An Ordered Subsets Algorithm for Transmission Tomography

Hakan Erdoğan[1] Gene Gualtieri[2] and Jeffrey A. Fessler[1]
[1] University of Michigan
[2] UGM Medical Systems
4415 EECS Bldg., University of Michigan, Ann Arbor, Mi 48109-2122

Abstract

The ordered subsets EM (OSEM) algorithm has enjoyed considerable interest for emission image reconstruction due to its acceleration of the original EM algorithm and ease of programming. The transmission EM reconstruction algorithm converges very slowly and is not used in practice, particularly because there are faster simultaneous update algorithms that converge much faster. We introduce such an algorithm called separable paraboloidal surrogates (SPS) in this paper which is also monotonic even with nonzero background counts. We demonstrate that the ordered subsets method can also be applied to the new algorithm to accelerate "convergence" for the transmission tomography problem, albeit with similar sacrifice of global convergence properties as OSEM. We implemented and evaluated this ordered subsets transmission (OSTR) algorithm. The results indicate that the OSTR algorithm speeds up the increase in the objective function by roughly the number of subsets in the early iterates when compared to the ordinary SPS algorithm. We compute mean square errors and segmentation errors for different methods and show that OSTR method is superior to OSEM applied to the logarithm of the transmission data. But, penalized-likelihood reconstructions yield the best quality images among all other methods tested.

I. Introduction

Attenuation is an important factor that should be corrected for in emission computed tomography. In modern systems, transmission scans are performed in addition to emission scans to correct for the effects of attenuation. Statistical methods can be used to reconstruct attenuation maps from these scans which in turn can be used in emission image reconstructions as attenuation correction factors (ACFs) to yield quantitatively accurate images.

Algorithms exist for maximum likelihood (ML) and maximum penalized likelihood (PL) transmission image reconstruction problems. Most of the recent ones [1, 2] are based on direct maximization of the objective function rather than the famous expectation maximization (EM) algorithm [3] due to the fact that EM algorithm for transmission case converges very slowly [4].

Recently, ordered subsets EM (OSEM) [5] for the emission problem has been used a lot in emission image reconstruction. This algorithm is frequently used because of the following reasons.

- OSEM provides order-of-magnitude acceleration over EM in ML problems.
- The reconstructed images look good after only a few iterations.
- OSEM is implemented by slightly modifying the well-known and well-established EM algorithm.
- OSEM is easily implemented with any type of system model.

However, although the images seem to look good, the resolution and variance properties of OSEM are unclear. In addition it does not converge and may cycle. Due to its popularity, OSEM for emission problem was even applied to transmission data after taking its logarithm. In the results section, we show that this approach yields lower quality images than the ordered subsets transmission (OSTR) algorithm we introduce in this paper.

Manglos *et al.* [6] applied the ordered subsets idea to the transmission EM method. Although ordered subsets accelerate convergence of the original transmission EM algorithm, it is still affected by the slow convergence of it. The separable praboloidal surrogates (SPS) algorithm we base our OSTR algorithm is much faster than the transmission EM algorithm.

The ordered subsets method can be applied to any algorithm which involves a sum over sinogram indices. The sum over all the sinogram indices are replaced by a sum over a subset of the data and an ordered subsets version of the algorithm is obtained. However, it is best to apply this idea to algorithms which update the parameters simultaneously at each iteration rather than greedy sequential update algorithms. This is due to the fact that greedy algorithms such as coordinate ascent tend to update high frequencies faster [1]. When only a subset of the data is used as in ordered subsets, there is no point in making high frequency details converge.

In this paper, we apply the ordered subsets idea to a separable paraboloidal surrogates (SPS) algorithm. Paraboloidal surrogate [7] is a quadratic function that is designed to lie below the log-likelihood. Using concavity [2], we get a separable quadratic function that lies below this paraboloid. This separable surrogate function can be maximized by a simple simultaneous update. The SPS algorithm is intrinsically monotonic. However, the ordered subsets version is not guaranteed to be monotonic and does not converge for number of subsets greater than one.

In the rest of the paper, we first introduce the problem and the OSTR algorithm for general penalized-likelihood (PL) objective. Then, we present results on real PET transmission data with maximum likelihood and (ML) and PL reconstructions. We analyze the algorithms in terms of their mean squared error. We also perform hard segmentation on the reconstructed images to analyze their tissue classification performance.

II. The Problem

For transmission scan measurements, it is realistic to assume the following model:

$$y_i \sim \text{Poisson}\{b_i e^{-[\boldsymbol{A}\mu]_i} + r_i\},\ i = 1, \ldots, N,$$

where N is the number of measurements, μ_j is the average linear attenuation coefficient in voxel j for $j = 1, \ldots, p$, and p denotes the number of voxels. The notation $[\boldsymbol{A}\mu]_i = \sum_{j=1}^{p} a_{ij}\mu_j$ represents the line integral of the attenuation map μ, and $\boldsymbol{A} = \{a_{ij}\}$ is the $N \times p$ system matrix. We assume that $\{b_i\}, \{r_i\}$ and $\{a_{ij}\}$ are known nonnegative constants, where r_i is the mean number of background events, b_i is the blank scan count and y_i represents the number of coincident transmission events counted by the i^{th} detector pair.

The log-likelihood function for the independent transmission data is:

$$L(\mu) = \sum_{i=1}^{N} h_i([\boldsymbol{A}\mu]_i),$$

where

$$h_i(l) = y_i \log(b_i e^{-l} + r_i) - (b_i e^{-l} + r_i),$$

ignoring the constant terms. Directly maximizing $L(\mu)$ (ML method) results in a very noisy estimate $\hat{\mu}$. Segmentation of the attenuation map is commonly performed to reduce noise afterwards. Penalized-likelihood (PL) (or MAP) methods regularize the problem and reduce the noise by adding a roughness penalty to the objective function as follows:

$$\hat{\mu} = \arg\max_{\mu \geq 0} \Phi(\mu), \quad \Phi(\mu) = L(\mu) - \beta R(\mu). \tag{1}$$

The roughness penalty R is of this form:

$$R(\mu) = \frac{1}{2} \sum_{j=1}^{p} \sum_{k \in \mathcal{N}_j} w_{jk} \psi(\mu_j - \mu_k), \tag{2}$$

where ψ penalizes neighboring pixel differences.

In the following discussion, we use the PL formulation to derive the new algorithm. If one wants to implement the ML method, setting $\beta = 0$ should yield the ML estimator.

III. The Algorithm

A. The Likelihood Part

We presented the paraboloidal surrogates algorithm for transmission tomography previously [7, 8]. We first find a one-dimensional surrogate parabola $q_i(l; l_i^n)$ that is tangent to the marginal log-likelihood function $h_i(l)$ at the current iterate $l_i^n = [\boldsymbol{A}\mu^n]_i$ and lies below it for all $l > 0$. Then, we sum up these parabolas to obtain an overall paraboloidal surrogate function for the log-likelihood as follows:

$$Q_1(\mu; \mu^n) \stackrel{\triangle}{=} \sum_{i=1}^{N} q_i([\boldsymbol{A}\mu]_i; l_i^n) \leq L(\mu),\ \forall \mu \geq 0,$$

where

$$q_i(l; l_i^n) \stackrel{\triangle}{=} h_i(l_i^n) + \dot{h}_i(l_i^n)(l - l_i^n) + 1/2 c_i(l_i^n)(l - l_i^n)^2.$$

The optimum curvature that provides fastest convergence rate while preserving monotonicity was shown to be [7]

$$c_i(l_i^n) = \begin{cases} \left[2\dfrac{h_i(0) - h_i(l_i^n) + \dot{h}_i(l_i^n)(l_i^n)}{(l_i^n)^2} \right]_+, & l_i^n > 0, \\ \left[\ddot{h}_i(0) \right]_+, & l_i^n = 0. \end{cases}$$

This new objective $Q_1(\mu; \mu^n)$ and each $q_i(l; l_i^n)$ are naturally concave. Previously, we used coordinate ascent to maximize this objective function [7]. This approach leads to a very fast and monotonic algorithm. However, the computational advantages only exist if the system matrix is precomputed and column accessible [9] [1]. However, this problem does not exist for algorithms that update all the parameters simultaneously. Simultaneous update algorithms are parallelizable whereas sequential updates are not. To maximize this new objective by simultaneously updating all the parameters, we can use the additive concavity trick employed in [2] which makes use of DePierro's arguments about concave functions [10, 11]. These arguments provide a separable surrogate function $Q_2(\mu; \mu^n)$ for $Q_1(\mu; \mu^n)$ as follows:

$$\begin{aligned} Q_1(\mu; \mu^n) &= \sum_{i=1}^{N} q_i \left(\sum_{j=1}^{p} \frac{a_{ij}}{\gamma_i} \left[\gamma_i(\mu_j - \mu_j^n) + [\boldsymbol{A}\mu^n]_i \right] \right) \\ &\geq \sum_{i=1}^{N} \sum_{j=1}^{p} \frac{a_{ij}}{\gamma_i} q_i \left(\gamma_i(\mu_j - \mu_j^n) + [\boldsymbol{A}\mu^n]_i ; l_i^n \right) \\ &\stackrel{\triangle}{=} Q_2(\mu; \mu^n),\ \forall \mu \in \mathbb{R} \end{aligned} \tag{3}$$

where $\gamma_i = \sum_{k=1}^{p} a_{ik}$ is the projection of an image of all ones. Note that the terms a_{ij}/γ_i sum up to unity (over j) and the inequality follows from the definition of concavity. The function $Q_2(\mu; \mu^n)$ is now separable in j and quadratic, so that the exact maximization is reduced to maximization of p 1D functions each of which depend on one pixel value μ_j only.

B. The Penalty Part

Since the maximum likelihood problem is ill-posed (*i.e.* small changes in data result in big changes in the estimates), some kind of regularization is necessary. The discussion in section A provided separable surrogate functions for the log-likelihood function. A similar separable surrogate is needed for the penalty part $R(\mu)$ for a simultaneous update.

We exploit the convexity of the potential function $\psi(t)$ to obtain the surrogate. For completeness, we repeat the arguments in [11, 12]:

$$\begin{aligned} \psi(\mu_j - \mu_k) &= \psi \left(\frac{1}{2} \left[2\mu_j - \mu_j^n - \mu_k^n \right] \right. \\ &\quad \left. + \frac{1}{2} \left[-2\mu_k + \mu_j^n + \mu_k^n \right] \right) \\ &\leq \hat{\psi}(\mu; \mu^n) \stackrel{\triangle}{=} \frac{1}{2}\psi(2\mu_j - \mu_j^n - \mu_k^n) \\ &\quad + \frac{1}{2}\psi(2\mu_k - \mu_j^n - \mu_k^n) \end{aligned} \tag{4}$$

[1] For real 3D problems system matrix is huge, so precomputation can be impractical.

Using this inequality, one gets the following separable surrogate function:

$$S(\mu;\mu^n) \stackrel{\triangle}{=} \sum_{j=1}^{p} \sum_{k\in\mathcal{N}_j} w_{jk}\hat{\psi}(\mu;\mu^n) \geq R(\mu),\ \forall \mu \in \mathbb{R} \quad (5)$$

C. The SPS Algorithm

We designed separable surrogate functions for both the likelihood and the penalty parts in the preceding sections. By combining those, we define the global surrogate function

$$\phi(\mu;\mu^n) \stackrel{\triangle}{=} Q_2(\mu;\mu^n) - \beta S(\mu;\mu^n).$$

This global separable surrogate function satisfies:

$$\Phi(\mu) = L(\mu) - \beta R(\mu) \geq \phi(\mu;\mu^n),\ \forall \mu \geq 0$$

and $\phi(\mu;\mu^n)$ is tangent to $\Phi(\mu)$ at current iterate μ^n. We maximize (or increase) the function $\phi(\mu;\mu^n)$ at each iteration and repeat the procedure iteratively. We call this algorithm separable paraboloidal surrogates (SPS) algorithm. One can show [7] that increasing the surrogate function $\phi(\mu;\mu^n)$ also increases the original objective function $\Phi(\mu)$. Hence, this algorithm is intrinsically monotonic. The maximization of $\phi(\mu;\mu^n)$ is easy. Due to the additive separability, the update for each parameter only involves the parameter itself. When a quadratic penalty is used, *i.e.* $\psi(t) = t^2/2$, the maximization can be done exactly in a single step via Newton's algorithm as follows:

$$\mu^{n+1} = \mu^n + \boldsymbol{D}^{-1}\nabla'\Phi(\mu^n) \quad (6)$$

where $\boldsymbol{D}$ is a $p \times p$ diagonal matrix with diagonal entries

$$\boldsymbol{D}_{jj} = d_j^n + 2\beta \sum_k w_{jk},\ \text{for } j = 1 \ldots p.$$

The factor 2 in the denominator comes from the separable surrogate $S(\mu;\mu^n)$ in (5). The denominator terms d_j^n are:

$$d_j^n = -\sum_{i=1}^{N} a_{ij}\gamma_i c_i(l_i^n). \quad (7)$$

However, for transmission tomography, it is advantageous to use edge-preserving nonquadratic penalties to get images with sharper edges. An example for nonquadratic penalties is [13]:

$$\psi(t) = \delta^2\left[|t/\delta| - \log(1+|t/\delta|)\right],$$

where δ is a twiddle factor. We used this penalty in our PL reconstruction results.

In the nonquadratic penalty case, exact maximization is not easy, but one can monotonically increase the surrogate objective by doing a few sub-iterations of the following type.

$$\hat{\mu}_j := \hat{\mu}_j + \frac{\frac{\partial}{\partial\mu_j}\phi(\hat{\mu};\mu^n)}{d_j^n + 2\beta\sum_{k\in\mathcal{N}_j} w_{jk}\omega_\psi(\hat{\mu}_j - \mu_k^n)}, \quad (8)$$

where

$$\omega_\psi(t) = \frac{\dot{\psi}(t)}{t}.$$

The detailed explanation of the $\omega_\psi(t)$ function can be found in [7, 14]. The partial derivative of the surrogate ϕ with respect to μ_j can be found as:

$$\begin{aligned}\frac{\partial}{\partial\mu_j}\phi(\hat{\mu};\mu^n) &= \sum_{i=1}^{N} a_{ij}\dot{h}_i(l_i^n) - d_j^n(\hat{\mu}_j - \mu^n) \\ &\quad -\beta\sum_{k\in\mathcal{N}_j} w_{jk}\dot{\psi}(\hat{\mu}_j - \mu^n), \end{aligned} \quad (9)$$

where $\dot{h}_i(l) = \left(\frac{y_i}{b_i e^{-l} + r_i} - 1\right) b_i e^{-l}$.

Next, we apply the ordered subsets idea to the simultaneous update algorithm developed above.

D. Ordered Subsets

Ordered subsets idea can be used with any algorithm as long as the algorithm involves a sum over sinogram indices. The SPS algorithm introduced above also contains sums over sinogram indices in computing the denominator d_j^n terms (7) and the gradient terms $\frac{\partial}{\partial\mu_j}\phi$ (9). We apply the ordered subsets idea to this algorithm.

Ordered subsets methods group projection data into an ordered sequence of blocks and processes each block at once. These blocks are usually chosen so that the projections within one block correspond to projections of the image with downsampled projection angles.

Let M be the number of subsets chosen in the projection domain. Let $S_1, \ldots, S_M$ denote the subsets in the order selected. At step m the following objective function corresponding to the subset S_m should be maximized (or increased):

$$\Phi_m(\mu;y) = M\left\{\sum_{i\in S_m} h_i([\boldsymbol{A}\mu]_i)\right\} - \beta R(\mu)$$

The scaling of the log-likelihood function is necessary in order that the effect of β value is independent of the number of subsets. One iteration is completed when the algorithm goes through all the projections by going through all the subsets. The modification of the SPS algorithm to incorporate ordered subsets idea is relatively easy. We call the resulting algorithm ordered subsets transmission (OSTR) algorithm. The outline of the algorithm is given in Table 1. The maximization step is similar to (8) but the curvature and gradient terms are computed using just a subset of the data.

The OSTR algorithm reduces to the SPS algorithm when $M = 1$. Since the projections and backprojections are performed for only the elements of a single block, processing of each block in an OSTR algorithm with M subsets (OSTR-M) roughly takes $1/M$ of time that it would take for one iteration of the SPS algorithm. For PL problem, actually it would take a little more than $1/M$ of the time since the CPU time required for computing the gradient and curvatures of the penalty surrogate do not change. Yet, one hopes that processing of one block increases the objective function as much as one iteration of the original algorithm. That is, the objective increase for M iterations of OSTR-1 should be close to that increase for one

for each iteration $n = 1, \ldots,$ niter

for each subset m=1,...,M

$$\hat{l}_i = \sum_{j=1}^{p} a_{ij}\hat{\mu}_j,\ \dot{h}_i = \left(\frac{y_i}{b_i e^{-\hat{l}_i} + r_i} - 1\right) b_i e^{-\hat{l}_i},\ \forall i \in S_m$$

$$\mu^{\text{old}} = \hat{\mu}$$

for $j = 1, \ldots, p$

$$\dot{L}_j = M \sum_{i \in S_m} a_{ij}\dot{h}_i,$$

$$d_j = -M \sum_{i \in S_m} a_{ij}\gamma_i c_i(\hat{l}_i)$$

for a couple sub-iterations

$$\hat{\mu}_j := \left[\hat{\mu}_j + \frac{\dot{L}_j - d_j(\hat{\mu}_j - \mu_j^{\text{old}}) - \beta \sum_k w_{jk}\dot{\psi}(\hat{\mu}_j - \mu_k^{\text{old}})}{d_j + 2\beta \sum_k w_{jk}\omega_\psi(\hat{\mu}_j - \mu_k^{\text{old}})}\right]_+$$

end

end

end

end

Table 1
OSTR algorithm outline

full iteration of OSTR-M. This intuition is verified in the initial iterations and for up to a reasonable number of subsets in the following results section.

IV. Results

We acquired a 15-hour blank scan (b_i's) and a 12-min transmission scan data (y_i's) using a Siemens/CTI ECAT EXACT 921 PET scanner with rotating rod sources for transmission scans. The phantom used was an anthropomorphic thorax phantom (Data Spectrum, Chapel Hill, NC). The projection space was 160 radial bins and 192 angles, and the reconstructed images were 128×128 with 4.5 mm. pixels. The system matrix a_{ij} was computed by using 6 mm. wide strip integrals with 3 mm. spacing, which roughly approximates the system geometry.

A. Reconstructions

The attenuation map is reconstructed for both ML and PL methods using OSTR algorithm with 1, 2, 4, 8, 16 and 32 subsets. Figure 1 shows objective function increase for the ML reconstructions initialized with a uniform image. The order-of-magnitude acceleration can be seen in this plot for initial iterations. One iteration of ML-OSTR-16 increases the objective almost as much as 16 iterations of ML-OSTR-1 and 4 iterations of ML-OSTR-4 for initial iterations. Although, when $M > 1$, the algorithm does not converge to the true ML solution, in practice one would only do a few iterations using ML-OSTR-M. In the ML problem, exact maximization is not desired since the maximizer image is exteremely noisy.

Figure 2 shows objective function increase versus iterations for PL reconstructions ($\beta = 2^{10}$ and nonquadratic Lange's penalty). The iterations are initialized with an FBP image. There is a speed-up in using more subsets, but as the number of subsets increase, the order-of-magnitude acceleration does not hold. For example, one iteration of PL-OSTR-16 increases the objective more than one iteration of PL-OSTR-32 (not shown). So, number of subsets greater than 16 did not seem to improve convergence for this configuration and data. For comparison, the image is also reconstructed with the paraboloidal surrogates coordinate ascent (PL-PSCA) method which is a fast monotonic algorithm [7]. The CPU times for one iteration of PL-PSCA and one iteration of PL-OSTR-1 are similar. It is clearly seen that PL-OSTR-M algorithms do not converge to the true maximum when $M > 1$. To use the speed-up property and converge at the end, one can sequentially decrease the number of subsets with each iteration. One such run is shown in the same figure where we start with 16 subsets and gradually decrease to 1 after which the algorithm should converge.

B. Mean Squared and Segmentation Errors

The reconstructions were done using real data. We wished to find mean squared errors and segmentation errors on the reconstructed images. The real image of course was not known to us. So, we acquired a long 14 hour scan of the thorax phantom which was almost noise free. We reconstructed the data with FBP with a sharp filter. Then, we performed a 4 level hard thresholding segmentation on this image with attenuation map parameters assumed to be average standard attenuation map values for air, soft tissue, lungs and bone. We obtained regions for each attenuation level. Then, we eroded these regions with a 5x5 window to get more conservative estimates of the regions and calculated the average value of the FBP image in these regions. These new values were assumed to be the true attenuation coefficient levels for the image (air = 0, lungs = 0.035, soft tissue = 0.093, bone = 0.164 cm^{-1}). Then, the FBP image was segmented by thresholding using new levels to obtain the "true" phantom image. Figure 3 shows the true phantom image.

We found normalized mean squared errors (NMSE) for each method using the true phantom image. The reconstructed images were also hard-segmented with the thresholds found above and we evaluated their segmentation performance by counting the number of misclassified pixels.

For this part, we also applied the emission ML-OSEM algorithm to logarithm of the transmission data $-log((y_i - r_i)/b_i$. Although, there is no theoretical basis for this approach, it is nevertheless used. The results show that this approach is inferior to the ML-OSTR method and that it should be avoided.

Figure 4 shows NMSE versus iterations for ML-OSTR, ML-OSEM, PL-OSTR and PL-PSCA methods. Figure 5 shows the percentage of segmentation errors versus iterations for the same methods. These results show that ML-OSTR algorithms get noisy after a certain number of iterations and that the iterations should be stopped before convergence. For example, it appears that the ML-OSTR-16 algorithm should be stopped at the third iteration for lowest NMSE. ML-OSEM applied to the logarithm of the transmission data is inferior in quality to all other methods we tried. PL reconstructions have better quality than ML reconstructions in terms of both lower mean squared errors and lower segmentation errors. Although, PL-OSTR-16 algorithm does not reach the true objective maximum, it appears to be comparable to PL-PSCA algorithm in terms of

mean squared error and segmentation performance.

Figure 6 shows reconstructed images and their segmentations for FBP, ML-OSTR, ML-OSEM, PL-OSTR and PL-PSCA methods. Each image is the best among their kind. For example, to obtain the FBP image, we performed 20 different FBPs with Hanning windows with different cutoff frequencies and picked the one with lowest NMSE. ML-OSTR image is obtained by 16 subsets at 3 iterations. ML-OSEM image is obtained by 8 subsets at 2 iterations. PL images are the images at 10^{th} iterations of their corresponding algorithms. The bars show the levels of NMSE and segmentation errors. We conclude that PL reconstruction images are much better than the images obtained using other methods.

V. Conclusion

We introduced a new ordered subsets algorithm for ML and PL image reconstruction in transmission tomography. Although the algorithm does not converge for number of subsets greater than one, it seems to rapidly increase the objective function value in the early iterations. The images reconstructed from real PET data with ML method are worse in quality than images reconstructed with PL method. However, ML-OSTR is superior to ML-OSEM applied to the logarithm of transmission data for this particular data. The new algorithm is easy to implement with any type of system model and does not require column access to the system matrix unlike sequential update algorithms such as coordinate ascent. It is also easily parallelizable.

We conclude that if an approximate maximum is acceptable due to practical time and programming constraints, then the OSTR algorithm offers faster convergence than prior methods. However, for guaranteed global convergence to the maximum, other methods must be used.

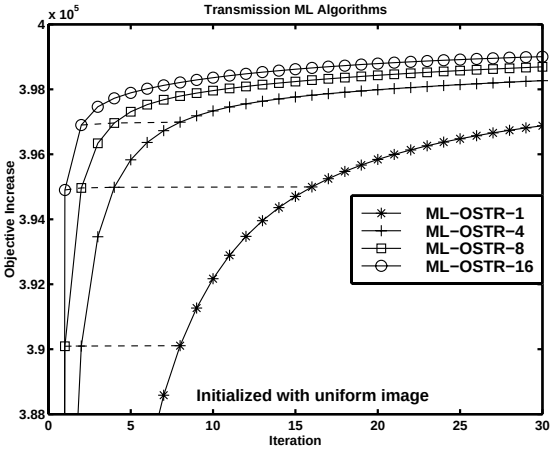

Figure 1: Maximum Likelihood $L(\mu^n) - L(\mu^0)$.

VI. References

[1] 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.

[2] J. A. Fessler, E. P. Ficaro, N. H. Clinthorne, and K. Lange, "Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction," *IEEE Tr. Med. Im.*, vol. 16, no. 2, pp. 166–75, April 1997.

[3] 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.

[4] J. M. Ollinger, "Maximum likelihood reconstruction of transmission images in emission computed tomography via the EM algorithm," *IEEE Tr. Med. Im.*, vol. 13, no. 1, pp. 89–101, March 1994.

[5] 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.

[6] S. H. Manglos, G. M. Gagne, A. Krol, F. D. Thomas, and R. Narayanaswamy, "Transmission maximum-likelihood reconstruction with ordered subsets for cone beam CT," *Phys. Med. Biol.*, vol. 40, no. 7, pp. 1225–41, July 1995.

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

[8] 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.

[9] J. A. Fessler. *ASPIRE: A Sparse, Precomputed, Iterative Reconstruction Library*, 1992.

[10] A. R. De Pierro, "On the relation between the ISRA and the EM algorithm for positron emission tomography," *IEEE Tr. Med. Im.*, vol. 12, no. 2, pp. 328–333, June 1993.

[11] A. R. De Pierro, "A modified expectation maximization

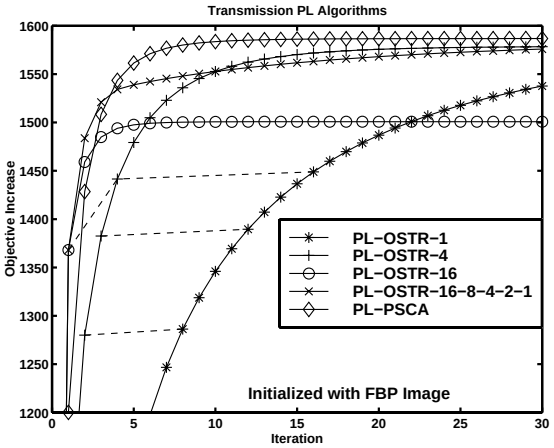

Figure 2: Penalized-likelihood $\Phi(\mu^n) - \Phi(\mu^0)$.

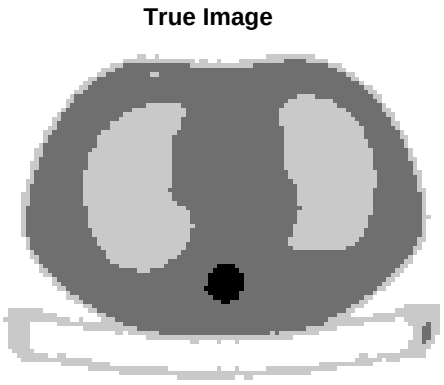

Figure 3: Image obtained by hard segmentation of the FBP image reconstructed from the 14-hour scan.

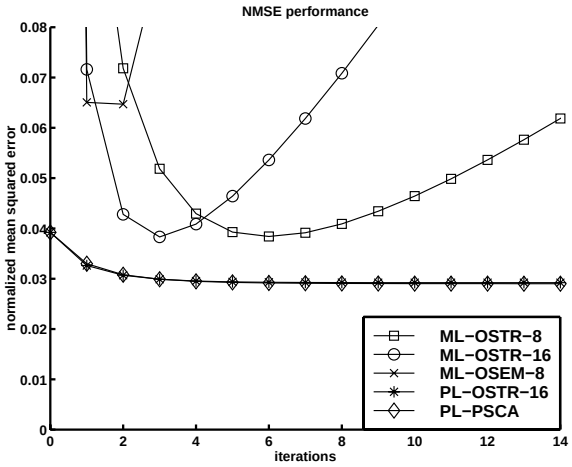

Figure 4: Normalized mean squared errors versus iterations for various methods of reconstruction.

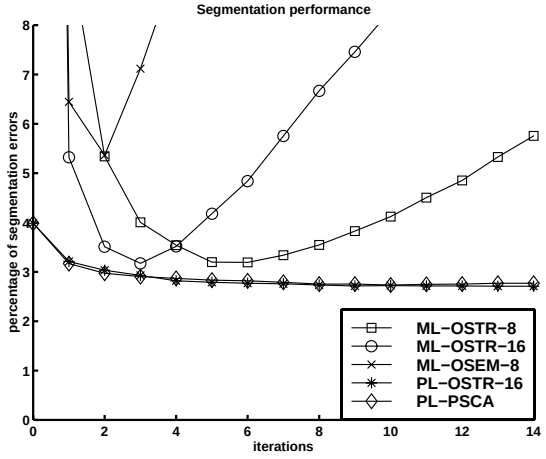

Figure 5: Segmentation errors versus iterations for various methods.

algorithm for penalized likelihood estimation in emission tomography," *IEEE Tr. Med. Im.*, vol. 14, no. 1, pp. 132–137, March 1995.

[12] 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.

[13] K. Lange, "Convergence of EM image reconstruction algorithms with Gibbs smoothing," *IEEE Tr. Med. Im.*, vol. 9, no. 4, pp. 439–446, December 1990. Corrections, June 1991.

[14] 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.

This work was supported in part by NIH grants CA-60711 and CA-54362, and by the Whitaker Foundation. First author was supported by TUBITAK-NATO Science Fellowship for doctoral studies.

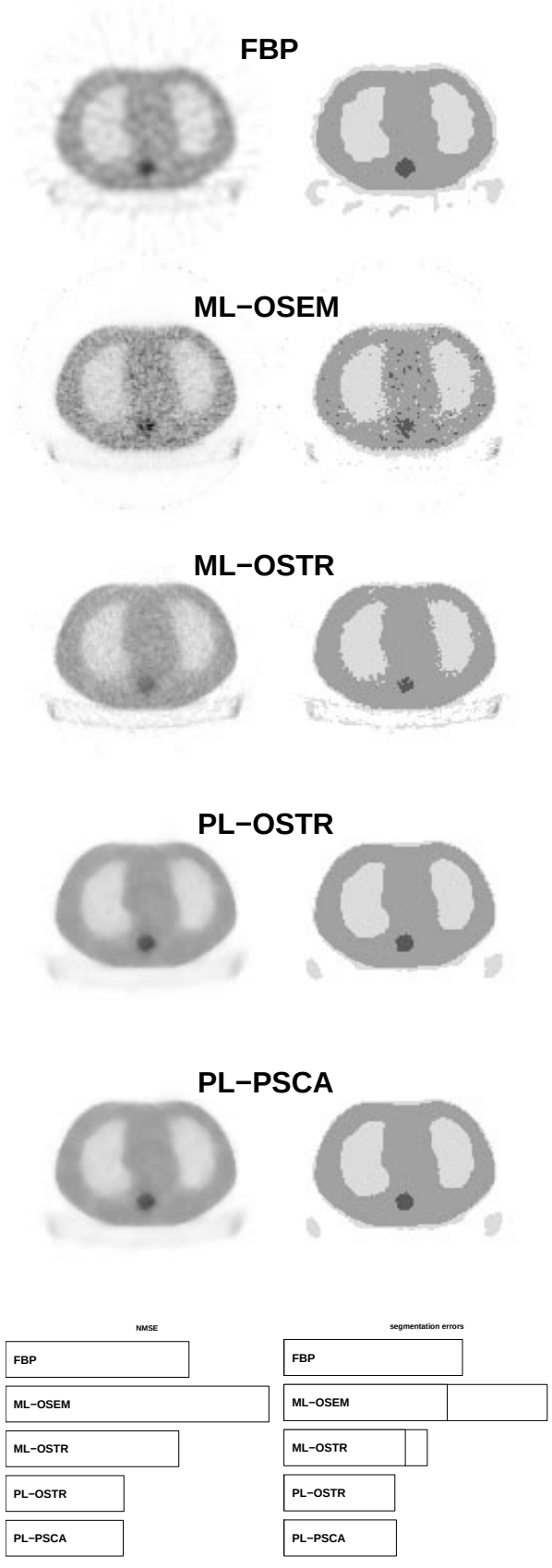

Figure 6: Reconstructed (left) and segmented (right) images using various methods. The bar plots show the relative NMSE and segmentation error levels. The middle lines in the right hand side bars for ML methods indicate the segmentation error reduction after median filtering.