Maximum Likelihood Transmission Image Reconstruction for Overlapping Transmission Beams

Daniel F. Yu, Jeffrey A. Fessler and Edward P. Ficaro
4415 EECS Bldg., University of Michigan, Ann Arbor, MI 48109-2122

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 SPECT 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.

Keywords: SPECT, attenuation maps, penalized-likelihood, multiple source.

I. Introduction

To reconstruct quantitatively accurate images of radioisotope emission distributions in 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 Figure 4 of [1]. It is natural to expect that higher and more uniform count profiles could lead to better reconstructed attenuation maps *if the overlap can be properly modeled by the reconstruction method*.

In both the scanning line source geometry and the geometry of Celler *et al* [1], there can be overlap of the beam footprints, as illustrated in Figure 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 2, we

give the statistical model and the proposed maximum-likelihood algorithm; in Section 3, we present some simulation results; in Section 4, we present the results from our preliminary study on optimal source collimation angles for a fixed system setup; and finally Section 5 is discussion and conclusion.

II. Statistical Model

Let Y_i denote the number of photons counted by the ith detector element[1] during the transmission scan, for $i = 1, \ldots, N$, 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 mean number of photons that would be observed during a transmission scan by the ith detector originating in the mth 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 $\mu^{\text{true}} = [\mu_1^{\text{true}}, \ldots, \mu_p^{\text{true}}]'$ 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 mth source and the ith detector location through the attenuating object is approximated by the following sum:

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

where $\boldsymbol{A}^m = \{a_{ij}^m\}$ 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 mth source in the direction of the i detector is (approximately) $\exp(-[\boldsymbol{A}^m \mu]_i)$.

We assume the Y_i's have independent Poisson distributions:

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

where the means are given by

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

The r_i's are nonnegative constants that one can include to account for the mean contributions of scatter, room background, and emission crosstalk [11]. We treat these r_i's as known constants, though in practice they must be determined experimentally. However, since scatter is a spatially smooth function, one can safely smooth scatter estimates fairly heavily, so generally the uncertainty in the r_i's can be made much smaller than that of the Y_i's.

The summation over m in (1) allows for arbitrary overlap of the beams transmitted from each source. Non-overlapping 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 $k \neq m$.

[1] Each "detector element" corresponds to a unique radial position and view angle, i.e., for typical 2D 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."

[2] Normalized by strip width

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

$$L(\mu, Y) = \sum_{i=1}^{N} h_i \left(\sum_{m=1}^{M} u_{im}(\mu) \right), \quad h_i(t) = y_i \log t - t, \tag{2}$$

$$u_{im}(\mu) = b_{im} \exp(-[\boldsymbol{A}^m \mu]_i) + r_i/M,$$

ignoring constants independent of μ. Since the form of this log-likelihood is sufficiently different from the usual models 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] 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 [13].

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:

$$\Phi(\mu, Y) = L(\mu, Y) - \beta R(\mu). \tag{3}$$

Our goal is to produce a penalized-likelihood estimate:

$$\hat{\mu} = \arg \max_{\mu \geq \underline{0}} \Phi(\mu, Y). \tag{4}$$

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

$$R(\mu) = \sum_{k=1}^{K} \psi_k([\boldsymbol{C}\mu]_k), \tag{5}$$

where the ψ_k's are potential functions acting as a norm on the "soft constraints" $\boldsymbol{C}\mu \approx 0$ and K is the number of such constraints. The functions ψ_k we consider here are convex, symmetric, nonnegative and differentiable [3].

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, 14] 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 nth iteration, the algorithm consists of repeatedly maximizing the surrogate function, *i.e.*

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

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

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

$$\frac{\partial Q}{\partial \mu_j}(\mu; \mu^n)\bigg|_{\mu=\mu^n} = \frac{\partial L}{\partial \mu_j}(\mu)\bigg|_{\mu=\mu^n}, \forall j = 1, \ldots, p \tag{8}$$

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

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

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 [15]. Because h_i is concave:

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

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

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

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

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

where

$$\begin{aligned} b_{im}^n &\triangleq \frac{\bar{y}_i{}^n}{u_{im}^n} b_{im}, \\ r_{im}^n &\triangleq \frac{\bar{y}_i{}^n}{u_{im}^n} \frac{r_i}{M} \\ g_{im}^n(l) &\triangleq y_i \log\left(b_{im}^n e^{-l} + r_{im}^n\right) - \left(b_{im}^n e^{-l} + r_{im}^n\right). \end{aligned}$$

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, 3] that the following paraboloidal function is a valid surrogate for Q_1:

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

where

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

and

$$\begin{aligned} \dot{g}_{im}^n(l) &\triangleq \frac{d}{dl} g_{im}^n(l) \\ \ddot{g}_{im}^n(l) &\triangleq \frac{d^2}{dl^2} g_{im}^n(l) \\ l_{im}^n = [A^m\mu^n]_i &= \sum_{j=1}^{p} \boldsymbol{A}_{ij}^m \mu_j^n. \end{aligned}$$

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. 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]:

$$c_{im}^n = \begin{cases} \left[-2\frac{g_{im}^n(0) - g_{im}^n(l_{im}^n) + \dot{g}_{im}^n(l_{im}^n)(l_{im}^n)}{(l_{im}^n)^2}\right]_+, & l_{im}^n > 0 \\ \left[-\ddot{g}_{im}^n(0)\right]_+, & l_{im}^n = 0 \end{cases} \tag{16}$$

$$= \begin{cases} \left[(2/(l_{im}^n)^2)\left\{b_{im}^n(1-e^{-l_{im}^n}) - y_i \log\frac{b_{im}^n + r_{im}^n}{\bar{y}_{im}^n} + b_{im}^n l_{im}^n e^{-l_{im}^n}\left(\frac{y_i}{\bar{y}_{im}^n} - 1\right)\right\}\right]_+, & l_{im}^n > 0 \\ \left[1 - \frac{y_i r_{im}^n}{(b_{im}^n + r_{im}^n)^2}\right]_+, & l_{im}^n = 0, \end{cases} \tag{17}$$

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].

Since our second surrogate Q_2 is a quadratic functional, it is easily maximized by a variety of algorithms, including the coordinate ascent algorithm [5, 16]. Adding a penalty function is straightforward. However, Q_2 is not separable and if we want to apply the ordered subsets idea, we must have a separable surrogate function.

A. Separable Paraboloidal Surrogate Algorithm

Now we derive a separable paraboloidal surrogate algorithm; we do not use this algorithm for the simulation described in this paper since it converges very slowly even though it is guaranteed to be monotonic. As noted by De Pierro in [15]:

$$[\boldsymbol{A}^m \mu]_i = \sum_{j=1}^p z_{ij}^m \left[\frac{a_{ij}^m}{z_{ij}^m}(\mu_j - \mu_j^n) + [\boldsymbol{A}^m \mu]_i\right], \tag{18}$$

provided $\sum_{j=1}^p z_{ij}^m = 1$ for all i and m. Thus if $z_{ij}^m \geq 0$ then it follows from the concavity of q_{im}^n that

$$q_{im}^n([\boldsymbol{A}^m \mu]_i) \geq \sum_{j=1}^p z_{ij}^m q_{im}^n\left(\frac{a_{ij}^m}{z_{ij}^m}(\mu_j - \mu_j^n) + [\boldsymbol{A}^m \mu]_i\right). \tag{19}$$

If we choose $z_{ij}^m = a_{ij}^m/\gamma_i^m$, where $\gamma_i^m = \sum_{j=1}^p a_{ij}^m$ (such that $\sum_{j=1}^p a_{ij}^m/\gamma_i^m = 1$ for all i and m), then from (12) and (19) we obtain our third surrogate function as follows:

$$Q_3(\mu; \mu^n) = \sum_{i=1}^N \sum_{m=1}^M \sum_{j=1}^p \frac{u_{im}^n a_{ij}^m}{\bar{y_i}^n \gamma_i^m} q_{im}^n(\gamma_i^m(\mu_j - \mu_j^n) + [\boldsymbol{A}^m \mu^n]_i). \tag{20}$$

This surrogate function is a quadratic form, and one that is trivial to maximize because it is a separable function. The partial derivatives of Q_3 are given as follows:

$$-\frac{\partial^2}{\partial \mu_j^2} Q_3(\mu; \mu^n) = \sum_{i=1}^N \sum_{m=1}^M \frac{u_{im}^n a_{ij}^m \gamma_i^m}{\bar{y_i}^n} c_{im}^n \triangleq d_j(\mu^n) \tag{21}$$

$$\frac{\partial}{\partial \mu_j} Q_3(\mu; \mu^n) = \sum_{i=1}^N \sum_{m=1}^M \frac{u_{im}^n a_{ij}^m}{\bar{y_i}^n} \dot{g}_{im}(l_{im}^n) - \sum_{i=1}^N \sum_{m=1}^M \frac{u_{im}^n a_{ij}^m \gamma_i^m}{\bar{y_i}^n} c_{im}^n (\mu_j - \mu_j^n). \tag{22}$$

Note that (8) is satisfied since

$$\frac{\partial}{\partial \mu_j} Q_3(\mu;\mu^n)\Big|_{\mu_j=\mu_j^n} = \sum_{i=1}^{N}\sum_{m=1}^{M} \frac{u_{im}^n a_{ij}^m}{\bar{y_i}^n} \dot{g}_{im}(l_{im}^n) = \frac{\partial}{\partial \mu_j} Q_2(\mu;\mu^n)\Big|_{\mu_j=\mu_j^n} = \frac{\partial}{\partial \mu_j} L(\mu, Y)\Big|_{\mu=\mu^n}. \tag{23}$$

From the above expressions for the partial derivatives of Q_3, the unconstrained maximizer of $Q_3(\mu;\mu^n)$ with regard to μ_j is given by:

$$\mu_j^n + \frac{1}{d_j(\mu^n)} \frac{\partial}{\partial \mu_j} Q_3(\mu;\mu^n)\Big|_{\mu_j=\mu_j^n}. \tag{24}$$

Since Q_3 is separable and concave, the iterative algorithm for the maximization of L is:

$$\mu^{n+1} = [\mu^n + \boldsymbol{D}(\mu^n)^{-1}\nabla' L(\mu^n, Y)]_+, \tag{25}$$

where $\nabla' = [\frac{\partial}{\partial \mu_1} \ldots \frac{\partial}{\partial \mu_p}]'$ denotes the (column) gradient operator, $[x]_+ = x$ for $x > 0$ and zero otherwise, and $\boldsymbol{D}(\mu^n)$ is a $p \times p$ diagonal matrix with jth diagonal entry $d_j(\mu^n)$. The ordered subsets idea could also be easily applied to this algorithm [17, 18] and adding regularization is straightforward as given in [18].

The iteration (25) monotonically increases the likelihood, however, it has very slow convergence due to the small curvatures of the surrogate parabolas. Applying the ordered subsets idea leads to faster "convergence" but monotonicity is no longer guaranteed.

B. Coordinate Ascent Algorithm

To obtain a monotonic algorithm that converges relatively quickly, we can apply coordinate ascent to the surrogate Q defined in (14), *i.e.,* sequentially update one pixel at a time while holding all other pixels fixed. First, we obtain the likelihood surrogate parabola for a particular pixel with every other pixel value fixed:

$$\begin{aligned} \hat{Q}_j^n(\mu_j) &\triangleq Q([\hat{\mu}_1, \ldots, \hat{\mu}_{j-1}, \mu_j, \hat{\mu}_{j+1}, \ldots, \hat{\mu}_p]; \mu^n) \\ &= Q(\hat{\mu};\mu^n) + \dot{Q}_j^n(\hat{\mu}_j)(\mu_j - \hat{\mu}_j) - \frac{1}{2} d_j^n (\mu_j - \hat{\mu}_j)^2, \end{aligned} \tag{26}$$

where $\hat{\mu}$ denotes the most recent estimate of μ, $\hat{\mu}_j$ denotes the jth entry of $\hat{\mu}$, μ^n denotes the estimate for μ after the nth iteration (with which $Q(\cdot,\mu^n)$ is constructed), and $\hat{Q}_j^n(\mu_j)$ is treated as a function of μ_j only. The derivative of the likelihood surrogate parabola at $\hat{\mu}_j$ is:

$$\dot{Q}_j^n(\hat{\mu}) \triangleq \frac{\partial}{\partial \mu_j} Q_j^n(\mu_j)\Big|_{\mu_j=\hat{\mu}_j} = \sum_{i=1}^{N}\sum_{m=1}^{M} \frac{u_{im}^n}{\bar{y}_i^n} a_{ij}^m [\dot{g}_{im}^n(l_{im}^n) - c_{im}^n(\hat{l}_{im} - l_{im}^n)], \tag{27}$$

where

$$\begin{aligned} \hat{l}_{im} &\triangleq [\boldsymbol{A}^m \hat{\mu}]_i = \sum_{j=1}^{p} \boldsymbol{A}_{ij}^m \hat{\mu}_j \\ \dot{g}_{im}^n(l) &= \left(\frac{y_i}{b_{im}^n e^{-l} + r_{im}^n} - 1 \right) b_{im}^n e^{-l}. \end{aligned}$$

The curvature of the parabola $\hat{Q}_j^n(\mu_j)$ is

$$d_j^n \triangleq \sum_{i=1}^{N}\sum_{m=1}^{M} \frac{u_{im}^n}{\bar{y}_i^n} (a_{ij}^m)^2 c_{im}^n. \tag{28}$$

Extension to the penalized likelihood case is straightforward following the methods in [2, 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. Combining the likelihood surrogate parabolas in (26) and the penalty surrogate parabolas, the maximization step of the coordinate ascent for pixel j is:

$$\hat{\mu}_j^{\text{new}} = \arg\max_{\mu_j \geq 0} \hat{Q}_j^n(\mu_j) - \beta \hat{R}_j(\mu_j) = \left[\hat{\mu}_j + \frac{\dot{Q}_j^n(\hat{\mu}) - \beta \dot{R}_j(\hat{\mu})}{d_j^n + \beta \hat{p}_j} \right]_+ . \tag{29}$$

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 (29) in a sequential order. We update the paraboloidal surrogate function after one iteration of coordinate ascent (CA), although one could also perform more than one CA iteration per surrogate. An outline of this algorithm is given in Table I, where ψ_k denote the potential function used in the penalty acting as a norm on the "soft constraints" $\boldsymbol{C}\mu \approx 0$, K is the number of such constraints, and $\hat{\psi}_k$ denotes the surrogate function used for ψ_k; see [3] for details.

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

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

Compute c_{im} according to (16)

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

for $j = 1, \ldots, p$

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

end

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

end

TABLE I

ALGORITHM OUTLINE

IV. Simulation Results

We compared the proposed reconstruction algorithm to 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 $\mathrm{Profile}^{\mathrm{TM}}$ system (Hoffman Estates, IL) [19]. The sources for the simulated system consisted of a multiple line source array with 14 sources, unequally spaced, located on a line parallel to the detector and 110 cm away from the detector plane. 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 pixels of size $3.56 \times 3.56\mathrm{mm}^2$. The sinogram size was 128×60 with detector bins of width $4.8\mathrm{mm}$ (*i.e.*, the simulated detector response was rectangular with width $4.8\mathrm{mm}$). 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, 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 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 5 (outlined by solid lines), and the large region used for the evaluation of spatial resolution (outlined by dashed lines) are shown in Figure 1. Figure 2 illustrates the system setup. The radial distribution of blank counts (at any projection angle) is shown in Figure 3.

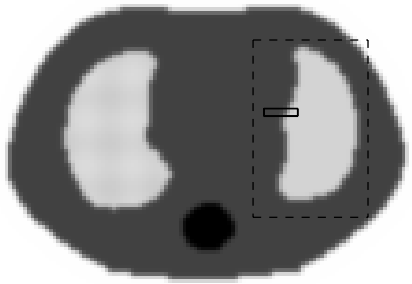


Fig. 1. Digital Phantom used in our simulations and the ROI used for collimation angle optimization.

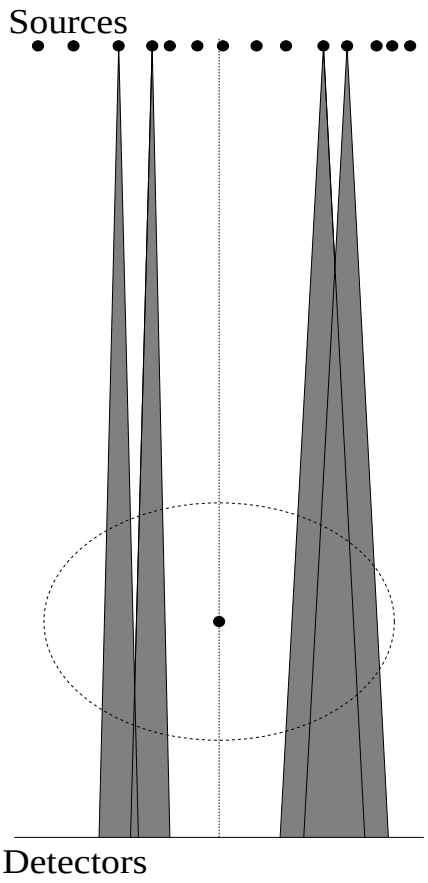
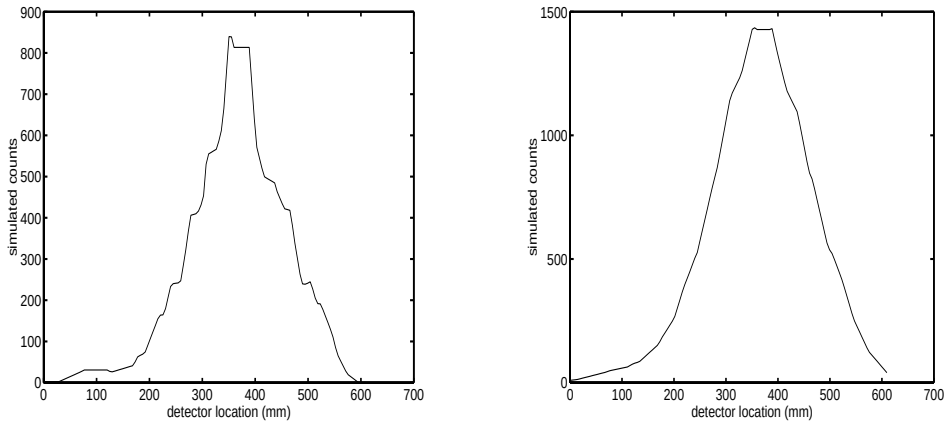


Fig. 2. Scaled illustration of the system setup; the two fan-beams on the left have collimation angle $2.6°$; the two fan-beams on the right side have collimation angle $5.6°$.



(a) (b)

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

Figures 4 and 5 show reconstructions of noisy data using FBP, the parallel algorithm[3], and the proposed algorithm, with new sources. Figure 4 shows reconstructions with $2.6°$ source collimation and Figure 5 shows reconstructions with $3.6°$ source collimation. Figures 6 and 7 show the same reconstructions with old sources. The spatial resolution

[3] We refer to the penalized-likelihood reconstruction assuming ideal normalized parallel "strip-integrals" (and ignoring beam overlap) as the parallel algorithm.

of the images in these figures 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:

$$\check{\mu} = \arg\max_{\mu \geq \underline{0}} \Phi(\mu, \bar{y}_i(\mu^{\text{true}})), \tag{30}$$

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

$$\arg\min_{\sigma} \sum_{j \in \mathcal{R}} \left| [G_\sigma \mu^{\text{true}}]_j - \check{\mu}_j \right|^2 \tag{31}$$

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 Figure 1.

We observe from Figures 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 5). 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 Figures 4-7 is due to regularization and noise. Figure 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 (Figure 8c).

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 collimation angle is 4.6° (Figure 8a)[4], 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.

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

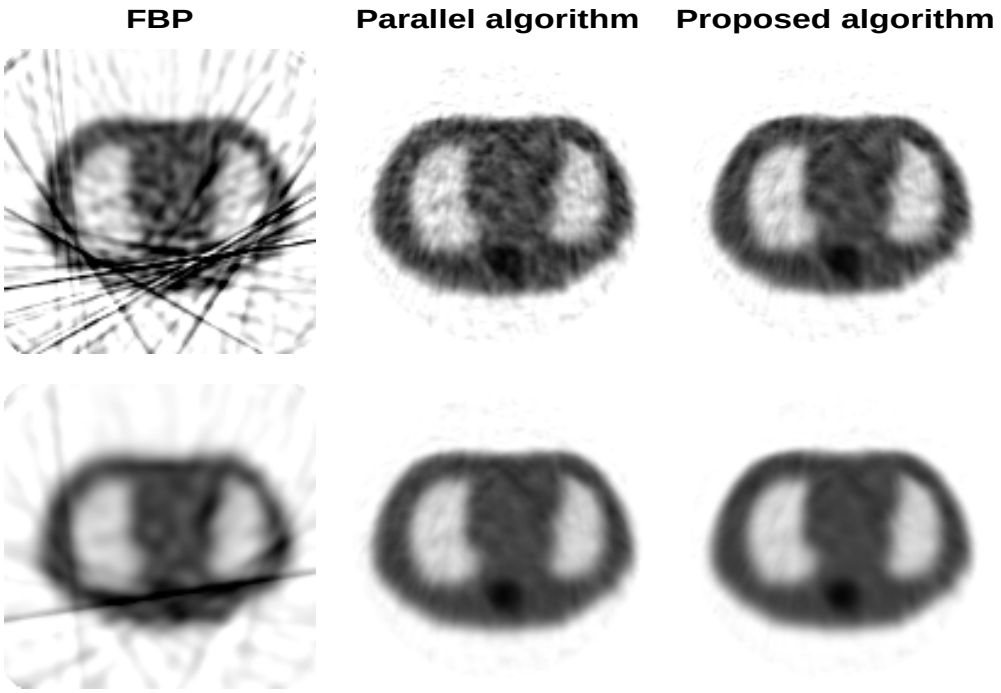


Fig. 4. New sources; collimation angle: 2.6°; 785,000 counts; top row: resolution 4.7 pixels; bottom row, resolution 6.8 pixels.

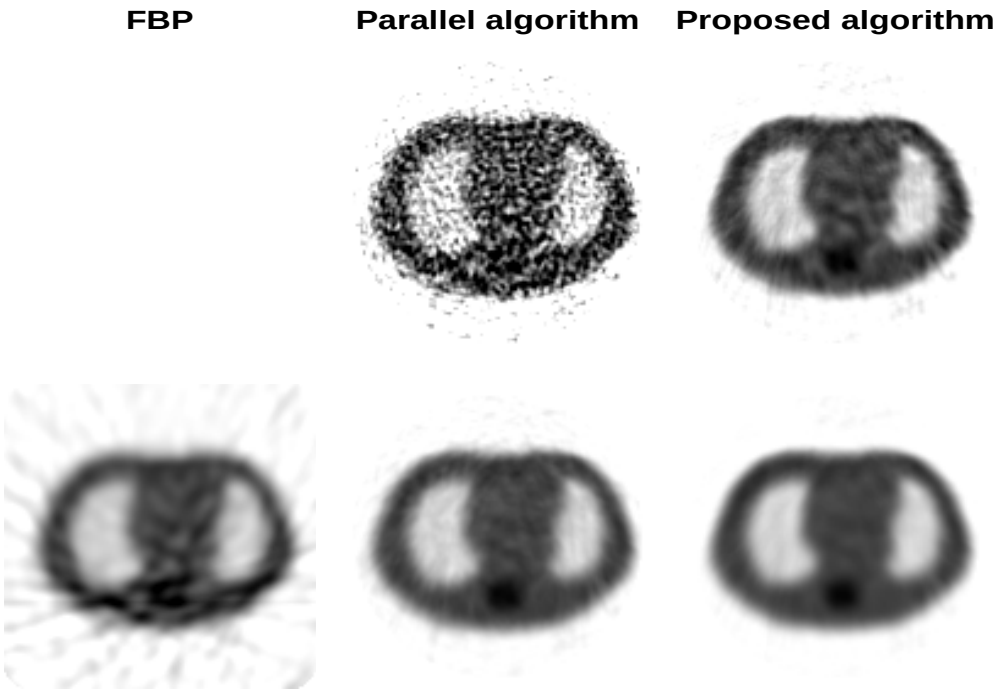


Fig. 5. New sources; collimation angle: 3.6°; 994,000 counts; top row: resolution 4.7 pixels; bottom row, resolution 6.8 pixels. A resolution of 4.7 pixels was not achievable with FBP in this case.

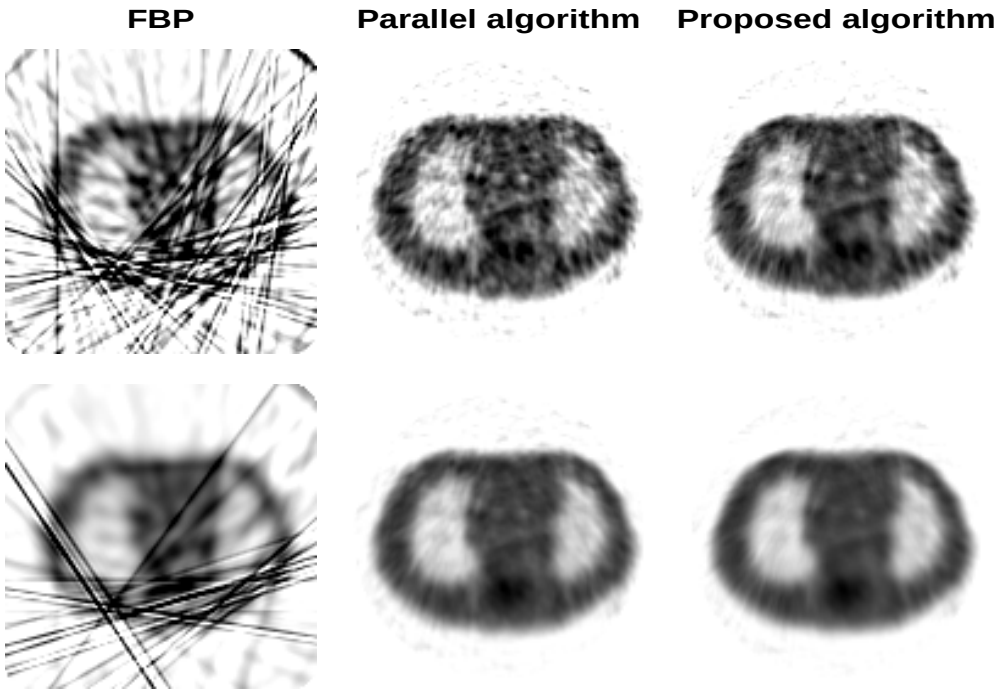


Fig. 6. Old sources; collimation angle: 2.6°; 392,000 counts; top row: resolution 4.7 pixels; bottom row, resolution 6.8 pixels.

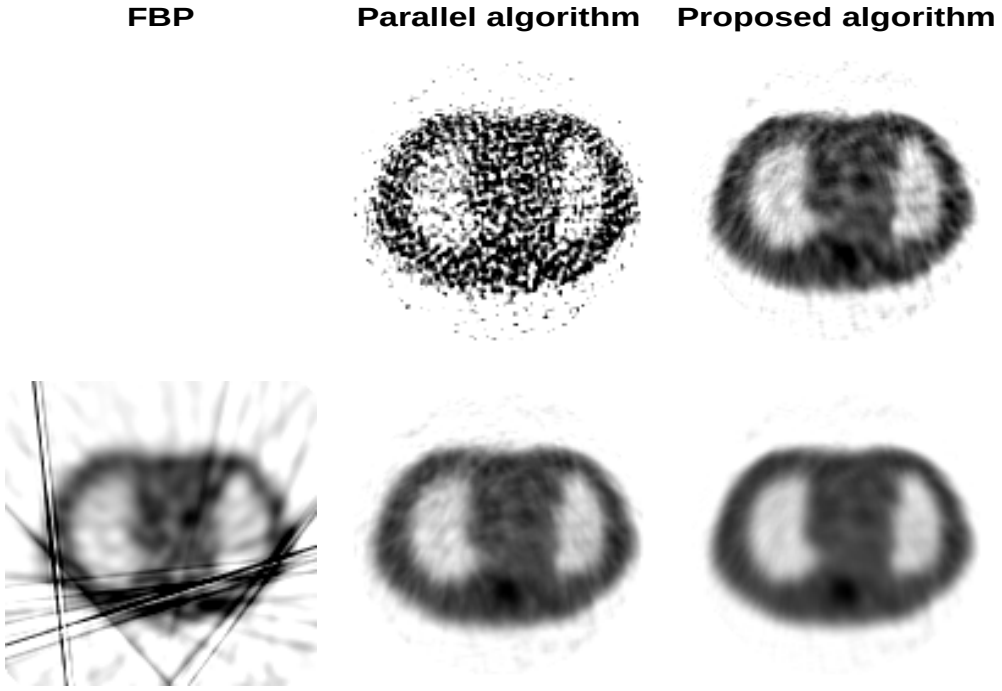
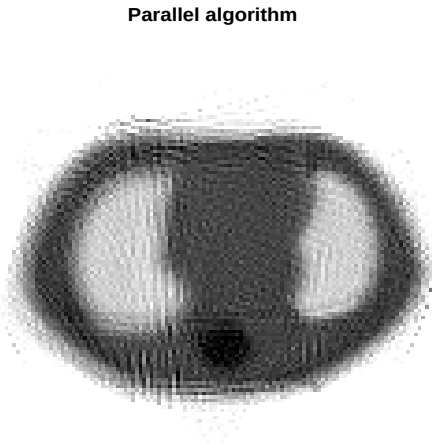
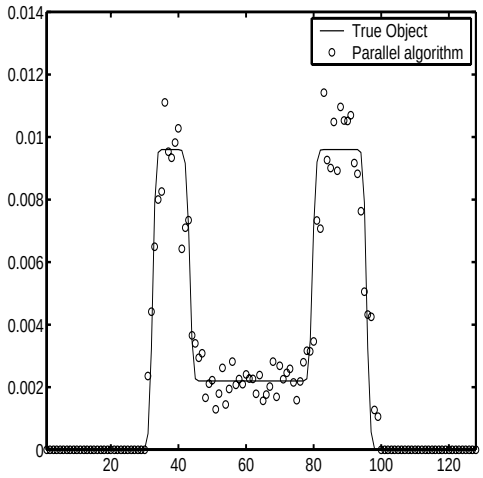
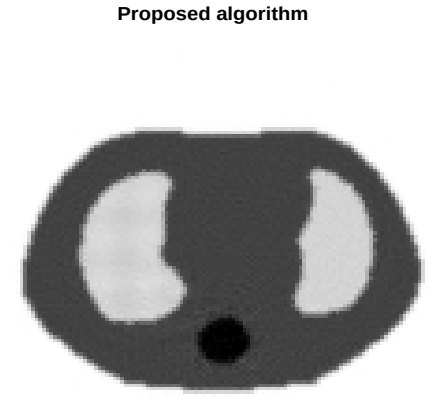


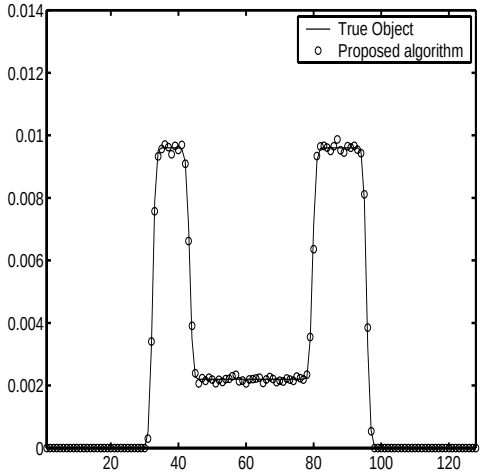
Fig. 7. Old sources; collimation angle: 3.6°; 497,000 counts; top row: resolution 4.7 pixels; bottom row, resolution 6.8 pixels. A resolution of 4.7 pixels was not achievable with FBP in this case.





(a) (b)





(c) (d)

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.

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 [20] to compute the variance of the ROI. For FBP, since numerical simulations are relatively inexpensive, we performed 2000 realizations for each data point.

The approximate covariance for an implicitly defined biased estimator is [20]:

$$\mathrm{Cov}(\hat{\mu}) \approx [-\nabla^{20}\Phi(\check{\mu},\bar{y})]^{-1}\nabla^{11}\Phi(\check{\mu},\bar{y})\mathrm{Cov}\{Y\}[\nabla^{11}\Phi(\check{\mu},\bar{y})]'[-\nabla^{20}\Phi(\check{\mu},\bar{y})]^{-1}, \tag{32}$$

where $\check{\mu}$ is defined in (30). 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

$$p_{im}(\mu) \triangleq b_{im}e^{-\sum_j a_{ij}^m \mu_j} \tag{33}$$

$$\bar{y}_i(\mu) = \sum_{m=1}^{M} p_{im}(\mu) + r_i, \tag{34}$$

then

$$\frac{\partial p_{im}(\mu)}{\partial \mu_j} = -a_{ij}^m p_{im}(\mu) = \frac{\partial \bar{y}_i(\mu)}{\partial \mu_j}. \tag{35}$$

Since

$$L(\mu,Y) = \sum_{i=1}^{N} y_i \log \bar{y}_i(\mu) - \bar{y}_i(\mu), \tag{36}$$

we have

$$\frac{\partial L(\mu,Y)}{\partial \mu_j} = \sum_{i=1}^{N} \frac{y_i}{\bar{y}_i(\mu)}\frac{\partial \bar{y}_i(\mu)}{\partial \mu_j} - \frac{\partial \bar{y}_i(\mu)}{\partial \mu_j} = \sum_{i=1}^{N}\left(1-\frac{y_i}{\bar{y}_i(\mu)}\right)\sum_{m=1}^{M} a_{ij}^m p_{im}(\mu), \tag{37}$$

and, applying the chain rule:

$$\frac{\partial^2 L(\mu,Y)}{\partial \mu_j \partial y_i} = -\frac{\sum_{m=1}^{M} a_{ij}^m p_{im}(\mu)}{\bar{y}_i(\mu)} \tag{38}$$

$$\begin{aligned}\frac{\partial^2 L(\mu,Y)}{\partial \mu_j \partial \mu_k} &= \frac{\partial}{\partial \mu_k}\sum_{i=1}^{N}\left(1-\frac{y_i}{\bar{y}_i(\mu)}\right)\sum_{m=1}^{M} a_{ij}^m p_{im}(\mu) \\ &= -\sum_{i=1}^{N}\left(1-\frac{y_i}{\bar{y}_i(\mu)}\right)\sum_{m=1}^{M} a_{ij}^m a_{ik}^m p_{im}(\mu) - \sum_{i=1}^{N}\frac{y_i}{\bar{y}_i^2(\mu)}\left(\sum_{m=1}^{M} a_{ij}^m p_{im}(\mu)\right)\left(\sum_{m=1}^{M} a_{ik}^m p_{im}(\mu)\right).\end{aligned} \tag{39}$$

Let $\mathbf{C}(\mu) \triangleq \{c_{ij}(\mu)\}$, where $c_{ij}(\mu) \triangleq \sum_{m=1}^{M} a_{ij}^m p_i^m(\mu)$ so $\mathbf{C}(\mu) = \sum_{m=1}^{M} \mathrm{diag}\{p_i^m(\mu)\}\boldsymbol{A}^m$. Then we obtain:

$$\nabla^{11}\Phi(\mu, Y) = -\mathbf{C}(\mu)'\mathrm{diag}\{1/\bar{y}_i(\mu)\} \tag{40}$$

$$\begin{aligned} -\nabla^{20}\Phi(\mu, Y) &= \sum_{m=1}^{M} (\boldsymbol{A}^m)'\mathrm{diag}\{(\bar{y}_i(\mu) - r_i)(1 - y_i/\bar{y}_i(\mu))\}\boldsymbol{A}^m \\ &\quad + T\mathbf{C}(\mu)'\mathrm{diag}\{y_i/\bar{y}_i^2(\mu)\}\mathbf{C}(\mu) + \beta\mathbf{R}(\mu), \end{aligned} \tag{41}$$

where $\mathbf{R}(\mu) = \nabla^2 R(\mu)$. (For a detailed derivation of the penalty part, see [20].) We use the following recipe to compute the approximate estimator (using the proposed algorithm) variance of a ROI:

(i) Compute $\check{\mu}$ by applying the proposed algorithm to noise-free data $\{\bar{y}_i(\mu^{\mathrm{true}})\}$;
(ii) Forward project $\check{\mu}$ to compute $\bar{y}_i(\check{\mu}) = \sum_{m=1}^{M} p_{im}(\check{\mu}) + r_i$;
(iii) Use an iterative method such as conjugate gradient [21] or Gauss-Siedel [22] to solve $[-\nabla^{20}\Phi(\check{\mu}, \bar{y})]u^{\mathrm{ROI}} = e^{\mathrm{ROI}}$, where e^{ROI} is a vector with entries $1/n_{\mathrm{ROI}}$ for pixels inside the region and zeros for pixels outside; n_{ROI} denotes the number of pixels in the ROI;
(iv) Compute $(u^{\mathrm{ROI}})'\nabla^{11}\Phi(\check{\mu}, \bar{y})\mathrm{Cov}\{Y\}[\nabla^{11}\Phi(\check{\mu}, \bar{y})]'u^{\mathrm{ROI}}$ by first forward projecting u^{ROI} to compute $v = \mathrm{diag}\{1/\bar{y}_i(\check{\mu})\}\mathbf{C}(\check{\mu})u^{\mathrm{ROI}}$, and then summing:

$$\mathrm{Var}(\hat{\mu}^{\mathrm{ROI}}) \approx \sum_{i=1}^{N} v_i^2 y_i(\mu^{\mathrm{true}}). \tag{42}$$

For the parallel algorithm, the computation of the approximate variance has been outlined in [20], except that we use $\mathrm{Cov}(Y) = \mathrm{diag}\{\bar{y}_i(\mu^{\mathrm{true}})\}$ in (23) of [20], where $\bar{y}_i$ is based on the overlapping beam model (34) rather than the parallel strip-integral model of [20].

We analyzed the resolution/variance tradeoffs for the ROI illustrated in Figure 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 4. 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. Figures 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 Figures 11 and 12. The proposed algorithm outperforms the parallel algorithm which in 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 $> 7°$ 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. Figure 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. Figure 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%.

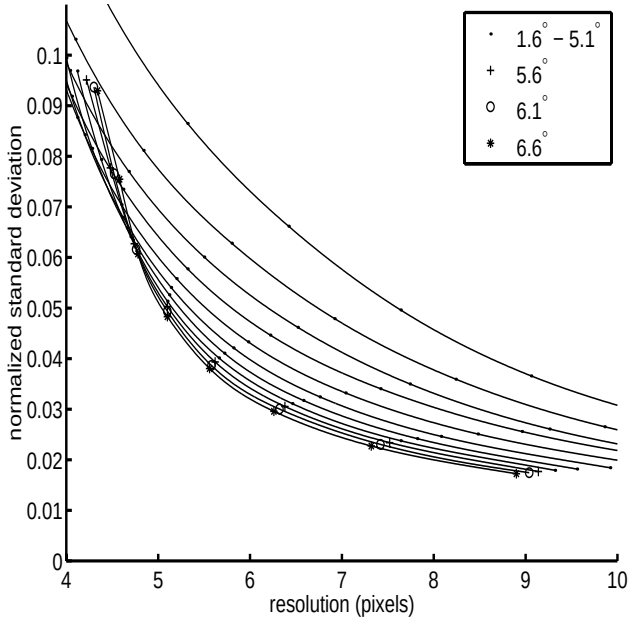


Fig. 9. High count case: ROI resolution/variance curves for the proposed algorithm.

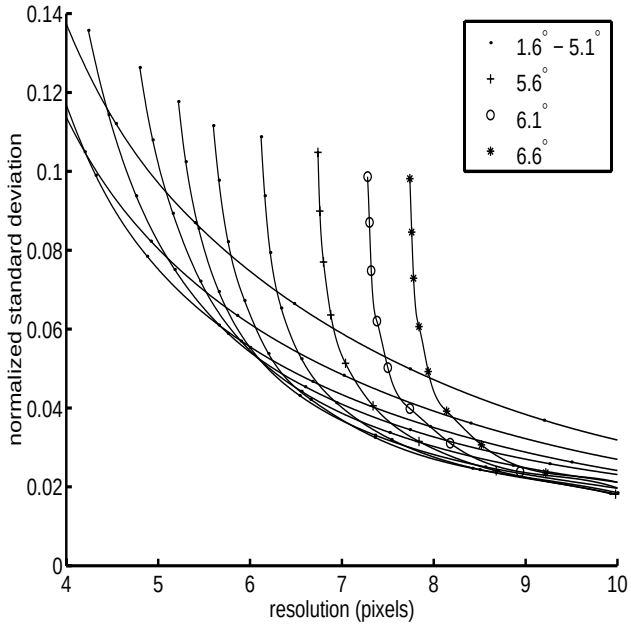


Fig. 10. High count case: ROI resolution/variance curves for the parallel algorithm.

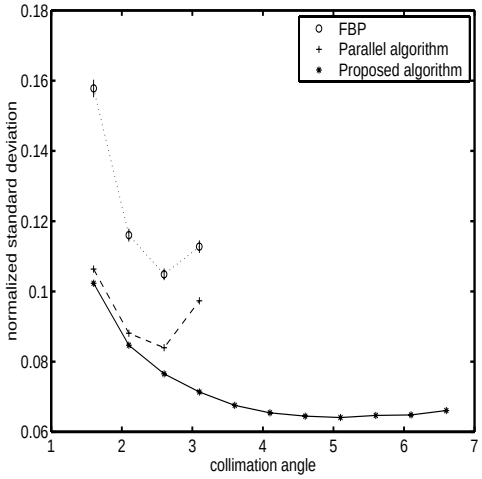


Fig. 11. High count case: ROI variances at desired spatial resolution 4.7 pixels.

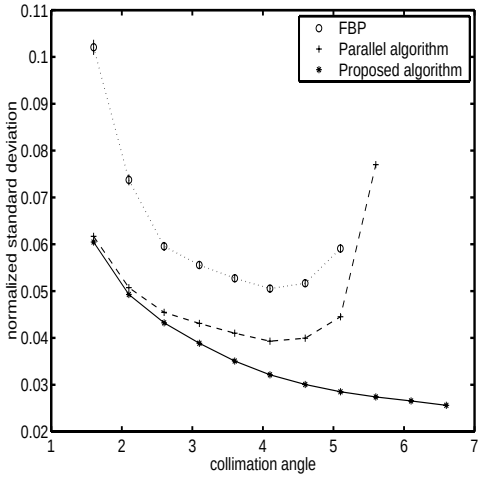


Fig. 12. High count case: ROI variances at desired spatial resolution 6.8 pixels.

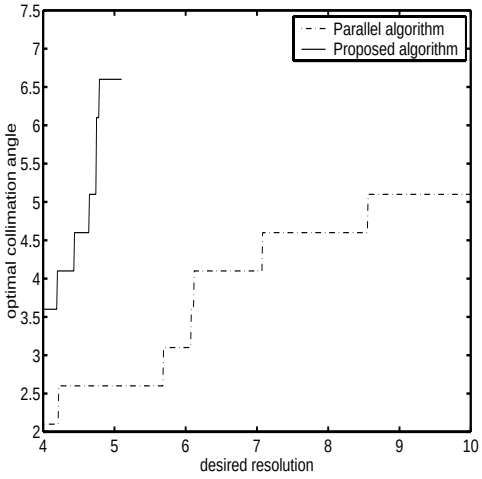


Fig. 13. High count case: A comparison of optimal collimation angles, at different resolutions for the proposed and parallel algorithms.

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 [23]. The algorithm is

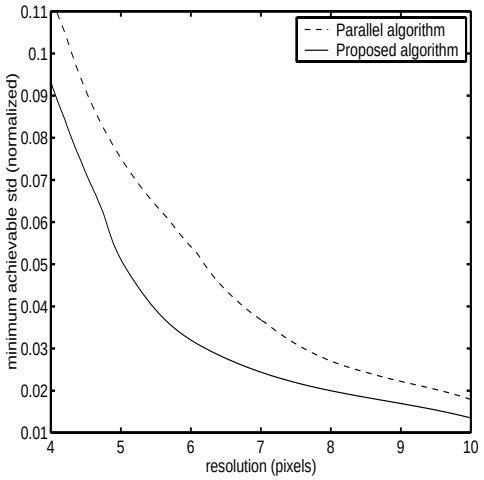


Fig. 14. High count case: A comparison of minimum achievable normalized standard deviation, at different resolutions for the proposed and parallel algorithms.

guaranteed to monotonically increase the objective function at each iteration, and consistently achieves better variance/resolution tradeoffs than "conventional" image reconstruction algorithms, both statistical (the parallel algorithm) and non-statistical (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.) 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.

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 seconds, compared to about 0.8 seconds needed for the parallel algorithm, on a Sun Ultra2 workstation. Furthermore, because of the overlap between transmission 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 (16) 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 3. With regard to the number of iterations necessary for convergence, *i.e.*, the smallest n such that $\Phi(\mu^n) - \Phi(\mu^0) > 0.999 \left[\Phi(\mu^*) - \Phi(\mu^0)\right]$, 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.

VII. Acknowledgment

The authors gratefully acknowledge Peggy Hawman for assistance with the Siemens Profile geometrical specifications. NIH grants CA-60711 and CA-54362.

References

[1] A. Celler, A. Sitek, and R. Harrop, "Reconstruction of multiple line source attenuation maps," *IEEE Tr. Nuc. Sci.*, vol. 44, no. 4, pp. 1503–8, August 1997.

[2] H. Erdoğan and J. A. Fessler, "Accelerated monotonic algorithms for transmission tomography," in *Proc. IEEE Intl. Conf. on Image Processing*, volume 2, pp. 680–4, 1998.

[3] H. Erdoğan and J. A. Fessler, "Fast monotonic algorithms for transmission tomography," *IEEE Tr. Med. Im.*, 1998. To appear.

[4] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *J. Comp. Assisted Tomo.*, vol. 8, no. 2, pp. 306–316, April 1984.

[5] K. Sauer and C. Bouman, "A local update strategy for iterative reconstruction from projections," *IEEE Tr. Sig. Proc.*, vol. 41, no. 2, pp. 534–548, February 1993.

[6] K. Lange and J. A. Fessler, "Globally convergent algorithms for maximum a posteriori transmission tomography," *IEEE Tr. Im. Proc.*, vol. 4, no. 10, pp. 1430–8, October 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 Tr. Med. Im.*, vol. 13, no. 3, pp. 687–701, December 1994.

[8] J. A. Fessler, "Grouped coordinate descent algorithms for robust edge-preserving image restoration," in *Proc. SPIE 3071, Im. Recon. and Restor. II*, pp. 184–94, 1997.

[9] R. R. Meyer, "Sufficient conditions for the convergence of monotonic mathematical programming algorithms," *J. Comput. System. Sci.*, vol. 12, pp. 108–21, 1976.

[10] A. H. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography," *IEEE Tr. Im. Proc.*, vol. 7, no. 2, pp. 204–221, February 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. 463–73, February 1996.

[12] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Stat. Soc. Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.

[13] H. Erdoğan, G. Gualtieri, and J. A. Fessler, "An ordered subsets algorithm for transmission tomography," in *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, 1998. Inadvertently omitted from proceedings. Available from web page.

[14] 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.

[15] A. R. De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Tr. Med. Im.*, vol. 14, no. 1, pp. 132–137, March 1995.

[16] J. A. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE Tr. Med. Im.*, vol. 13, no. 2, pp. 290–300, June 1994.

[17] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Tr. Med. Im.*, vol. 13, no. 4, pp. 601–609, December 1994.

[18] H. Erdoğan and J. A. Fessler, "Ordered subsets algorithms for transmission tomography," *Phys. Med. Biol.*, 1999.

[19] 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," *J. Nuc. Med. (Abs. Book)*, vol. 40, no. 5, pp. 4, May 1999.

[20] J. A. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography," *IEEE Tr. Im. Proc.*, vol. 5, no. 3, pp. 493–506, March 1996.

[21] D. M. Young, *Iterative solution of large linear systems*, Academic Press, New York, 1971.

[22] N. H. Clinthorne, T. S. Pan, P. C. Chiao, W. L. Rogers, and J. A. Stamos, "Preconditioning methods for improved convergence rates in iterative reconstructions," *IEEE Tr. Med. Im.*, vol. 12, no. 1, pp. 78–83, March 1993.

[23] 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. 2183–9, December 1998.

This work was supported in part by NIH grants CA-60711 and CA-54362.