Treatment of Compton Scattering in Maximum-Likelihood, Expectation-Maximization Reconstructions of SPECT Images

James E. Bowsher and Carey E. Floyd, Jr.

Department of Radiology, Duke University Medical Center, Durham, North Carolina

This paper studies the extent to which lesion contrasts and signal-to-noise ratios in maximum-likelihood, expectation-maximization estimates of SPECT images can be improved by considering Compton scattering when calculating the photon detection probability matrix. Matrices are generated using a Monte Carlo code that realistically models a SPECT imaging system. For cold lesions having true contrasts of one, thirty-six combinations of three lesion sizes, three lesion locations, and four projection-data count levels are considered. Both hot and cold, 2-cm-radius lesions having various values of true contrast are studied at a few count levels and lesion locations. With scatter in the matrix, scatter effects are completely removed from the central regions of lesions whose radius is greater than the full width at half maximum of the spatial resolution. Resolution recovery is initially slower but eventually superior. Percent improvements in contrast are greater for lower-contrast cold lesions and higher-contrast hot lesions. In most cases, signal-to-noise ratios are comparable or better. Correct absolute activity levels are obtained.

J Nucl Med 1991; 32:1285–1291

Compton scattering reduces the contrast of lesions in single-photon emission computed tomography (SPECT) images. Presently, almost all clinical SPECT images are reconstructed using the filtered backprojection (FBP) algorithm. Several scatter compensation techniques primarily intended for use with SPECT images reconstructed by FBP have been considered. (*1–8*) However, much research has been performed on a second class of reconstruction techniques. These techniques form images by solving the set of equations:

$$\sum_i T_{ij}A_i = P_j \qquad \text{Eq. 1}$$

for $\vec{A}$. In Equation 1, A_i is the number of photons emitted from source voxel i and $\vec{A}$ is, therefore, the distribution of radioactive decays that one wishes to image, P_j is the number of photons detected at projection pixel j, and T_{ij} is the probability that a photon emitted from voxel i is detected at pixel j. Included in this second class of reconstruction techniques are algorithms for obtaining maximum-likelihood (ML) (*9–13*), maximum entropy (*14*), and Bayesian (*15–17*) estimates of $\vec{A}$. These techniques are potentially superior to FBP because they allow for a unified, simultaneous treatment of statistical uncertainty in the projection data, prior information about $\vec{A}$, such as non-negativity and nearest-neighbor constraints on the source voxels, and sophisticated models of the SPECT imaging system (*11,18*). The modeling of the system occurs in the calculation of the matrix T of Equation 1. The calculation may, for example, consider finite energy resolution in the gamma camera, spatially varying collimator resolution, and photon attenuation and scatter.

This paper studies the extent to which lesion contrasts and signal-to-noise ratios (SNRs) can be improved by considering Compton scattering when calculating T. Monte Carlo simulations are employed to generate two matrices, T, which are identical except that only one includes scatter effects. Source distributions $\vec{A}$ are simulated, and projection data are obtained as prescribed by Equation 1. The above two matrices are then utilized to reconstruct $\vec{A}$ from $\vec{P}$. Thus, we are evaluating the performance of this method of treating scatter in the ideal case of an exactly correct matrix T, except that direct-plus-scatter data are sometimes reconstructed using the scatter-free matrix. This procedure allows us to evaluate the best possible performance of the method; we know, for example, that any shortcomings which we find in this technique are not caused by an inadequate modeling of the actual Compton scattering or of other aspects of the imaging system such as the detector's resolution. The ultimate test of the method will be its ability to improve clinical images. However, optimally effective clinical implementations will require overcoming several practical problems. We briefly address these problems in the Methods and Discussion sections.

The expectation-maximization (EM) algorithm (*11–13*) for approximating ML estimates of $\vec{A}$ is used. For cold lesions having true contrasts of one, thirty-six combinations of three lesion sizes, three lesion locations, and four projection-data count levels are considered. Both hot and

Received Jun. 6, 1990; revision accepted Nov. 19, 1990.

For reprints contact: James E. Bowsher, PhD, Box 3808, Department of Radiology, Duke University Medical Center, Durham, NC 27710.

cold, 2-cm-radius lesions having various values of true contrast are studied at a few count levels and lesion locations. Reconstructions of point-sources from 100-million-count data are also examined. Except for images reconstructed from noise-free or 100-million-count data, ensembles of images, rather than single images, are generated and evaluated. The use of ensembles reduces statistical fluctuations in some measured quantities, such as contrast, and allows for definitions of noise, signal-to-noise, and root-mean-squared (RMS) error that are more directly related to the tasks of detection and quantitation.

METHODS

Monte Carlo Simulation of a SPECT System

The matrices T were generated using an experimentally verified Monte Carlo code (*19, 20*) and realistic modeling of a SPECT system, including spatially varying collimator resolution and photon attenuation and scatter. The calculations assumed 140 keV photons (which are emitted by ^{99m}Tc, a 20% energy window —126–154 keV), a high-resolution collimator, and a 12-cm radius of rotation for the gamma camera.

Figure 1 shows an appropriate geometry for Monte Carlo calculations that include realistic scatter effects. Photons are allowed to originate with equal probability from any point in a water-filled cylinder of radius 11.3 cm and height 22.0 cm. Detection of these photons is restricted to the vertical range -0.45 cm $\leq y \leq 0.45$ cm of a 28.9-cm^2 projection plane. Direct photons can be detected within the 0.9-cm high projection slice only if they originate within or near the vertical range -0.45 cm $\leq y \leq$ 0.45 cm, whereas scattered photons originating many cm above or below the 0.9-cm high projection slice can be detected. Hence, obtaining the full effect of scatter requires including a broad vertical range of the source. (The matrices employed for this paper were generated assuming a different geometry which was more efficient for Monte Carlo calculations and which yielded essentially the same matrix elements as did calculations based on the geometry of Fig. 1.)

Fully realistic treatments of scatter require allowing for different activity levels at different y values. The number of events

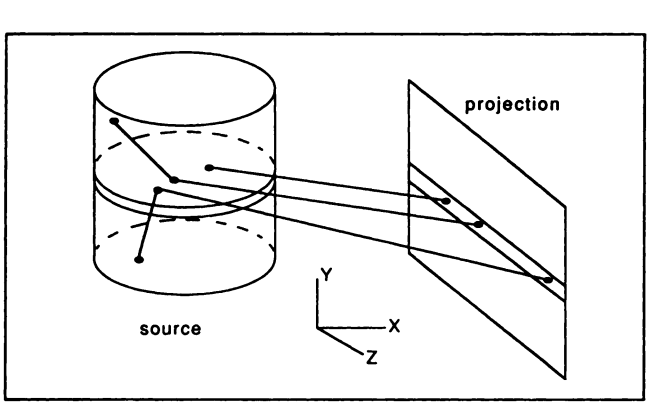

FIGURE 1. A collimator (not shown) allows only those direct photons that originate within the narrow source slice shown to be detected within the projection slice shown. Scattered photons originating over a broad range of y-values can contribute to this projection slice. Hence, obtaining the full effect of scatter requires including a broad vertical range of the source in the Monte Carlo calculations.

detected at projection pixel j and originating in source pixels having the same (z, x) values but different y values would then be $\sum_y A_{zxy}T_{zxy}(j)$, where A_i and T_{ij} of Equation 1 have been rewritten as A_{zxy} and $T_{zxy}(j)$ so that source voxel i may be identified by its z, x, y coordinates. However, this direct extension of the matrix into a third dimension produces a prohibitively large matrix. If when calculating the above summation one assumes that the activities $A_{zxy} = A_{zx}$ are independent of y, this summation reduces to $A_{zx} \sum_y T_{zxy}(j) = A_{zx}T_{zx}(j)$. Thus, we obtained the true, three-dimensional effects of scatter in two-dimensional reconstructions by invoking the above assumption and using phantoms in which the activity did not vary with y.

To calculate the $T_{zx}(j)$, the z-x plane of the source was partitioned into a 64 × 64 grid 28.9 cm wide, and likewise the projection plane was split into 64 pixels along the z-axis. Matrix elements for 60 different projection angles equally spaced over 360° were obtained by exploiting the cylindrical symmetry of the water-filled cylinder. Matrix elements associated with source pixels falling outside of the 11.3-cm radius cylinder were set identically to zero and were not stored, yielding a 60 × 64 × 1976-element, 30-megabyte matrix.

Phantoms, Projection Data, and Reconstructions

Simulated phantom activity distributions were generated on a 64 × 64 grid. Three phantoms, each composed of a single hot pixel located 0, 4, or 8 cm from the image center in an otherwise cold, water-filled cylinder of radius 11.3 cm were utilized because profiles through reconstructions of these phantoms provide a direct observation of resolution recovery. Each of three additional phantoms consisted of three equal-sized, roughly cylindrical, cold lesions having true contrasts of one and centered at the coordinates (z, x) = (0 cm, 0 cm), (4 cm, 0 cm), and (0 cm, 8 cm) and embedded in a uniformly active, hot cylinder of radius 11.3 cm. The cross-sectional areas of the lesions were equivalent to the cross-sectional areas of cylinders having 0.5-cm, 1-cm, and 2-cm radii. Several similar phantoms, composed of 2-cm radius lesions having various amounts of contrast against a uniformly active background were also considered and are described in the Results sections.

Only two matrices were generated, a "direct-only" matrix, which assumed non-scattered photons only, and a "direct-plus-scatter" matrix, which considered both scattered and non-scattered photons. The two calculations were otherwise identical. Direct-only and direct-plus-scatter projection data were obtained by multiplying the vectors $\vec{A}$ specifying the phantom activity distributions into the direct-only and direct-plus-scatter matrices, respectively, as indicated by Equation 1. These data were scaled to a specified total number of counts. Then, the noise-included value of each projection pixel was obtained by sampling from a Poisson distribution having a mean at the original, scaled pixel value. One-hundred-million-count, direct-plus-scatter and direct-only projection sets were generated for most of the above phantoms. For some phantoms, multiple, independent Poisson samplings were performed to obtain ensembles of twenty 50,000-count, 200,000-count, and 800,000-count, direct-plus-scatter projection sets.

Images were reconstructed by the ML-EM algorithm using the same two matrices that were employed to generate the projection data. This procedure assured that the reconstruction matrix modeled the photon transport exactly with the one exception that in some cases direct-plus-scatter projection data were reconstructed with the direct-only matrix. Ensembles of images were recon-

structed from the ensembles of projection data. We refer to reconstructions as either "direct-only reconstructions" or "direct-plus-scatter reconstructions" depending on which matrix was employed for the reconstruction.

For reference, some projection sets and ensembles of projection sets were also reconstructed by FBP. For the FBP reconstructions, the usual "ramp" filter was applied to the projection data; low-pass filtering was not performed.

Measurements

For each lesion in a single reconstructed image, the signal S was defined as the average pixel value over a region of interest (ROI) covering part or all of the lesion. The background B was defined as the average pixel value in a ring surrounding, and extending a few cm beyond, the lesion. Contrast was defined as

$$C = \frac{|B - S|}{B}.$$

For ensembles of images reconstructed from 50,000-count, 200,000-count, and 800,000-count data, the average $\bar{C}$ and standard deviation ΔC of contrasts across the ensembles of 20 reconstructions were determined as a function of iteration. Noise was evaluated as ΔC. A SNR was defined as $\bar{C}/\Delta C$. Also considered was the RMS error in contrast, which may be estimated from the ensemble of images as:

$$\text{RMS error} = \sqrt{\frac{\sum_i (C_T - C_i)^2}{N}}, \qquad \text{Eq. 2}$$

where C_T is the true contrast, C_i is the contrast measured from a single image and N is the number of images in the ensemble. The RMS error can be shown to equal

$$\sqrt{\Delta C^2 + \text{bias}^2},$$

where bias = $(C_T - \bar{C})$ is the bias in the contrast measurement.

Ensembles of images allow one to include low-count projection sets in studies of scatter effects as a function of lesion location, lesion size, and number of projection counts. Random fluctuations in contrast can be comparable to or greater than the variations in contrast with respect to the above parameters. In such cases, one must study contrasts averaged over ensembles of images.

The use of ensembles also allows the above definitions of noise, RMS error, and SNR. Regarding quantitation, ΔC is the statistical error in contrast measurements, and the RMS error $\sqrt{\Delta C^2 + \text{bias}^2}$ is arguably a better measure of performance at the quantitation task than are $\bar{C}$ and the bias, which consider only the accuracy and not the precision of the measurement. As for detection, definitions of SNR similar to $\bar{C}/\Delta C$ have been discussed by other authors (*21*, *22*). Here, we note briefly that a lesion will be consistently detectable only if its average contrast ($\bar{C}$) against the background significantly exceeds random fluctuations (ΔC) in that contrast.

RESULTS

For the phantom composed of a single point source located 4 cm from the image center, Figure 2 compares profiles through the images at iteration 500 of a direct-only reconstruction of direct-only data (large circles, dotted line), a direct-only reconstruction of direct-plus-scatter data (dashed line), and a direct-plus-scatter reconstruction of direct-plus-scatter data (solid line). Comparison of the dotted and dashed curves shows the deleterious effects of untreated scatter. A long scatter tail is added, and the narrow, high-amplitude portion of the profile, which is influenced primarily by the spatial resolution of the detector's collimator and crystal, is broadened. Both the spatial resolution of the detector and the high amplitude portion of the profile may be roughly characterized by their FWHM. Beyond about 100 iterations, the FWHM of the profiles continue to improve but are in the range of the FWHM of the detector's spatial resolution, about 1 cm. By iteration 500, the direct-plus-scatter reconstruction (solid curve) provides a more narrow FWHM than does the direct-only reconstruction and removes essentially all of the scatter tail. However, at the earliest iterations, direct-plus-scatter reconstructions are slower to improve resolution. Direct-only reconstructions provide smaller scatter tails until about iteration 4 and more narrow FWHM up through iterations 50, 23, and 17 for point sources located 0 cm, 4 cm, and 8 cm from the image center, respectively. Below, we show that similar effects are present in reconstructions of phantoms comprised of lesions in a hot background.

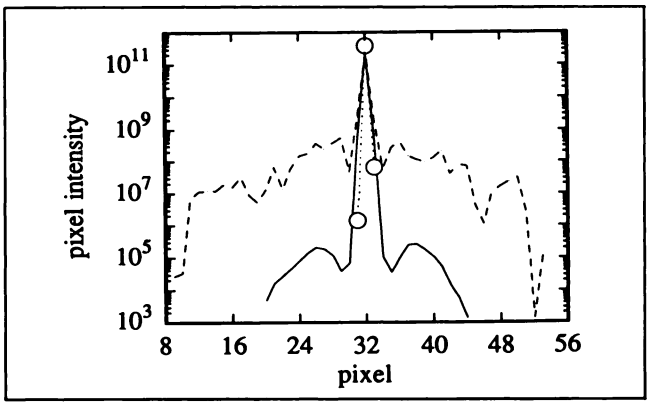

FIGURE 2. Profiles through the images at iteration 500 of reconstructions of a point source located 4 cm from the image center. The dotted, dashed, and solid curves correspond to direct-only reconstructions of direct-only data, direct-only reconstructions of direct-plus-scatter data, and direct-plus-scatter reconstructions of direct-plus-scatter data, respectively.

We now consider the problem of obtaining quantitative estimates of lesion contrast. Often considered particularly important is the removal of scatter effects from that central region of a lesion where every pixel is at least one FWHM away from the lesion's edge, because in this region, the signal is only weakly biased by the finite spatial resolution of the detector, and, if scatter effects can be removed, an accurate quantitative measurement of the signal can be made. To study whether direct-plus-scatter reconstructions can produce unbiased estimates of the signal in this central region, 1.4-cm radius ROIs were employed to determine the contrasts of 2-cm radius, hot and cold lesions having true contrasts of 4.0 and 0.5, respectively. For the hot lesion phantom, the solid, dashed, and dotted curves of Figure 3 plot profiles through the true image and through

iteration 500 of direct-plus-scatter and direct-only reconstructions from noise-free data. The direct-only reconstruction is renormalized by 0.7. With noise-free data, the direct-plus-scatter reconstructions overestimate the contrast by 4% (4%) and 2% (1%) for the cold (hot) lesions located at the image center and 8 cm from the center, respectively. These slight overestimations are primarily, if not completely, artifacts of the ringing, as seen in Figure 3, that occurs along the edges of the lesions. Somewhat larger ROIs yield slightly greater overestimations, while somewhat smaller ROIs result in slight underestimations. The direct-plus-scatter reconstructions apparently remove from the central regions of lesions all bias associated with scatter effects. This result is not surprising given the essentially complete removal of the scatter tail seen in the plots of Figure 2.

In general, noise affects the bias only weakly if at all. For reconstructions of the $C_T = 0.5$, cold-lesion phantom from 200,000-count data, the overestimates in contrast at iteration 500 change from 4% of the true contrast (with noise-free data) to 2% ± 5% and from 2% to 1% ± 5% for lesions located at the center and 8 cm from the center, respectively; these changes may simply reflect the statistical uncertainties in the measured biases for the 200,000-count data. However, whereas bias generally decreases with iteration, ΔC increases, and minimizing the RMS error requires evaluating the contrast at an earlier iteration and thus accepting a higher bias. For example, the minimum RMS error in the contrast of the lesion located 8 cm from the center occurs at iteration 38, where the bias, ΔC, and RMS error are 4%, 10%, and 11%, respectively, of the true contrast, compared with a bias of 1% at iteration 500.

The $C_T = 1$, cold lesions are a special case in which noise does affect bias. By iteration 500 of direct-plus-scatter reconstructions from 50,000-count, 200,000-count, and 800,000-count data, biases roughly equal ΔC. This result suggests that these biases are simply a consequence of ΔC; since $C_T = 1$, all fluctuations must be in the direction of less contrast. At all three count levels, RMS errors are approximately constant and near their minima from iteration 200 to 500. Direct-plus-scatter RMS errors are 25%–70% lower than direct-only RMS errors, with the larger percent reductions occurring for lesions closer to the image edge and for higher count levels.

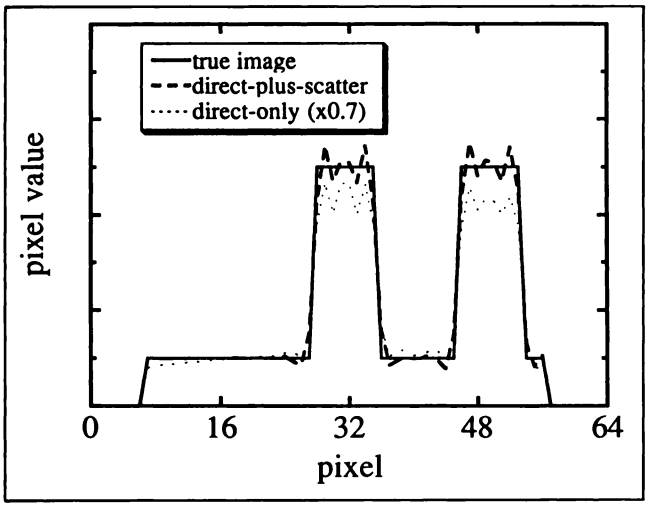

FIGURE 3. The solid curve is a profile through a phantom that contains two $C_T = 4$, hot lesions. The heavy, dashed and light, dotted curves are profiles through the images at iteration 500 of direct-plus-scatter and direct-only reconstructions. Noise-free, direct-plus-scatter projection data were used. The profile associated with the direct-only reconstruction is renormalized by 0.7.

The detectability of lesions probably depends not only on restoring the true contrast in the central regions of lesions but also on the sharpness with which the edges of lesions are recovered. Therefore, below we employ full ROIs, which cover the entire lesion, to evaluate contrast. These contrast determinations are influenced by scatter but also by the limited spatial resolution of the detector, which causes blurring across the edges of lesions.

The three panels of Figure 4 plot average contrast, ΔC, and signal-to-noise ratio as a function of iteration for direct-only reconstructions (dotted lines) and direct-plus-scatter reconstructions (solid lines) from 200,000-count, direct-plus-scatter projection data of 1-cm radius, $C_T = 1$, cold lesions located at the image center (open diamonds) and 8 cm from the center (dark circles). The shapes of these curves are representative of the shapes seen for most lesion sizes and locations and for most count levels. The horizontal lines indicate FBP results.

At early iterations, direct-only reconstructions yield better average contrasts than do direct-plus-scatter reconstructions. Crossovers to better contrasts for direct-plus-scatter reconstructions occur later for smaller lesions and for lesions located closer to the image center. The location dependence can be seen in Figure 4A; the crossover for the lesion at the image center (open diamonds) occurs at iteration 125, compared with iteration 18 for the lesion located 8 cm from the center (dark circles). Crossovers for

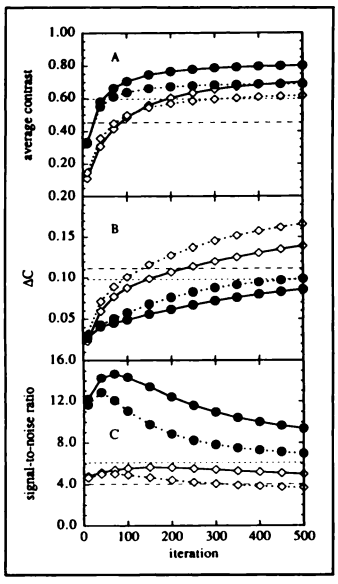

FIGURE 4. Average contrast, ΔC, and SNR as a function of iteration for reconstructions from 200,000-count, 1-cm radius lesion, direct-plus-scatter projection data. The open diamonds and dark circles correspond to the lesions located at the image center and 8 cm from the center, respectively, with the dotted and solid connecting curves indicating, respectively, direct-only and direct-plus-scatter, reconstructions. The long-dashed and short-dashed horizontal lines give the results for the central and edge lesions, respectively, of FBP reconstructions of the same data.

TABLE 1
Iterations at which Full-ROI, $C_T = 1$, Cold-Lesion Contrasts Obtained from Direct-Plus-Scatter Reconstructions First Exceed Those Generated by Direct-Only Reconstructions

Location	Lesion radius (cm)		
	2.0	1.0	0.5
8 cm out	10, 10, 10*	18, 18, 15	20, 20, 20
4 cm out	40, 40, 35	80, 80, 70	130, 140, 100
center	60, 65, 55	125, 125, 120	160, –, –†

* The three numbers are for 800,000-, 200,000-, and 50,000-count data, respectively.
† The direct-only contrasts are always better.

various lesion sizes and locations are listed in Table 1. In general, the locations of these crossovers are roughly independent of the total number of projection counts. The varied results associated with the 0.5-cm-radius lesion located at the image center are perhaps caused by statistical uncertainties in determining these crossovers. By iteration 500, direct-plus-scatter contrasts are about 11% better for 2-cm radius lesions and 15% better for 1 cm radius lesions. These percentage improvements are roughly independent of lesion location and of the total number of projection counts.

Figure 5 provides another perspective on the ability, as a function of lesion size and location, of direct-plus-scatter and direct-only reconstructions to recover full-ROI, cold-lesion contrasts. Full-ROI contrasts obtained at iteration 500 of reconstructions from 100-million-count, $C_T = 1$, cold lesion projection data are plotted against lesion radius. The dotted, dashed, and solid curves are associated with direct-only reconstructions of direct-only data, direct-only reconstructions of direct-plus-scatter data, and direct-plus-scatter reconstructions of direct-plus-scatter data, respectively. Comparison of the dotted and dashed curves indicates the deleterious effects of scatter. It is interesting to note that the additive decrease in contrast is greatest with lesions whose radius roughly equals the FWHM of the detector's resolution. This decrease is typically 0.2. Except for smaller lesions located at or near the image center, direct-plus-scatter reconstructions of direct-plus-scatter data (solid lines) do about as well as direct-only reconstructions of scatter-free data (dotted lines). The open circles and triangles plot contrasts for FBP reconstructions of the direct-only and direct-plus-scatter data, respectively. The ML-EM reconstructions yield contrasts which are significantly better than those obtained with FBP.

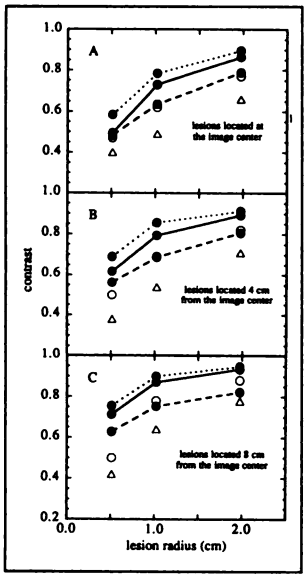

FIGURE 5. Contrasts obtained at iteration 500 of reconstructions of 100-million-count projection data plotted against lesion radius. Panels A, B, and C are associated with lesions located at the image center, 4 cm from the center, and 8 cm from the center. The dotted, dashed, and solid connecting curves correspond to, respectively, direct-only reconstructions of direct-only data, direct-only reconstructions of direct-plus-scatter data, and direct-plus-scatter reconstructions of direct-plus-scatter data. The open circles and triangles are associated with FBP reconstructions of direct-only and direct-plus-scatter data, respectively.

The observations that resolution is recovered more slowly near the image center and that the inclusion of scatter effects in the reconstruction matrix slows down resolution recovery at early iterations allow for a coherent understanding of many of the above results. Consistent with the later observation, lesion contrasts and profiles through reconstructions of point sources are initially better with direct-only reconstructions. Consistent with both observations, crossovers to better contrasts and profiles for direct-plus-scatter reconstructions occur later for lesions located closer to the image center. The dependence of contrast crossovers on lesion size is perhaps due to the increasingly dominant influence of finite resolution on successively smaller lesions. Both the size and location dependence of contrast crossovers underlie the results of Figure 5; for the smallest lesions located at or near the image center, the direct-plus-scatter reconstructions yield, by iteration 500, contrasts which are only slightly better than those obtained with direct-only reconstructions.

How do the above full-ROI results differ at other values of true contrast? For direct-plus-scatter and direct-only reconstructions, respectively, the solid and dashed curves of Figure 6 plot, against true contrast, contrasts obtained at iteration 500 divided by true contrasts; cold lesion results are presented in Figure 6A while hot lesion results are given in Figure 6B. The reconstructions utilized 100-million-count, direct-plus-scatter projection data generated from phantoms consisting of a single, 2-cm radius lesion located 5.4 cm from the image center. The reductions in contrast caused by untreated scatter (dashed

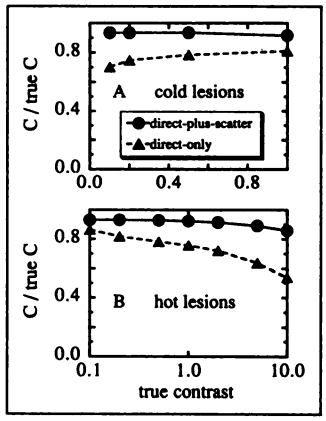

FIGURE 6. For direct-plus-scatter and direct-only reconstructions, respectively, the solid and dashed curves plot, against true contrast, contrasts obtained at iteration 500 divided by true contrasts; cold (hot) lesion results are presented in part A (B). The reconstructions utilized 100-million-count, direct-plus-scatter projection data generated from phantoms consisting of a single, 2-cm radius lesion located 5.4 cm from the image center.

curves) are greater for lower-contrast cold lesions and for higher-contrast hot lesions. Except with $C_T = 5$ and 10 hot lesions, direct-plus-scatter reconstructions recover better than 90% of the true contrast. Direct-plus-scatter contrasts are better than direct-only contrasts by 33%, 25%, 19%, and 13% for $C_T = 0.1$, 0.2, 0.5, and 1 cold lesions and by 8%, 14%, 19%, 22%, 27%, 39%, and 59% with $C_T = 0.1$, 0.2, 0.5, 1, 2, 5, and 10 hot lesions.

For $C_T = 1$ cold lesions, ΔC is typically 10% to 20% lower with direct-plus-scatter reconstructions (Fig. 4B), but at other values of true contrast, direct-plus-scatter reconstructions are noisier than direct-only reconstructions. A phantom containing four 2-cm radius, cold lesions placed 5.4 cm off center and having contrasts of 0.1, 0.2, 0.5, and 1 and a similar phantom containing four hot lesions with contrasts of 0.2, 1, 5, and 10 were considered. For each phantom, ensembles of twenty, 200,000 count projection sets were reconstructed. For the $C_T = 0.1$, 0.2, and 0.5 cold lesions, direct-plus-scatter uncertainties ΔC first exceed direct-only uncertainties at about iteration 25 and are, beyond 100 iterations, 10%–15% larger. With the hot lesions, direct-plus-scatter uncertainties surpass direct-only uncertainties somewhere between 20 and 50 iterations. Higher true contrast implies a higher ratio of direct-plus-scatter ΔC values to direct-only ΔC values. For $C_T = 0.2$, 1, 5, and 10, direct-plus-scatter uncertainties ΔC are, at iteration 200, greater than direct-only uncertainties by 17%, 18%, 60%, and 90%, respectively.

With the $C_T = 1$, cold-lesion data, direct-plus-scatter SNRs are better at all iterations and, at some iterations, exceed direct-only SNRs by 20% to 40% (Fig. 4C). Direct-plus-scatter SNRs peak later and more gradually than direct-only SNRs. This effect occurs because direct-plus-scatter contrasts peak more slowly and thus balance out increases in noise over a broader range of iterations. The peak direct-plus-scatter SNRs are typically 10%–14% higher than the peak direct-only SNRs. The influence of direct-plus-scatter reconstructions on SNRs was also studied at other values of true contrast using the 200,000-count data discussed in the previous paragraph. For $C_T = 0.1$, 0.2, and 0.5 cold lesions, direct-plus-scatter SNRs are, at all iterations, better than direct-only SNRs, typically by 2% to 12%. Results are more varied with the hot lesions, but briefly, direct-plus-scatter and direct-only reconstruction yield similar SNRs for $C_T = 0.2$ and 1, whereas with $C_T = 5$ and 10 direct-plus-scatter SNRs are initially slightly (<10%) better than direct-only SNRs but become worse after about 60 iterations and are, by iteration 500, 15% and 25% worse, respectively.

Some applications of SPECT imaging require determinations of absolute activity levels, rather than merely contrasts (which are measures of relative activities in two regions of an image). The inclusion of scatter effects in the matrix T allows for these determinations. Scatter causes the detection of photons that would otherwise be excluded by the collimator. For the geometry considered here, a water-filled cylinder of radius 11.3 cm, scatter accounts for 23% of all detected photons and for 40% of the detected photons that originate near the center of the phantom. If scatter effects are ignored when calculating the photon detection probability, these probabilities will be too low, and the reconstruction will overestimate activity levels. For example, direct-only reconstructions of a uniformly active, 11.3-cm-radius source distribution overestimate the activities at the image center and 8 cm from the image center by 38% and 31%, respectively, whereas the direct-plus-scatter reconstructions yield correct absolute activity levels at all locations.

DISCUSSION

We began this paper by noting a class of reconstruction techniques that generate images by solving Equation 1, and we have specifically considered the EM algorithm for approximating the ML solution to Equation 1. FBP reconstructions assume that the detected events in projection pixels represent line integrals through the source activity distribution; after the image is formed, the deleterious effects of this false assumption can be alleviated somewhat by, for example, attenuation and scatter compensation methods. On the other hand, techniques that solve Equation 1 can incorporate a realistic model of the SPECT imaging system into the matrix T and thus into the reconstruction itself. Figures 4 and 5 suggest that ML estimates are better than standard FBP reconstructions. In this paper, we have studied the inclusion of Compton scattering effects into the matrix T. Several important results were obtained. Perhaps most significantly, scatter effects are completely removed from those central regions of lesions where each pixel is at least one FWHM away from the lesion edge, thus allowing unbiased, quantitative measurements of contrast for lesions whose radius exceeds one FWHM. Another interesting and important effect is that although direct-plus-scatter reconstructions eventually yield superior resolution, their resolution recovery is slower at early iterations. Consequently, for smaller lesions located closer to the image center, many iterations may be required before direct-plus-scatter reconstructions yield better contrasts than do direct-only reconstructions. The significance of scatter was shown to depend on the lesion's true contrast; losses in contrast caused by untreated scatter are greater for lower-contrast cold lesions and higher-contrast hot lesions, but whatever the true contrast, direct-plus-scatter reconstructions recover most of the contrast lost to scatter, except in the case of smaller lesions located closer to the image center. For $C_T = 1$, cold lesions direct-plus-scatter reconstructions produce smaller uncertainties ΔC, but at other values of true contrast, direct-plus-scatter reconstructions are noisier than direct-only reconstructions. In most cases, SNRs for direct-plus-scatter reconstructions are comparable to or better than those for direct-only reconstructions. Finally, direct-plus-scatter reconstructions yield correct absolute activity levels.

The inclusion of Compton scattering effects in the reconstruction matrix does present several practical problems. First, for a one-source-slice to one-projection-slice matrix such as that considered here, about 95% of the matrix elements are non-zero when scatter effects are considered, compared with only about 5% when scatter effects are ignored. If scatter effects are not included in the matrix, data compression techniques can be employed to reduce the memory requirements of the matrix, perhaps by a factor of ten, say from 30 megabytes to 3. Furthermore, as discussed in the Methods section, fully realistic treatments of scatter will require matrix elements specifying the probability of photons that were emitted from a given source slice (that is, a given y value) being detected in a different projection slice. Such a matrix would probably be impractical since it would be many times larger than the 30-megabyte matrix considered here. On the other hand, more compact representations of this scatter information may be possible. Scatter introduces a small but non-zero probability for a photon emitted from a given source location being accepted at almost any position on the detector. This probability varies slowly as a function of source location and detection location and therefore could perhaps be stored using coarse source and projection grids; alternatively, one might store the coefficients of low-order polynomial fits to these probabilities. Second, methods for obtaining, in an acceptable amount of computation time, the scattered-photon detection probabilities of a specific patient have not been worked out. Research on these problems is underway. This paper examines the degree of and the nature of the improvements in image quality that can be expected from these efforts.

ACKNOWLEDGMENT

This work was supported by NIH grant CA46856.

REFERENCES

1. Hamill JJ, Devito RP. Scatter reduction with energy-weighted acquisition. *IEEE Trans Nucl Sci* 1989;36:1334–1339.
2. Gagnon D, Todd-Pokropek A, Arsenault A, Dupras G. Introduction to holospectral imaging in nuclear medicine for scatter subtraction. *IEEE Trans Med Imag* 1980;8:245–250.
3. Gilardi MC, Bettinardi V, Todd-Pokropek A, Milanesi L, Fazio F. Assessment and comparison of three scatter correction techniques in single photon emission computed tomography. *J Nucl Med* 1988;29:1971–1979.
4. Msaki P, Axelsson B, Dahl CM, Larsson SA. Generalized scatter correction method in SPECT using point scatter distribution functions. *J Nucl Med* 1987;28:1861–1869.
5. Jaszczak RJ, Floyd CE, Coleman RE. Scatter compensation techniques for SPECT. *IEEE Trans Nucl Sci* 1985;32:786–793.
6. Floyd CE, Jaszczak RJ, Greer KL, Coleman RE. Deconvolution of Compton scatter in SPECT. *J Nucl Med* 1985;26:403–408.
7. Axelsson B, Msaki P, Israelsson A. Subtraction of Compton-scattered photons in single-photon emission computerized tomography. *J Nucl Med* 1984;25:490–494.
8. Jaszczak RJ, Greer KL, Floyd CE, Harris CC, Coleman RE. Improved SPECT quantification using compensation for scattered photons. *J Nucl Med* 1984;25:893–900.
9. Llacer J, Veklerov E. Feasible images and practical stopping rules for iterative algorithms in emission tomography. *IEEE Trans Med Imag* 1989;8:186–193.
10. Snyder DL, Miller MI, Thomas LJ Jr, Politte DG. Noise and edge artifacts in maximum-likelihood reconstructions for emission tomography. *IEEE Trans Med Imag* 1987;6:228–238
11. Lange K, Carson R. EM reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomogr* 1984;8:306–316.
12. Shepp LA, Vardi Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imag* 1982;1:113–122.
13. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc B* 1977;39:1–38.
14. Kemp M. Maximum entropy reconstructions in emission tomography. *Medical Radionuclide Imaging* 1980;1:313–323.
15. Liang Z, Jaszczak RJ, Greer KL. On Bayesian image reconstruction from projections: Uniform and nonuniform *a priori* source information. *IEEE Trans Med Imag* 1989;8:227–235.
16. Geman S, McClure DE. Statistical methods for tomographic image reconstruction. *Proc 46th ISI* 1987;52.
17. Levitan E, Herman GT. A maximum *a posteriori* probability expectation maximization algorithm for image reconstruction in emission tomography. *IEEE Trans Med Imag* 1987;6:185–192.
18. Floyd CE, Jaszczak RJ, Greer KL, Coleman RE. Inverse Monte Carlo as a unified reconstruction algorithm for ECT. *J Nucl Med* 1986;27:1577–1585.
19. Manglos SH, Floyd CE, Jaszczak RJ, Greer KL, Harris CC, Coleman RE. Experimentally measured scatter fractions and energy spectra as a test of Monte Carlo simulations. *Phys Med Biol* 1987;32:335–343.
20. Beck JW, Jaszczak RJ, Coleman RE, Starmer CF, Nolte LW. Analysis of SPECT including scatter and attenuation using sophisticated Monte Carlo modeling methods. *IEEE Trans Nucl Sci* 1982;29:506–511.
21. Wagner RF. Decision theory and the detail signal-to-noise ratio of Otto Schade. *Photo Sci Eng* 1978;22:41–46.
22. Rose A. *Vision: human and electronic.* New York: Plenum Press; 1973.

EDITORIAL

Correction for Patient Compton Scattering—Current Status

Methods of correcting for the physical fact of Compton scattering of gamma rays within a patient with subsequent detection by an Anger camera are always extra work. The justifications for this extra work are (1) anticipated improvement of contrast in the image and (2) the potential for accurate quantification if the attenuation correction is also correct.

Techniques for Compton-scatter correction can be classified as pre-, during- or post-reconstruction. The extra work involved in the method is not directly related to the type. Pre-reconstruction methods include those of Gagnon et al. (*1*) and Koral et al. (*2*), which require acquisition of separate energy spectra for individual locations on the face of the Anger camera. The early one-dimensional projection convolution followed by subtraction of Axelsson et al. (*3*) and the later two-dimensional version by Msaki et al. (*4*) are also pre-reconstruction methods.

During-reconstruction methods are represented by (1) the true de-

Received Mar. 7, 1991; accepted Mar. 7, 1991.

For reprints contact: Kenneth F. Koral, Division of Nuclear Medicine, University of Michigan Medical School, Ann Arbor, MI 48109-0552.