (64 × 64 × 64, 0.5 cm voxels) for 60 keV x-ray beam on 30cm$^3$ water phantom. Reconstruction done with 72 projections.

### 3. Results

#### 3.1. A water cylinder

As a first test of the proposed scatter correction algorithm a simple homogeneous water cylinder (20 cm diameter × 20 cm length) is considered with the x-ray source rotating around the symmetry axis. Since the phantom is cylindrically symmetric, it is sufficient to compute the 0° projection only, obtaining all other projections by simply copying the 0° scan (128 projections are used in this example). In this first test the correction method given in Eq. (3) with $\alpha = 0$ is employed. Figure 2 (left panel) shows a profile through the middle slice of the reconstructed attenuation image after the 0’th iteration (i.e. using the uncorrected scans $r_i$), 1’st and 5’th iterations. One can see how with each iteration step, the values of the attenuation coefficient approach the correct value of $\mu$ (60 keV) = 0.21 cm$^{-1}$. The large fluctuations observed in the middle of the profile are due to the strong correlation when re-using the 0° projection for all 128 projections. The central points along the rotation axis are the most affected since exactly the same points are used in the reconstruction for this region. If one uses the LASG denoising algorithm to reduce the statistical noise, a much smoother profile is obtained and a faster convergence to the correct value of the attenuation coefficient is observed, as can be seen in the right panel of figure 2.
Figure 5. 0° scans of a water cube with 5 bone inserts with a 64 × 64 detector. (a) Initial uncorrected scan with full scatter. (b) Scatter-free scan obtained from direct simulation which is the goal of the scatter corrections. (c) Result of correcting the scatter without smoothing. (d) A smoothing procedure is used with 10 times less histories as when not smoothing.

3.2. Water cube with 5 bone rods
A more challenging arrangement is a 30 cm³ water cube with five bone inserts as shown in figure 3. In this case the correction algorithm given in Eq. (3) with α = 0 converges too slowly. Using α = 1 initially improves the convergence rate, but after 6 iteration the change in each subsequent iteration is very small. Fig. 4 shows a comparison between the scatter-free scan (black line) and the estimated scatter-free scans after 0, 1 and 5 iterations along a line in the 0° scan. The measured scan (used to obtain the phantom in the 0'th iteration) is also shown with the blue line.

The influence of denoising is illustrated in Fig. 5, which shows the full-scatter scan (a), scatter-free scan (b), and reconstructed scatter-free scans after 5 iterations without (c) and with (d) denoising. The image in Fig. 5d is obtained with 10 times fewer particle histories compared to Fig. 5c, thus indicating that denoising brings another factor of 10 improvement in simulation speed.

Figure 6. The middle slice of the reconstructed attenuation image (left panel) and a profile through the center (right panel) of the phantom described in section 3.3.
3.3. Water sphere with spherical inserts of different composition

The FDK reconstruction algorithm employed in this study introduces significant artifacts near sharp edges. As a consequence, the actual phantom composition in these regions is not correctly reproduced. This in turn introduces errors in the estimated scatter distributions which are exponentially amplified in regions with significant attenuation according to equation 5. Therefore, a geometry with smooth boundaries, which is also more representative for practical applications than the water box with bone inserts from the previous section, is considered as a final example: a 15 cm radius water sphere with four spherical inserts of varying composition (\( \rho = 1.1, 1.5, 1.6, 2.0 \) g/cm\(^3\)). The scatter correction algorithm defined in Eq. (4) with \( \beta = 0.5 \) is employed. However, it is found that the scatter distribution computed from a voxelized representation of the phantom must be corrected before use in Eq. (4). The correction is a simple shift determined from the requirement that the computed scan reproduces the measured scan in regions with no attenuation (i.e. regions that are not behind the scanned object). Figure 6 shows a profile across a central slice of the reconstructed sphere after the third iteration. In this case the algorithm converges faster and is also capable of faithfully reproducing the inserts of different density.

4. Conclusions

The proposed approach to correct CBCT images for scatter using only the results of a Monte Carlo simulation is shown to correctly reproduce the attenuation coefficients. A new EGSnrc code named egs_cbct for performing CBCT scatter computations is developed. The set of variance reduction techniques implemented in egs_cbct increases the efficiency of the simulation by a factor of 4000 compared to a fully analog simulation. Use of denoising for the computed scatter distribution reduces the simulation time by another factor of 10. To put this in perspective, if one were to repeat the calculations described by Jarry et al [2] using egs_cbct, the calculation time would be 39 seconds instead of the 430 hours reported. Due to the limitations of the FDK reconstruction algorithm, objects with round boundaries can be better reproduced than objects with sharp edges. The investigations presented in this paper show that using the signal beyond the objects dimensions as representative of the actual scatter to shift the computed scatter distribution, together with the correction method defined in Eq. (4) produces the best convergence and contrast. Since results reported here can be considered only preliminary, extensive benchmarking of the algorithm with real CBCT images, and a better model of the x-ray source and detector array is needed.

Acknowledgments

We would like to thank Dr. Glenn Wells of the Ottawa Heart Institute for kindly providing the FDK image reconstruction algorithm used in our study.

References

[1] Ning R, Tang X and Conover D 2004 Med. Phys. 31 1195 – 1202
[2] Jarry G, Graham S A, Moseley D J, Jaffray D A, Siewerdsen J H and Verhaegen F 2006 Med. Phys. 33 4320 – 4329
[3] Feldkamp L A, Davis L C and Kress J W 1984 J. Opt. Soc. Am. A1 612 – 619
[4] Kawrakow I 2000 Med. Phys. 27 485 – 498
[5] Kawrakow I 2005 EGSnrc C++ class library Technical Report PIRS–898 National Research Council of Canada Ottawa, Canada
[6] Mainegra-Hing E and Kawrakow I 2006 Med. Phys. 33 3340 – 3347
[7] Kawrakow I 2002 Phys. Med. Biol. 47 3087 – 3104