

Maximum-Likelihood Transmission Image Reconstruction for Overlapping Transmission Beams

Daniel F. Yu, Jeffrey A. Fessler*, *Senior Member, IEEE*, and Edward P. Ficaro

Abstract—In many transmission imaging geometries, the transmitted “beams” of photons overlap on the detector, such that a detector element may record photons that originated in different sources or source locations and thus traversed different paths through the object. Examples include systems based on scanning line sources or on multiple parallel rod sources. The overlap of these beams has been disregarded by both conventional analytical reconstruction methods as well as by previous statistical reconstruction methods. We propose a new algorithm for statistical image reconstruction of attenuation maps that explicitly accounts for overlapping beams in transmission scans. The algorithm is guaranteed to monotonically increase the objective function at each iteration. The availability of this algorithm enables the possibility of deliberately increasing the beam overlap so as to increase count rates. Simulated single photon emission tomography transmission scans based on a multiple line source array demonstrate that the proposed method yields improved resolution/noise tradeoffs relative to “conventional” reconstruction algorithms, both statistical and nonstatistical.

Index Terms—Attenuation maps, penalized-likelihood, multiple source, SPECT.

I. INTRODUCTION

TO reconstruct quantitatively accurate images of radioisotope emission distributions in single photon emission tomography (SPECT), one must compensate for the effects of photon absorption or attenuation. Accurate attenuation correction requires good attenuation maps, and one can reconstruct such maps from transmission scan measurements obtained either prior to or simultaneously with the SPECT emission scan.

Several source/detector configurations for SPECT transmission scans have been investigated, including a single fixed line source opposite a symmetric fan-beam collimator, used in triple-head SPECT cameras, a scanning line source for orthogonal dual-head cameras, and offset line sources opposite asymmetric fan-beam collimators. Cellar *et al.* [1] describe an alternative geometry based on several fixed-position collimated line sources opposing a parallel-beam collimator. In that system design, the source collimation was selected to minimize overlap on the detector of the transmitted “fan-beams.” They

then applied the filtered back-projection (FBP) algorithm to reconstruct the attenuation map (an ART algorithm was also mentioned without details). This source collimation has the undesirable consequence of very nonuniform count profiles, as shown in [1, Fig. 2]. It is natural to expect that higher and more uniform count profiles could lead to better reconstructed attenuation maps *if* the reconstruction method properly models the beam overlap.

In both the scanning line source geometry and the geometry of Cellar *et al.* [1], there can be overlap of the beam footprints, as illustrated in Fig. 2. Previously published statistical algorithms for transmission tomography, e.g., [2]–[8], are inapplicable to the multiple source problem when the beams overlap. In this paper we formulate a statistical model for multiple-source transmission measurements with arbitrary overlapping beams, and then derive an iterative algorithm for maximizing the likelihood (or a regularized variant thereof). The log-likelihood is not necessarily globally concave, which usually precludes proofs of convergence to a global maximum. The algorithm that we present is guaranteed to increase the likelihood at every iteration, and the set of fixed points of the algorithm is the same as the set of stationary points of the objective function. The algorithm also satisfies the continuity conditions of Meyer [9]. Therefore, by the convergence results in [9], the proposed algorithm produces a sequence of estimates that converge from any nonnegative initial image to a stationary point of the objective, provided the set of stationary points is not a continuum. This is nearly as strong of a convergence result as one might expect for a possibly nonconcave objective function. In Section II, we give the statistical model and the proposed maximum-likelihood algorithm; in Section III, we present some simulation results; in Section IV, we present the results from our preliminary study on optimal source collimation angles for a fixed system setup; and finally Section V is discussion and conclusion.

II. STATISTICAL MODEL

Let Y_i denote the number of photons counted by the i th detector element¹ during the transmission scan, for $i = 1, \dots, N$, where N is the number of measurement elements. Each detector element conceivably may count photons that originated in any of the $M \geq 1$ sources. We assume that separate blank scans are available for each of the sources (or source positions for a scanning line source). This information is essential for unscrambling the multiplexing of overlapping beams. Let b_{im} denote the

Manuscript received October 21, 1999; revised August 15, 2000. This work was supported in part by the National Institutes of Health (NIH) under Grants CA-60 711 and CA-54 362. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was V. Johnson. *Asterisk indicates corresponding author.*

D. F. Yu and E. P. Ficaro are with the University of Michigan, Ann Arbor, MI 48109-2122 USA.

*J. A. Fessler is with the Department of Electrical and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122 USA (e-mail: fessler@umich.edu).

Publisher Item Identifier S 0278-0062(00)09747-0.

¹Each “detector element” corresponds to a unique radial position and view angle, i.e., for typical two-dimensional reconstruction $N = N_r N_\theta$ where N_r is the number of radial samples along the detector and N_θ is the number of view angles or “steps.”

mean number of photons that would be observed during a transmission scan by the i th detector originating in the m th source in the absence of any patient in the scanner. Typically, the b_{im} s would be determined by a periodic calibrating “blank scan,” performed separately for each of the M sources, and then scaled by the relative durations of the blank scan and transmission scans. However, we ignore any statistical uncertainty in the b_{im} s and treat them as known constants. This assumption is reasonable provided the blank scans are sufficiently lengthy.

Let $\mu^{\text{true}} = [\mu_1^{\text{true}}, \dots, \mu_p^{\text{true}}]$ denote the vector of unknown attenuation coefficients for each of the p pixels or voxels in the attenuation map (or the coefficients of some other basis for the attenuation distribution such as B-splines [10]). The line integral between the m th source and the i th detector location through the attenuating object is approximated by the following sum:

$$[\mathbf{A}^m \mu]_i = \sum_{j=1}^p a_{ij}^m \mu_j$$

where $\mathbf{A}^m = \{a_{ij}^m\}$ is a $N \times p$ matrix with nonnegative elements and the a_{ij}^m s represent line-lengths or normalized strip-intersection areas.² Thus, by Beer’s law the “survival probability” for a photon transmitted from the m th source in the direction of the i th detector is (approximately) $\exp(-[\mathbf{A}^m \mu]_i)$.

We assume the Y_i s have independent Poisson distributions

$$Y_i \sim \text{Poisson}\{\bar{y}_i(\mu^{\text{true}})\}$$

where the means are given by

$$\bar{y}_i(\mu) = \left[\sum_{m=1}^M b_{im} \exp(-[\mathbf{A}^m \mu]_i) \right] + r_i. \quad (1)$$

The r_i s are nonnegative constants that one can include to account for the mean contributions of scatter,³ room background, and emission crosstalk [11].

The summation over m in (1) allows for arbitrary overlap of the beams transmitted from each source. Nonoverlapping beams would correspond to the assumption that if $b_{im} \neq 0$, then $b_{ik} = 0$ for all $k \neq m$, i.e., $b_{im}b_{ik} = 0$ for all $k \neq m$.

Under the above statistical model, given a particular measurement realization $Y = [y_1, \dots, y_N]$, we can write the log-likelihood for μ in the following convenient form:

$$L(\mu, Y) = \sum_{i=1}^N h_i \left(\sum_{m=1}^M u_{im}(\mu) \right), \quad h_i(t) = y_i \log t - t, \\ u_{im}(\mu) = b_{im} \exp(-[\mathbf{A}^m \mu]_i) + r_i/M \quad (2)$$

ignoring constants independent of μ . Since the form of this log-likelihood is sufficiently different from the usual models

²Normalized by strip width.

³One could, for example, estimate the scatter contribution to the r_i s by the triple energy window method. Since scatter is a spatially smooth function, one can safely smooth such scatter estimates \hat{r}_i fairly heavily, so generally the variability in the \hat{r}_i s will be much smaller than that of the Y_i s. Thus we treat the \hat{r}_i s as constants, hence the notation r_i .

for emission tomography and transmission tomography [4], previously derived algorithms for maximum-likelihood estimation are not directly applicable to this problem.

One could easily derive an expectation-maximization (EM) algorithm [12], [13] that would monotonically increase the likelihood $L(\mu, Y)$ for this problem, generalizing [4]. However, the convergence would be as painfully slow and the M-step as difficult as the usual transmission EM algorithm. Instead, we propose an algorithm based on an extension of our recent work on paraboloidal surrogates methods [2], [3]. For even faster “convergence” one could apply ordered subsets ideas [14].

Because of the ill-posedness of the reconstruction problem, a penalty term is usually added to the likelihood to encourage piecewise smoothness in the reconstructed image, resulting in the following objective function:

$$\Phi(\mu, Y) = L(\mu, Y) - \beta R(\mu). \quad (3)$$

Most roughness penalties $R(\mu)$ can be expressed in the following general form:

$$R(\mu) = \sum_{k=1}^K \psi_k([\mathbf{C}\mu]_k) \quad (4)$$

where the ψ_k s are potential functions acting as a norm on the “soft constraints” $\mathbf{C}\mu \approx 0$ and K is the number of such constraints. The functions ψ_k we consider here are convex, symmetric, nonnegative and differentiable [3]. Our goal is to produce a penalized-likelihood estimate

$$\hat{\mu} = \arg \max_{\mu \geq \underline{0}} \Phi(\mu, Y). \quad (5)$$

III. ALGORITHM

We focus on the unregularized maximum-likelihood problem; the regularized approach easily follows from [3]. Since maximizing the log-likelihood directly is difficult to do, we apply the principle of optimization transfer [6], [15] and define a “surrogate function” $Q(\mu; \mu^n)$ that is easier to maximize. Since this surrogate function depends on the previous estimate μ^n at the n th iteration, the algorithm consists of repeatedly maximizing the surrogate function, i.e.,

$$\mu^{n+1} = \arg \max_{\mu \geq \underline{0}} Q(\mu; \mu^n). \quad (6)$$

Note that the maximization is constrained to enforce the non-negativity constraint. The key algorithm design requirement is to choose Q functions that satisfy the following conditions:

$$Q(\mu^n; \mu^n) = L(\mu^n, Y), \quad \forall \mu^n \geq \underline{0} \quad (7)$$

$$\left. \frac{\partial Q}{\partial \mu_j}(\mu; \mu^n) \right|_{\mu=\mu^n} = \left. \frac{\partial L}{\partial \mu_j}(\mu) \right|_{\mu=\mu^n}, \quad \forall j = 1, \dots, p \quad (8)$$

$$Q(\mu; \mu^n) \leq L(\mu, Y), \quad \forall \mu \geq \underline{0}. \quad (9)$$

These conditions ensure that the proposed iteration monotonically increases the likelihood.

TABLE I
ALGORITHM OUTLINE

Initialize: $\hat{\mu} = \text{FBP}\{\log((\sum_{m=1}^M b_{im})/(y_i - r_i))\}_{i=1}^N$ and $\hat{l}_{im} = \sum_{j=1}^p a_{ij}^m \hat{\mu}_j$, $i = 1, \dots, N$
for each iteration $n = 1, \dots, \text{niter}$

$$\begin{aligned} u_{im} &= b_{im} \exp(-[A\hat{\mu}]_i) + r_i/M \\ \bar{y}_i &= \sum_{m=1}^M b_{im} \exp(-[A\hat{\mu}]_i) + r_i \\ b_{im}^n &= \frac{\bar{y}_i}{u_{im}} b_{im}, r_{im}^n = \frac{\bar{y}_i}{u_{im}} r_i \end{aligned}$$

Compute c_{im} according to (30)

$$\dot{q}_{im} = \dot{g}_{im} = \left(\frac{y_i}{b_{im}^n e^{-\hat{l}_{im}} + r_{im}^n} - 1 \right) b_{im}^n e^{-\hat{l}_{im}}, \quad \text{for } i = 1, \dots, N$$

for $j = 1, \dots, p$

$$\begin{aligned} \dot{Q}_j &= \sum_{m=1}^M \sum_{i=1}^N \frac{u_{im}}{\bar{y}_i} a_{ij}^m \dot{q}_{im} \\ d_j &= \sum_{m=1}^M \sum_{i=1}^N \frac{u_{im}}{\bar{y}_i} (a_{ij}^m)^2 c_{im} \\ \mu_j^{\text{old}} &= \mu_j \\ \hat{\mu}_j &:= \left[\hat{\mu}_j + \frac{\dot{Q}_j - \beta \sum_{k=1}^K c_{kj} \psi_k([\mathbf{C}\hat{\mu}]_k)}{d_j + \beta \sum_{k=1}^K c_{kj}^2 \psi_k([\mathbf{C}\hat{\mu}]_k)} \right]_+ \\ \dot{q}_{im} &:= \dot{q}_{im} - a_{ij}^m c_{im} (\hat{\mu}_j - \mu_j^{\text{old}}) \end{aligned}$$

end

$$\hat{l}_{im} := \hat{l}_{im} + \frac{\dot{q}_{im} - \dot{g}_{im}}{c_{im}}, \quad \text{for } i = 1, \dots, N$$

end

A difficulty in maximizing L is the sum over m within the logarithm in (2). To move the summation outside of the logarithm, we first adapt De Pierro's multiplicative convexity trick [16]. Because h_i is concave

$$h_i \left(\sum_{m=1}^M u_{im}(\mu) \right) = h_i \left(\sum_{m=1}^M \frac{u_{im}^n}{\bar{y}_i^n} \frac{u_{im}(\mu)}{u_{im}^n} \bar{y}_i^n \right) \quad (10)$$

$$\geq \sum_{m=1}^M \frac{u_{im}^n}{\bar{y}_i^n} h_i \left(\frac{u_{im}(\mu)}{u_{im}^n} \bar{y}_i^n \right) \quad (11)$$

where $u_{im}^n \triangleq u_{im}(\mu^n)$, and $\bar{y}_i^n \triangleq \bar{y}_i(\mu^n) = \sum_{m=1}^M u_{im}^n$. This inequality leads to our first surrogate function

$$Q_1(\mu; \mu^n) \triangleq \sum_{i=1}^N \sum_{m=1}^M \frac{u_{im}^n}{\bar{y}_i^n} h_i \left(\frac{u_{im}(\mu)}{u_{im}^n} \bar{y}_i^n \right) \quad (12)$$

$$= \sum_{i=1}^N \sum_{m=1}^M \frac{u_{im}^n}{\bar{y}_i^n} g_{im}^n([A^m \mu]_i) \quad (13)$$

where

$$b_{im}^n \triangleq \frac{\bar{y}_i^n}{u_{im}^n} b_{im}$$

$$r_{im}^n \triangleq \frac{\bar{y}_i^n}{u_{im}^n} r_i$$

$$g_{im}^n(l) \triangleq y_i \log(b_{im}^n e^{-l} + r_{im}^n) - (b_{im}^n e^{-l} + r_{im}^n).$$

The surrogate function Q_1 remains too difficult to maximize directly because the argument of each h_i still depends on u_{im} , which has a complicated exponential form. However, it follows easily from the results in [2] and [3] that the following paraboloidal function is a valid surrogate for Q_1 :

$$Q_2(\mu; \mu^n) \triangleq \sum_{i=1}^N \sum_{m=1}^M \frac{u_{im}^n}{\bar{y}_i^n} q_{im}^n([\mathbf{A}^m \mu]_i) \quad (14)$$

where

$$q_{im}^n(l) \triangleq g_{im}^n(l_{im}^n) + \dot{g}_{im}^n(l_{im}^n) \times (l - l_{im}^n) - \frac{1}{2} c_{im}^n (l - l_{im}^n)^2 \quad (15)$$

and

$$\dot{g}_{im}^n(l) \triangleq \frac{d}{dl} g_{im}^n(l)$$

$$\ddot{g}_{im}^n(l) \triangleq \frac{d^2}{dl^2} g_{im}^n(l)$$

$$l_{im}^n = [A^m \mu^n]_i = \sum_{j=1}^p \mathbf{A}_{ij}^m \mu_j^n.$$

To ensure (9), we must choose the curvatures $\{c_{im}^n\}$ appropriately [2], [3]. As discussed in [3], for the fastest convergence rate, we would like to choose the curvatures as small as possible, subject to the constraint that the surrogate function q_{im}^n lies below the functions g_{im}^n ; see Appendix A for the expression for the optimum curvature c_{im}^n . We can also separate Q_2 to obtain another surrogate function Q_3 , to which we could apply the ordered subsets idea ([17], p. 107).

To obtain a monotonic algorithm that converges relatively quickly, we apply coordinate ascent to the surrogate Q_2 defined in (14), i.e., we sequentially update one pixel at a time while holding all other pixels fixed. The derivative of the likelihood surrogate parabola at $\hat{\mu}_j$ is

$$\begin{aligned} \dot{Q}_j^n(\hat{\mu}) &\triangleq \left. \frac{\partial}{\partial \mu_j} Q_2^n(\mu; \mu^n) \right|_{\mu=\hat{\mu}} \\ &= \sum_{i=1}^N \sum_{m=1}^M \frac{u_{im}^n}{\bar{y}_i^n} a_{ij}^m \left[\dot{g}_{im}^n(l_{im}^n) - c_{im}^n (\hat{l}_{im} - l_{im}^n) \right] \end{aligned} \quad (16)$$

where $\hat{l}_{im} \triangleq [\mathbf{A}^m \hat{\mu}]_i$ and

$$\dot{g}_{im}^n(l) = \left(\frac{y_i}{b_{im}^n e^{-l} + r_{im}^n} - 1 \right) b_{im}^n e^{-l}.$$

The curvature of the parabola Q_2 is

$$\begin{aligned} d_j^n &\triangleq \frac{\partial^2}{\partial \mu_j^2} Q_2^n(\mu; \mu^n) \\ &= \sum_{i=1}^N \sum_{m=1}^M \frac{u_{im}^n}{\bar{y}_i^n} (a_{ij}^m)^2 c_{im}^n. \end{aligned} \quad (17)$$

Extension to the penalized likelihood case is straightforward following the methods in [2] and [3], so we omit the details and only comment on notations: $\hat{R}_j(\mu_j)$ denotes the penalty surrogate parabola for μ_j , and \hat{p}_j denotes its curvature. The maximization step of the coordinate ascent for pixel j is

$$\begin{aligned} \hat{\mu}_j^{\text{new}} &= \arg \max_{\mu_j \geq 0} Q([\hat{\mu}_1, \dots, \hat{\mu}_{j-1}, \mu_j, \hat{\mu}_{j+1}, \dots, \hat{\mu}_p]; \mu^n) \\ &- \beta \hat{R}_j(\mu_j) = \left[\hat{\mu}_j + \frac{\dot{Q}_j^n(\hat{\mu}) - \beta \dot{R}_j(\hat{\mu})}{d_j^n + \beta \hat{p}_j} \right]_+. \end{aligned} \quad (18)$$

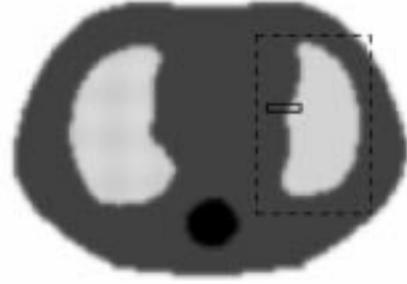


Fig. 1. Digital phantom used in our simulations and the ROI used for collimation angle optimization.

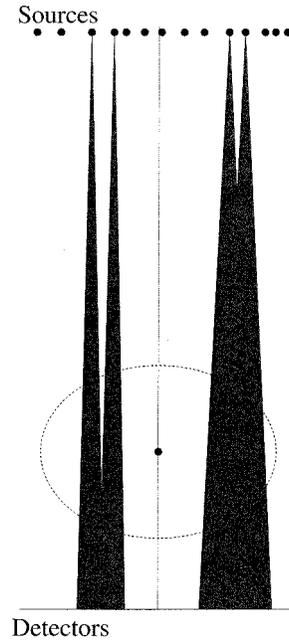


Fig. 2. Scaled illustration of the system setup; the two fan-beams on the left have collimation angle 2.6° ; the two fan-beams on the right side have collimation angle 5.6° .

Because of our construction based on surrogate functions that satisfy (9), this update is guaranteed to monotonically increase the value of Φ . One iteration is finished when all pixels are updated via (18) in a sequential order. We update the paraboloidal surrogate function after one full sweep of coordinate ascent (CA) through all pixels, although one could also perform more than one CA sweep per surrogate. An outline of this algorithm is given in Table I, where ψ_k denotes the potential function used in the penalty acting as a norm on the “soft constraints” $\mathbf{C}\mu \approx 0$, K is the number of such constraints, and $\hat{\psi}_k$ denotes the surrogate function used for ψ_k ; see [3] for details.

IV. SIMULATION RESULTS

We compared the proposed reconstruction algorithm with the “conventional” reconstruction algorithms (statistical and FBP) that treat the transmission measurements simply as ideal normalized parallel “strip-integrals.” The system geometry corresponded fairly closely to the Siemens Profile system (Hoffman Estates, IL) [18]. The sources for the simulated system consisted of a multiple line source array with 14

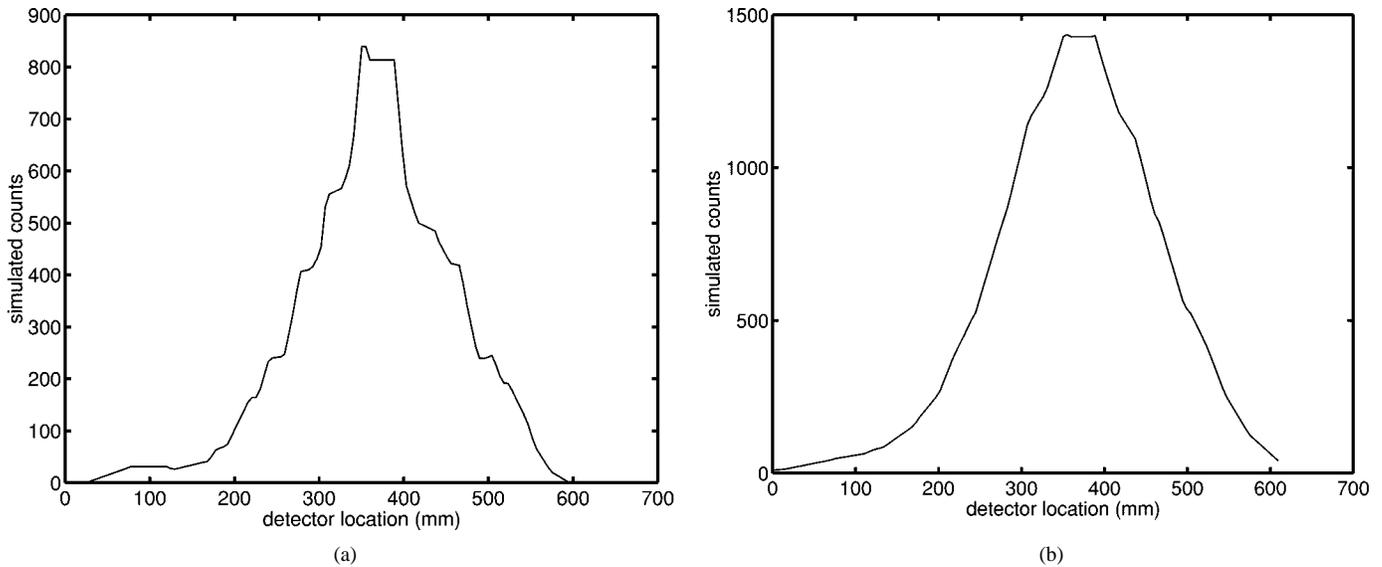


Fig. 3. Distribution of blank counts (a) collimation angle 2.6° (b) collimation angle 5.6° .

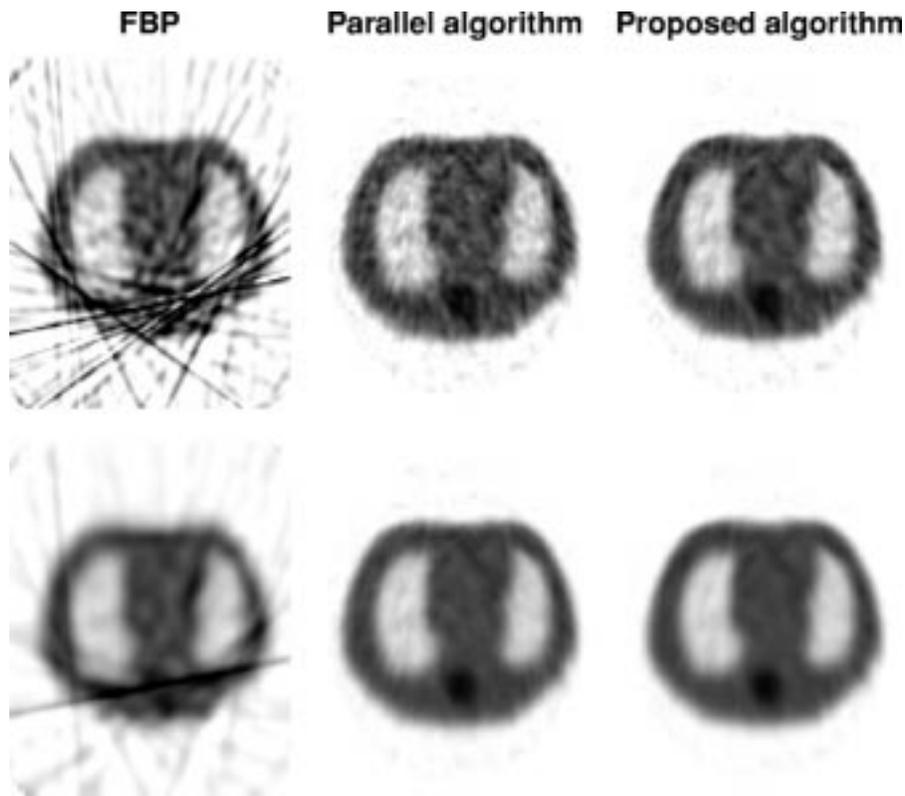


Fig. 4. New sources; collimation angle: 2.6° ; 785 000 counts; top row: resolution 4.7 pixels; bottom row, resolution 6.8 pixels.

sources, unequally spaced, located on a line parallel to the detector and 110 cm away from the detector plane.⁴ The detector plane was located 22 cm away from the center of rotation. We simulated a range of source collimation angles, from 1.6° (with almost no overlap in the transmission beams) to 6.6° , and disregarded detector collimation. The image consisted 128×128

⁴In addition to transaxial overlap of transmission source radiation, there are some designs for transmission sources in which the overlap is primarily in the axial direction. This algorithm can be easily modified to suit such designs. However, storage and computational issues might make this algorithm prohibitively expensive to perform in reality.

pixels of size $3.56 \times 3.56 \text{ mm}^2$. The sinogram size was 128×60 with detector bins of width 4.8 mm (i.e., the simulated detector response was rectangular with width 4.8 mm). We performed the simulation for two levels of transmitted counts, one corresponding to a system whose center rods have just been replaced (the new source case), and the other with sources that have all decayed by one half-life (the old source case). At 1.6° , we simulated 321 000 transmitted counts for the new source case (160 000 counts for the old source case), and 263 000 background counts (on average). As the source collimation angle increases, the number of transmitted counts increases,

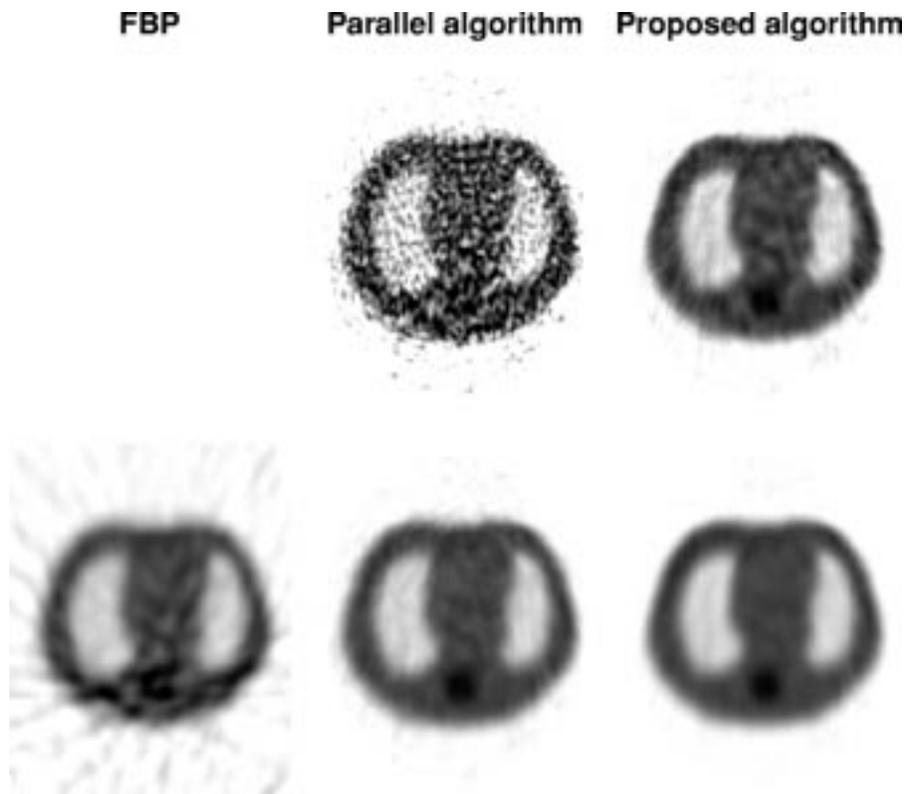


Fig. 5. New sources; collimation angle: 3.6° ; 994 000 counts; top row: resolution 4.7 pixels; bottom row, resolution 6.8 pixels. A resolution of 4.7 pixels was not achievable with FBP in this case.

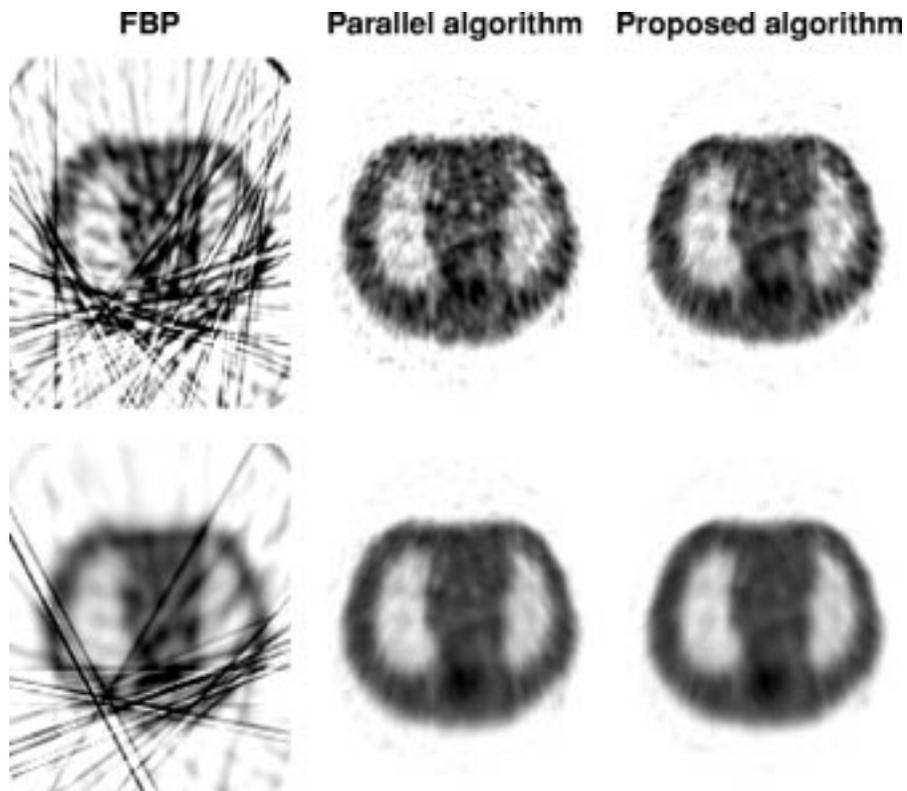


Fig. 6. Old sources; collimation angle: 2.6° ; 392 000 counts; top row: resolution 4.7 pixels; bottom row, resolution 6.8 pixels.

naturally, but the number of background counts remains the same; at 2.6° , there are 523 000 transmitted counts for the

new source case (261 000 counts for the old source case); and at 6.6° , there are 1 396 000 transmitted counts for the new

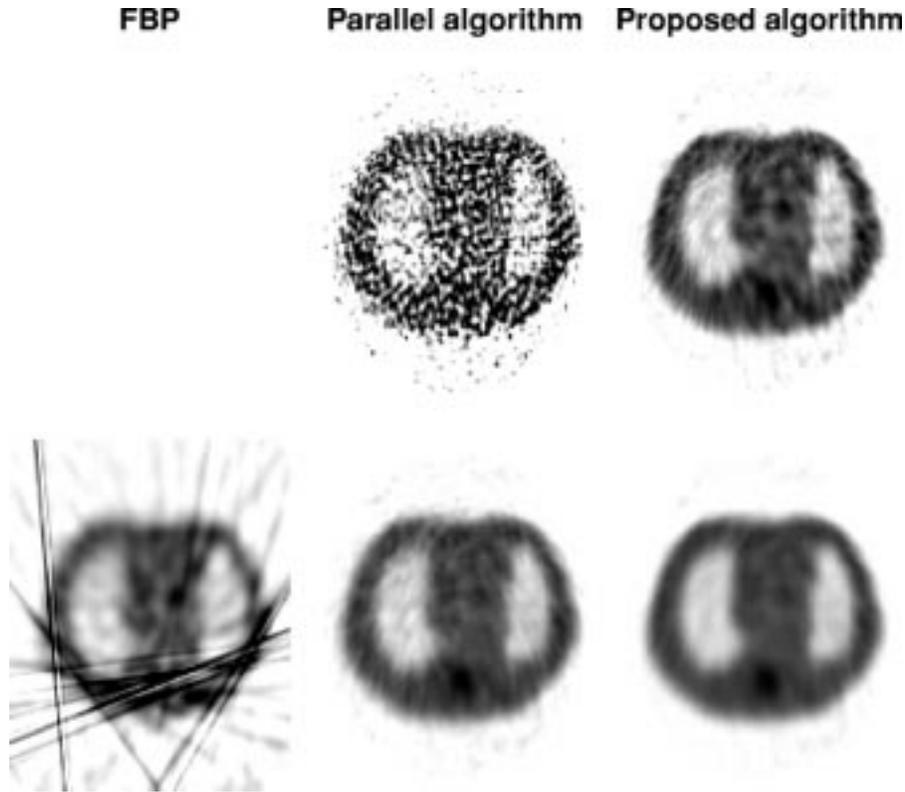


Fig. 7. Old sources; collimation angle: 3.6° ; 497 000 counts; top row: resolution 4.7 pixels; bottom row, resolution 6.8 pixels. A resolution of 4.7 pixels was not achievable with FBP in this case.

source case. For simplicity, we used a space-invariant quadratic penalty over first-order neighbors throughout our simulations. The phantom used in our simulations, the ROI used for the evaluation of variance in Section V (outlined by solid lines), and the large region used for the evaluation of spatial resolution (outlined by dashed lines) are shown in Fig. 1. Fig. 2 illustrates the system setup. The radial distribution of blank counts (at any projection angle) is shown in Fig. 3.

Figs. 4 and 5 show reconstructions of noisy data using FBP, the parallel algorithm⁵, and the proposed algorithm, with new sources. Fig. 4 shows reconstructions with 2.6° source collimation and Fig. 5 shows reconstructions with 3.6° source collimation. Figs. 6 and 7 show the same reconstructions with old sources. Since FBP and statistical algorithms have different methods for controlling the resolution/variance tradeoff, we must display images with similar resolutions (or similar variances) to have a meaningful comparison between different reconstruction algorithms. We chose β so that the spatial resolution of the images in Figs. 4–7 is 4.7 pixels on the top row and 6.8 pixels on the bottom row. We use the following simple method to determine the resolution of a particular (noiseless) reconstruction. Given the ideal image μ^{true} and the reconstructed image $\check{\mu}$ using ideal data

$$\check{\mu} = \arg \max_{\mu \geq 0} \Phi(\mu, \bar{y}_i(\mu^{\text{true}})) \quad (19)$$

⁵We refer to the penalized-likelihood reconstruction assuming ideal normalized parallel “strip-integrals” (and ignoring beam overlap) as the parallel algorithm.

we define the resolution of $\check{\mu}$ to be

$$\arg \min_{\sigma} \sum_{j \in \mathcal{R}} |[G_{\sigma} \mu^{\text{true}}]_j - \check{\mu}_j|^2 \quad (20)$$

where G_{σ} represents a Gaussian smoothing filter with FWHM σ , and the \mathcal{R} is a large region encompassing both the right lung and our ROI as illustrated in Fig. 1.

We observe from Figs. 4–7 that the proposed algorithm consistently produces less noisy reconstructions than both the parallel algorithm and FBP (this conclusion from anecdotal evidence will be confirmed in Section V). The noise reduction is especially significant when the collimation angle is large (i.e., 3.6°) and the desired spatial resolution is high (i.e., 4.7 pixels). In such cases, FBP simply cannot produce a reconstructed image of the desired resolution even with an unapodized ramp filter. Since the parallel algorithm is based on an incorrect system and statistical model, one expects artifacts due to model mismatch. The absence of apparent artifacts in Figs. 4–7 is due to regularization and noise. Fig. 8 shows the reconstructed images from noiseless data (4.6° collimation angle) using the parallel and proposed algorithms with almost no regularization ($\beta = 2^{-10}$). The reconstructed image from noiseless data using the parallel algorithm shows severe artifacts resulting from model mismatch, which are absent in the reconstructed image from noiseless data using the proposed algorithm [Fig. 8(c)].

As the collimation angle increases, the artifacts generated by the parallel algorithm worsen. In fact, even the noiseless reconstruction has a spatial resolution of about 5.6 pixels when the

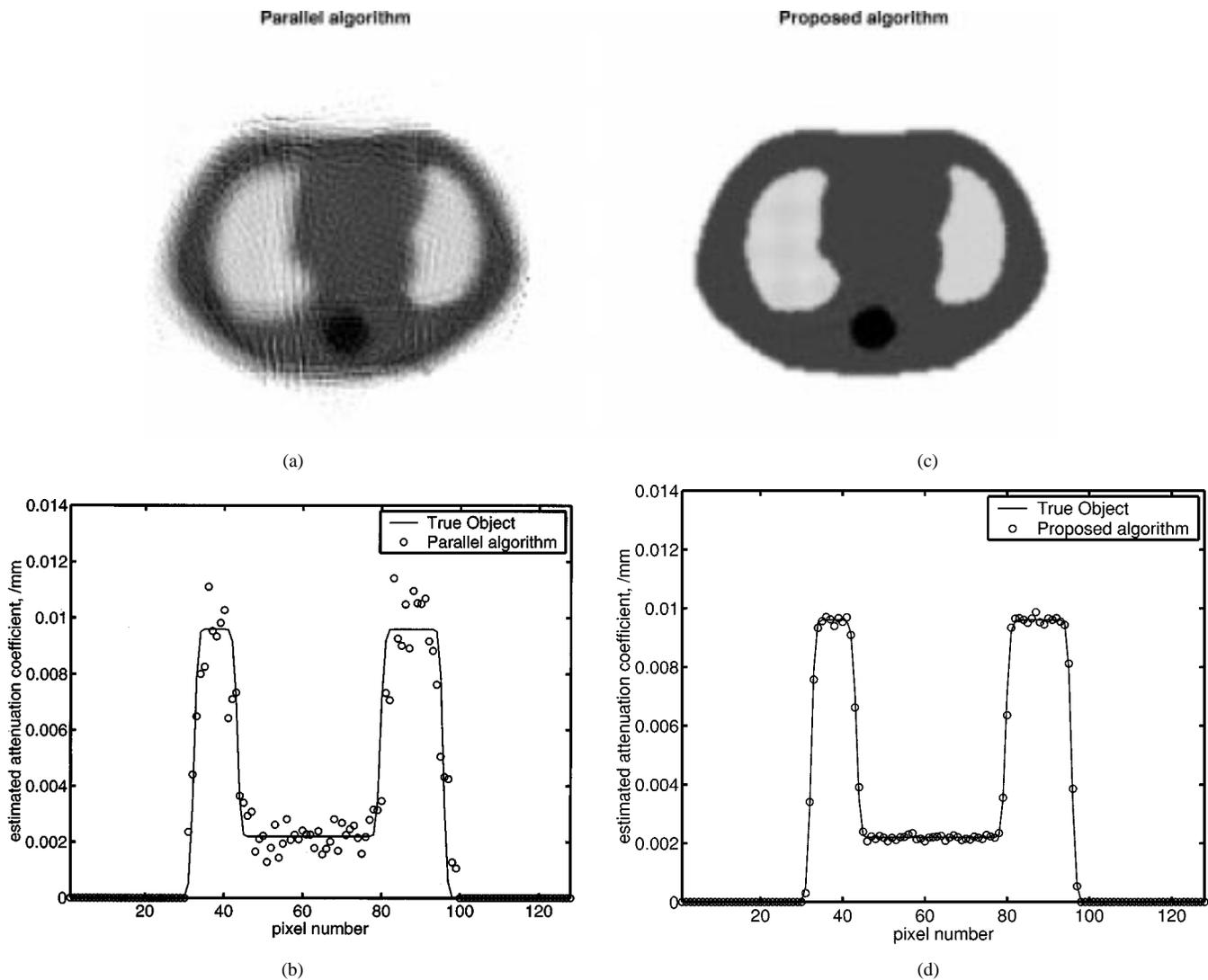


Fig. 8. Reconstruction using the parallel and proposed algorithms with almost no regularization; collimation angle 4.6° (a) (b) Parallel algorithm (c) (d) Proposed algorithm. The profile displayed in (b) and (d) is of vertical column number 100, through the right lung.

collimation angle is 4.6° [Fig. 8(a)],⁶ and at 6.6° , the noiseless reconstruction has a spatial resolution of 7.7 pixels. Thus for system setups with significantly overlapping transmission beams, the parallel algorithm simply cannot produce a high-resolution reconstruction, no matter how many counts one collects.

The proposed algorithm is more time-consuming and uses more memory than conventional statistical algorithms. For our simulations, we used 14 system matrices (one for each source) with appropriate collimation angles. The system matrices with collimation angle 2.6° occupy 20 MBytes of disk space, and the system matrices with collimation angle 7° occupy 50 MBytes of disk space, compared to 8 MBytes occupied a single system matrix used by the parallel algorithm. It also takes longer to project or backproject an image in the proposed algorithm than the parallel algorithm; in fact, for collimation angle 2.6° , each iteration of the proposed algorithm takes about 1.9 s, compared to about 0.8 s needed for the parallel algorithm, on a Sun Ultra2 workstation. Furthermore, because of the overlap between trans-

mission beams, the convergence rate of the proposed algorithm is also slower than the parallel algorithm; as the overlap between transmission beams increases, i.e., the collimation angle gets larger, the curvatures c_{im} , as given in (30) at the bottom of the page, get larger, hence, the convergence rate becomes even slower. In fact, if one goes from a collimation angle of 1.6° to 5.6° , the average curvature of the surrogate parabolas increases by a factor of three. With regard to the number of iterations necessary for convergence, i.e., the smallest n such that $\Phi(\mu^n) - \Phi(\mu^0) > 0.999[\Phi(\mu^*) - \Phi(\mu^0)]$, where $\Phi(\mu^*)$ is the largest objective value obtained in 40 iterations, the parallel algorithm required 22 iterations and the proposed algorithm required 30 iterations for collimation angle 2.6° and a desired spatial resolution of 4.7 pixels. Hence, the total amount time required by the proposed algorithm would be 3.3 times that of the parallel algorithm for collimation angle 2.6° and a desired spatial resolution of 4.7 pixels. On a typical three-dimensional data set, with 40 slices and running 20 iterations for each slice, it would take approximately 25 min to perform a complete reconstruction on a Sun Ultra2 workstation.

⁶The noiseless reconstruction [Fig. 8(c)] using the proposed algorithm has a spatial resolution of 1.4 pixels.

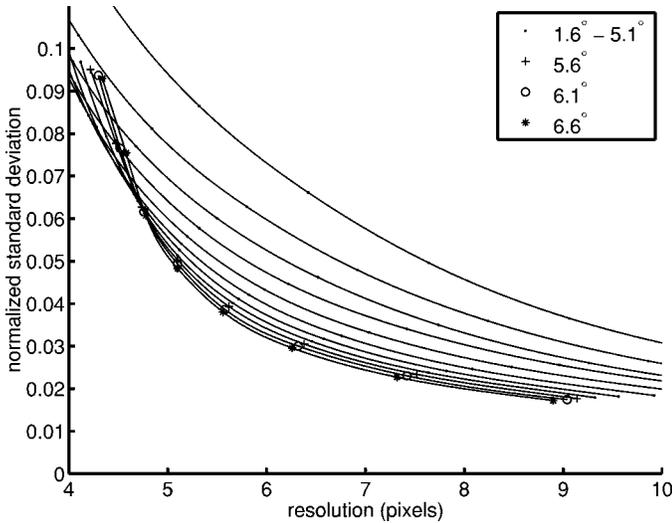


Fig. 9. High-count case: ROI resolution/variance curves for the proposed algorithm.

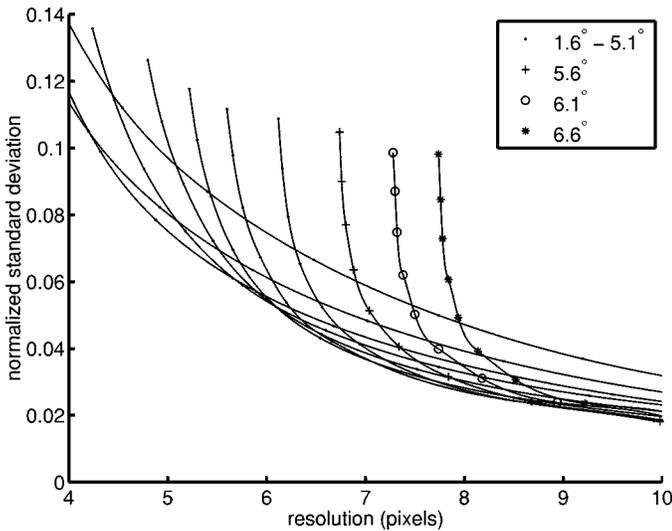


Fig. 10. High-count case: ROI resolution/variance curves for the parallel algorithm.

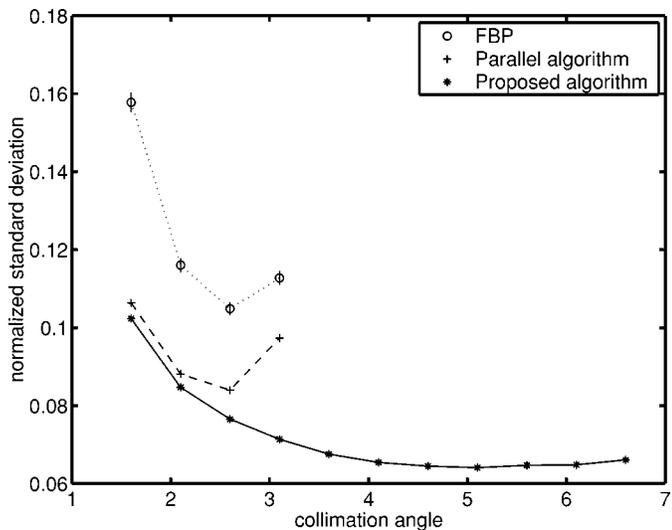


Fig. 11. High-count case: ROI variances at desired spatial resolution 4.7 pixels.

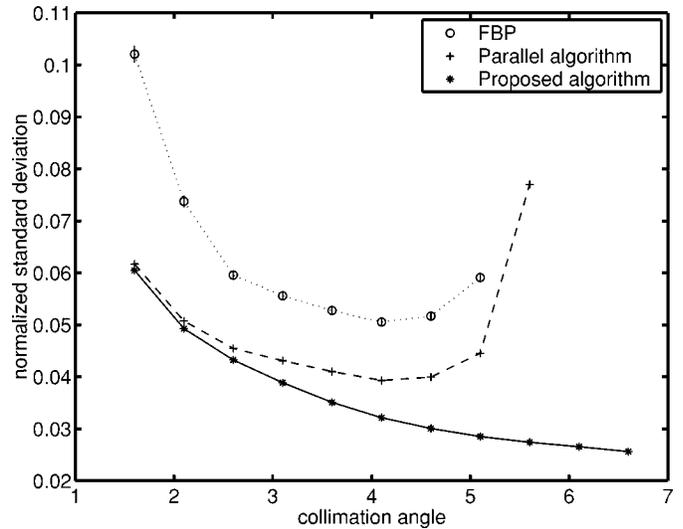


Fig. 12. High-count case: ROI variances at desired spatial resolution 6.8 pixels.

V. COLLIMATION ANGLE OPTIMIZATION

We performed a preliminary study of the optimal source collimation angle given that the system configuration, source strength, and background counts remain constant. As source collimators open up, i.e., the collimation angle increases, there would be more counts, but there would also be more overlap of transmission beams. Hence, initially, when the transmission beams widen from no overlap to some overlap, we expect better resolution/variance tradeoffs; however, as the transmission beams open up more and more, we expect less and less improvements, and eventually worse resolution/variance tradeoff since eventually each detected photon hitting will yield very little information about where it originated. We want to obtain the collimation angle that minimizes a region of interest (ROI) variance for a fixed spatial resolution. We investigated the proposed algorithm, the parallel algorithm (i.e., the conventional statistical algorithm), and FBP (derived assuming ideal parallel “line-integrals”). For the statistical algorithms, instead of performing numerical simulations, we used the approach outlined in [19] to compute the variance of the ROI. For FBP, since numerical simulations are relatively inexpensive, we performed 2000 realizations for each data point.

We will now derive an algorithm for approximately calculating the variance of the ROI of the proposed estimator. The approximate covariance for an implicitly defined biased estimator is [19]

$$\text{Cov}(\hat{\mu}) \approx [-\nabla^{20}\Phi(\check{\mu}, \bar{y})]^{-1} \nabla^{11}\Phi(\check{\mu}, \bar{y}) \text{Cov}\{Y\} \times [\nabla^{11}\Phi(\check{\mu}, \bar{y})]' [-\nabla^{20}\Phi(\check{\mu}, \bar{y})]^{-1} \quad (21)$$

where $\check{\mu}$ is defined in (19), the (j, k) th element of the operator ∇^{20} is $\partial^2 / \partial \check{\mu}_j \partial \check{\mu}_k$, and the (j, k) th element of the operator ∇^{11} is $\partial^2 / \partial \check{\mu}_j \partial \bar{y}_k$. We first derive $\nabla^{20}\Phi(\mu, Y)$ and $\nabla^{11}\Phi(\mu, Y)$, which are needed for computing the variance of a ROI for the proposed algorithm. Defining

$$p_{im}(\mu) \triangleq b_{im} e^{-\sum_j a_{ij}^m \mu_j} \quad (22)$$

$$\bar{y}_i(\mu) = \sum_{m=1}^M p_{im}(\mu) + r_i \quad (23)$$

then

$$\frac{\partial p_{im}(\mu)}{\partial \mu_j} = -a_{ij}^m p_{im}(\mu) = \frac{\partial \bar{y}_i(\mu)}{\partial \mu_j}. \quad (24)$$

Since

$$L(\mu, Y) = \sum_{i=1}^N y_i \log \bar{y}_i(\mu) - \bar{y}_i(\mu) \quad (25)$$

we have

$$\begin{aligned} \frac{\partial L(\mu, Y)}{\partial \mu_j} &= \sum_{i=1}^N \frac{y_i}{\bar{y}_i(\mu)} \frac{\partial \bar{y}_i(\mu)}{\partial \mu_j} - \frac{\partial \bar{y}_i(\mu)}{\partial \mu_j} \\ &= \sum_{i=1}^N \left(1 - \frac{y_i}{\bar{y}_i(\mu)}\right) \sum_{m=1}^M a_{ij}^m p_{im}(\mu) \end{aligned} \quad (26)$$

and, applying the chain rule, we have, after simplification

$$\begin{aligned} \nabla^{11} \Phi(\mu, Y) &= -\mathbf{C}(\mu)' \text{diag}\{1/\bar{y}_i(\mu)\} \\ -\nabla^{20} \Phi(\mu, Y) &= \sum_{m=1}^M (\mathbf{A}^m)' \\ &\quad \text{diag}\{(\bar{y}_i(\mu) - r_i)(1 - y_i/\bar{y}_i(\mu))\} \mathbf{A}^m \\ &\quad + \mathbf{C}(\mu)' \text{diag}\{y_i/\bar{y}_i^2(\mu)\} \mathbf{C}(\mu) + \beta \mathbf{R}(\mu) \end{aligned} \quad (28)$$

where $\mathbf{R}(\mu) = \nabla^2 R(\mu)$, $\mathbf{C}(\mu) \triangleq \{c_{ij}(\mu)\}$, and $c_{ij}(\mu) \triangleq \sum_{m=1}^M a_{ij}^m p_{im}(\mu)$, i.e., $\mathbf{C}(\mu) = \sum_{m=1}^M \text{diag}\{p_{im}(\mu)\} \mathbf{A}^m$. (For a detailed derivation of the penalty part, see [19].) We use the following recipe to compute the approximate estimator (using the proposed algorithm) variance of a ROI:

- 1) compute $\check{\mu}$ by applying the proposed algorithm to noise-free data $\{\bar{y}_i(\mu^{\text{true}})\}$;
- 2) forward project $\check{\mu}$ to compute $\bar{y}_i(\check{\mu}) = \sum_{m=1}^M p_{im}(\check{\mu}) + r_i$;
- 3) use the conjugate gradient method [20] to solve, for u^{ROI} , the system of equations $[-\nabla^{20} \Phi(\check{\mu}, \bar{y})] u^{\text{ROI}} = e^{\text{ROI}}$, where e^{ROI} is a vector with entries $1/n_{\text{ROI}}$ for pixels inside the region and zeros for pixels outside; n_{ROI} denotes the number of pixels in the ROI;
- 4) Compute $(u^{\text{ROI}})' \nabla^{11} \Phi(\check{\mu}, \bar{y}) \text{Cov}\{Y\} [\nabla^{11} \Phi(\check{\mu}, \bar{y})]' u^{\text{ROI}}$ by first forward projecting u^{ROI} to compute $v = \text{diag}\{1/\bar{y}_i(\check{\mu})\} \mathbf{C}(\check{\mu}) u^{\text{ROI}}$, and then summing

$$\text{Var}(\hat{\mu}^{\text{ROI}}) \approx \sum_{i=1}^N v_i^2 y_i(\mu^{\text{true}}). \quad (29)$$

For the parallel algorithm, the computation of the approximate variance has been outlined in [19], except that we use $\text{Cov}(Y) = \text{diag}\{\bar{y}_i(\mu^{\text{true}})\}$ in [19, (23)], where \bar{y}_i is based on the overlapping beam model (23) rather than the parallel strip-integral model of [19].

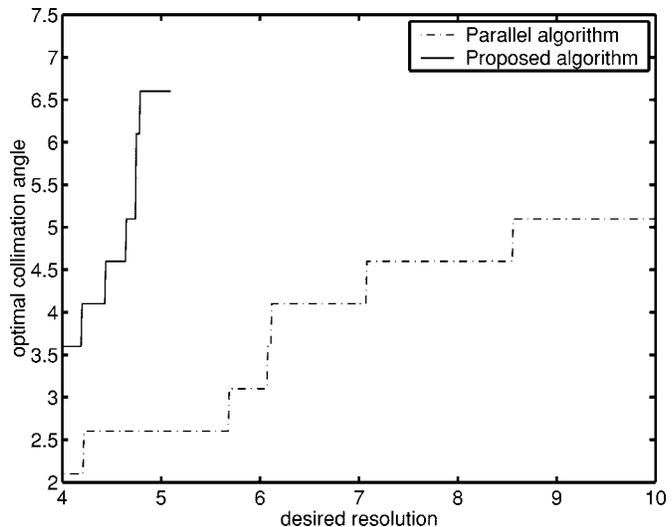


Fig. 13. High count case: A comparison of optimal collimation angles, at different resolutions for the proposed and parallel algorithms.

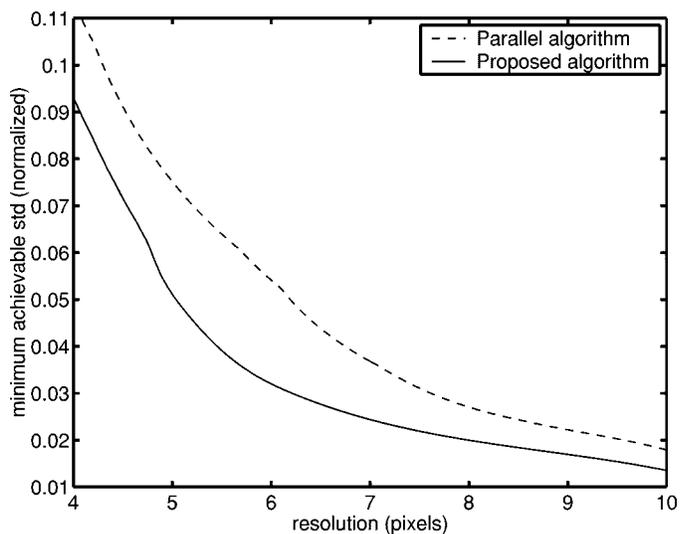


Fig. 14. High-count case: A comparison of minimum achievable normalized standard deviation, at different resolutions for the proposed and parallel algorithms.

For simplicity, we consider a first order quadratic penalty for the results below. We analyzed the resolution/variance tradeoffs for the ROI illustrated in Fig. 1: it is a 2×9 region that goes across the boundary of the soft tissue, the lung, and the heart. We performed the analysis at two levels of transmitted counts, one with new sources (the high count case) and the other with old sources (the low count case), i.e., the same as was done for Section IV. The background count level for both cases remains the same throughout. We only discuss the high count case; the low count case gives similar results. Figs. 9 and 10 show the resolution/variance curves of the proposed algorithm and the parallel algorithm, respectively, for various collimation angles ranging from 1.6° to 6.6° . Based on these two figures, we calculated (using cubic spline interpolation when needed) the variances of the ROI versus collimation angle at two fixed resolutions: 4.7 pixels and 6.8 pixels, as shown in Figs. 11 and 12. The proposed algorithm outperforms the parallel algorithm which in

$$c_{im}^n = \begin{cases} \left[-2 \frac{g_{im}^n(0) - g_{im}^n(l_{im}^n) + \dot{g}_{im}^n(l_{im}^n)(l_{im}^n)}{(l_{im}^n)^2} \right]_+, & l_{im}^n > 0 \\ [-\dot{g}_{im}^n(0)]_+, & l_{im}^n = 0 \end{cases} \quad (30)$$

$$= \begin{cases} \left[\left(2 / (l_{im}^n)^2 \right) \left\{ b_{im}^n (1 - e^{-l_{im}^n}) - y_i \log \frac{b_{im}^n + r_{im}^n}{\bar{y}_{im}^n} + b_{im}^n l_{im}^n e^{-l_{im}^n} \left(\frac{y_i}{\bar{y}_{im}^n} - 1 \right) \right\} \right]_+, & l_{im}^n > 0 \\ \left[1 - \frac{y_i r_{im}^n}{(b_{im}^n + r_{im}^n)^2} \right]_+, & l_{im}^n = 0 \end{cases} \quad (31)$$

turn outperforms FBP at both resolutions, and at all collimation angles. The performance gain of the proposed algorithm over the parallel algorithm is more impressive at larger collimation angles, e.g., more than 2.5° . Using the proposed algorithm, it seems that the optimal collimation angle for a resolution of 4.7 pixels is around 5° , and $>7^\circ$ for a resolution of 6.8 pixels, both much larger than typical collimation angles found on SPECT cameras. However, if the suboptimal parallel algorithm were used for reconstruction, then the optimal collimation angle would be 2.6° for a resolution of 4.7 pixels, and 4.1° for a resolution of 6.8 pixels. Even though the optimal collimation angle for the parallel algorithm is closer to what is typically found on SPECT cameras than the proposed algorithm, the proposed algorithm outperforms the parallel algorithm at all collimation angles. Furthermore, for collimation angles larger than 3.5° , a resolution requirement of 4.7 pixels would probably mean very high variances in the parallel reconstruction. Fig. 13 shows the optimal collimation angle for the proposed and parallel algorithms at different desired spatial resolutions. Naturally, as the desired spatial resolution improves, the optimal collimation angle decreases. Fig. 14 shows the minimum achievable normalized standard deviation for the the proposed and parallel algorithms at different desired spatial resolutions. As expected, the proposed algorithm consistently outperforms the parallel algorithm, resulting in reduction in the standard deviation by as much as 40%.

VI. DISCUSSION AND CONCLUSION

We have presented a new algorithm for statistical image reconstruction of attenuation maps that explicitly accounts for overlapping beams in transmission scans; an example of such a system can be found in [21].⁷ The algorithm is guaranteed to monotonically increase the objective function at each iteration, and consistently achieves better variance/resolution tradeoffs than the two “conventional” image reconstruction algorithms tested in this paper, both statistical (the parallel algorithm) and nonstatistical (FBP).

From our preliminary study on the optimal collimation angle, we find that it is desirable to open up the source collimators and allow beam overlap—provided the overlap is modeled appropriately in the reconstruction algorithm. (This conclusion applies to our ideal simulation settings where system geometry is known perfectly. Robustness of the proposed algorithm to model

mismatch needs investigation, since in practical clinical situations, the a_{ij}^m s and b_{im} s are never known exactly.) However, detector collimation currently limits usable source collimation angles to $\geq 2.6^\circ$, so alternatives to widening source collimation are needed to improve the counting statistics on this system.

A possible alternative to the coordinate ascent algorithm presented in this paper is the ordered subsets algorithm. Although the ordered subsets algorithm is not guaranteed to converge to a local maximum of the objective function, it often produces reasonable images quickly and is more parallelizable than CA, so it may be more practical for clinical use.

APPENDIX

For completeness, we include the following formula for the optimum curvature, i.e., $c_{im}^n = \min\{c \geq 0 : g_{im}^n(l) \leq q_{im}^n(l) \forall l \geq 0\}$ derived in [3]; see (30) and (31) at the top of the page) where

$$\bar{y}_{im}^n \triangleq b_{im}^n e^{-l_{im}^n} + r_{im}^n.$$

Other curvature choices that lead to even faster convergence (but do not guarantee monotonicity) can be found in [3].

ACKNOWLEDGMENT

The authors would like to thank P. Hawman for assistance with the Siemens Profile geometrical specifications.

REFERENCES

- [1] A. Celler, A. Sitek, and R. Harrop, “Reconstruction of multiple line source attenuation maps,” *IEEE Trans. Nucl. Sci.*, vol. 44, pp. 1503–1508, Aug. 1997.
- [2] H. Erdoğan and J. A. Fessler, “Accelerated monotonic algorithms for transmission tomography,” *Proc. IEEE Int. Conf. Image Processing*, vol. 2, pp. 680–684, 1998.
- [3] —, “Monotonic algorithms for transmission tomography,” *IEEE Trans. Med. Imag.*, vol. 18, pp. 801–814, Sept. 1999.
- [4] K. Lange and R. Carson, “EM reconstruction algorithms for emission and transmission tomography,” *J. Comput. Assist. Tomogr.*, vol. 8, no. 2, pp. 306–316, Apr. 1984.
- [5] K. Sauer and C. Bouman, “A local update strategy for iterative reconstruction from projections,” *IEEE Trans. Signal Processing*, vol. 41, pp. 534–548, Feb. 1993.
- [6] K. Lange and J. A. Fessler, “Globally convergent algorithms for maximum a posteriori transmission tomography,” *IEEE Trans. Image Processing*, vol. 4, pp. 1430–1438, Oct. 1995.
- [7] E. U. Mumcuoglu, R. Leahy, S. R. Cherry, and Z. Zhou, “Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images,” *IEEE Trans. Med. Imag.*, vol. 13, pp. 687–701, Dec. 1994.

⁷The reconstruction algorithm in [21] ignores overlap of the beam footprints.

- [8] J. A. Fessler, "Grouped coordinate descent algorithms for robust edge-preserving image restoration," in *Proc. SPIE 3071, Image Reconstruction and Restor. II*, 1997, pp. 184–194.
- [9] R. R. Meyer, "Sufficient conditions for the convergence of monotonic mathematical programming algorithms," *J. Comput. Syst. Sci.*, vol. 12, pp. 108–121, 1976.
- [10] A. H. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: An application to limited-angle tomography," *IEEE Trans. Image Processing*, vol. 7, pp. 204–221, Feb. 1998.
- [11] E. P. Ficaro, J. A. Fessler, P. D. Shreve, J. N. Kritzman, P. A. Rose, and J. R. Corbett, "Simultaneous transmission/emission myocardial perfusion tomography: Diagnostic accuracy of attenuation-corrected 99m-Tc-Sestamibi SPECT," *Circulation*, vol. 93, no. 3, pp. 463–473, Feb. 1996.
- [12] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc. Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.
- [13] J. E. Bowsher, M. P. Tornai, D. R. Gilland, D. E. G. Trotter, and R. J. Jaszczak, "An EM algorithm for modeling multiple or extended TCT sources," presented at the IEEE Nuc. Sci. Symp. Med. Im. Conf., 1999.
- [14] H. Erdoğan, G. Gualtieri, and J. A. Fessler, "Ordered subsets algorithms for transmission tomography," *Phys. Med. Biol.*, vol. 44, no. 11, pp. 2835–2851, Nov. 1999.
- [15] M. P. Becker, I. Yang, and K. Lange, "EM algorithms without missing data," *Stat. Meth. Med. Res.*, vol. 6, no. 1, pp. 38–54, 1997.
- [16] A. R. De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Trans. Med. Imag.*, vol. 14, pp. 132–137, Mar. 1995.
- [17] F. Yu, "Statistical Methods for Transmission Image Reconstruction with Nonlocal Edge-Preserving Regularization," Ph.D. dissertation, Univ. Michigan, Ann Arbor, M.I., 2000.
- [18] E. P. Ficaro, J. A. Fessler, J. N. Kritzman, P. C. Hawman, S. W. DeBruin, and J. R. Corbett, "Multiple line source array TCT/ECT system: A technical phantom evaluation," in *J. Nucl. Med. (Abs. Book)*, May 1999, vol. 40.
- [19] J. A. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum-likelihood): Applications to tomography," *IEEE Trans. Image Processing*, vol. 5, pp. 493–506, Mar. 1996.
- [20] D. M. Young, *Iterative Solution of Large Linear Systems*. New York: Academic, 1971.
- [21] A. Celler, A. Sitek, E. Stoub, P. Hawman, R. Harrop, and D. Lyster, "Multiple line source array for SPECT transmission scans: Simulation, phantom and patient studies," *J. Nuc. Med.*, vol. 39, no. 12, pp. 2183–2189, Dec. 1998.