Parallel Computation of Sequential Pixel Updates in Statistical Tomographic Reconstruction *†

Ken D. Sauer, Sean Borman
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556
(219) 631-6999

Charles A. Bouman
School of Electrical Engineering
Purdue University
West Lafayette, IN 47907-0501
(317) 494-0340

While Bayesian methods can significantly improve the quality of tomographic reconstructions, they require the solution of large iterative optimization problems. Recent results indicate that the convergence of these optimization problems can be improved by using sequential pixel updates, or Gauss-Seidel iterations.

However, Gauss-Seidel iterations may be perceived as less useful when parallel computing architectures are use. In this paper, we show that for degrees of parallelism of typical practical interest, the Gauss-Seidel iterations updates may be computed in parallel with little loss in convergence speed. In this case, the theoretical speed up of parallel implementations is nearly linear with the number of processors.

1 Introduction

Statistical methods of tomographic image reconstruction offer significant improvement in quality over deterministic approaches such as filtered backprojection (FBP), but entail far more computation for the resulting large-scale, iterative optimizations. Most optimization techniques for solving these problems are closely related to gradient ascent, with updates of all pixels' values computed in parallel at each iteration[1-3]. However, recent results have established that sequential updates of pixels, in the manner of Gauss-Seidel iterations for solving partial differential equations, converge very rapidly in the tomographic problem, and can trivially enforce constraints on the solution[4, 5]. The Gauss-Seidel method may be applied to Bayesian reconstruction in both emission and transmission cases [6]. While we use the name Gauss-Seidel for our approach to the optimization problem, iterative coordinate descent (ICD) and iterative conditional modes (ICM) are alternative names for the same method.

Because computation time is an important limiting factor in the industrial and clinical application of statistical reconstruction techniques, it is natural to consider parallel implementations. This is by definition at odds with the form of sequential updates. But it is intuitively clear that pixels which are spatially distant from one another influence each other minimally through the forward and backward projection. In this paper, we quantify this property, and use it to formulate a parallel implementation of Gauss-Seidel methods for tomographic reconstruction.

We include experimental results on synthetic phantoms which indicate that for practical levels of parallelization, pixels may be updated in parallel with little loss in convergence speed. In this case, the theoretical speed up of parallel implementations is nearly linear with the number of processors.

2 Parallel Computation of Sequential Updates

In this section, we present the form of provably convergent parallel pixel update computation. This method is similar to the technique employed by De Pierro to allow complete parallel updates for Bayesian tomographic emission reconstructions using a modified expectation-maxi- mization (EM) algorithm[3].

The photon counts Y which form the raw tomographic data are modeled as independent Poisson-dist- ributed random variables, dependent on projections $\sum_j A_{ij}x_j$, where x is the unknown cross section, and A is the projection matrix. The Poisson parameters depend on both input dosage and attenuation of the each ray in the transmission case, while the parameters are expressed directly by projections for the emission problem.

*This work was supported by the National Science Foundation under Grant no. MIP93-00560.

†Appeared in the Proceedings of *1995 IEEE International vol. 2, pp. 93-96, Conference on Image Processing*, Washington, DC, Oct. 22-25, 1995

DePierro has designed a method for fully parallel EM-type updates under a spatially connected *a priori* model for X. This method consists of replacing the logarithm of the prior density on X by an alternative cost function $C(x;x^n)$ at the $n+1$-th update, having the properties

$$C(x;x) = \log \mathcal{P}(X=x)$$
$$C(x;x^n) \leq \log \mathcal{P}(X=x).$$

Solving the optimization with $C(x;x^n)$ in place of $\log \mathcal{P}(X=x)$ guarantees that at each step, the log posterior density increases, since any increase in the former must result in an equal or greater increase in the latter.

With degrees of parallelism which are likely to be practical in our applications, and Markov random field (MRF) *a priori* image models, it is easy to choose parallel updates which are not coupled through $\log \mathcal{P}(X=x)$. We apply this approach to only the log likelihood function, $\log \mathcal{P}(Y=y|X=x)$, since through this function, essentially all pixel pairs are coupled.

The log likelihood for tomographic problems can be shown to have the form

$$\log \mathcal{P}(Y=y|X=x) = \sum_i -f_i\left(\sum_j A_{ij}x_j\right), \quad (1)$$

where $f_i(\cdot)$ is a convex function which depends on the data y. The first summation is over all entries in the data vector, while the second is over the image vector. Intuitively, the argument of $f_i(\cdot)$ is the i^{th} projection of the image x.

Suppose now that we consider the parallel update of a collection of pixels whose indeces form the set S at iteration $n+1$, with the remainder of the image fixed at x^n. We may then view the log likelihood at this step as a function of only $\{x_j; j \in S\}$. If we define

$$W_{S,i} = \sum_{k\in S} A_{ik},$$

then we may express the dependence on $\{x_j; j \in S\}$ by using the modified convex function $f_{Si}(\cdot)$.

$$\begin{aligned}\log \mathcal{P}(Y=y|X=x) & \\ &= -\sum_i f_{S,i}\left(\sum_{j\in S} A_{ij}x_j\right) \\ &= -\sum_i f_{S,i}\left(\sum_{j\in S}\frac{A_{ij}}{W_{S,i}}\left[\sum_{l\in S} A_{il}x_l^n + W_{S,i}(x_j - x_j^n)\right]\right).\end{aligned} \quad (2)$$

Applying Jensen's inequality results in the expression

$$-\log \mathcal{P}(Y=y|X=x) \leq \sum_i \sum_{j\in S}\frac{A_{ij}}{W_{S,i}} f_{S,i}\left(\sum_{l\in S} A_{il}x_l^n + W_{S,i}(x_j - x_j^n)\right) \quad (3)$$

This applies to common likelihood functions for the tomographic problems. Note that this is a summation over S, each term involving only one x_j, which allows for simple optimization.

If we define

$$p_{S,i}(x^n, x_j) = \sum_l A_{il}x_l^n + W_{S,i}(x_j - x_j^n)$$

and substitute into (3) for the standard Poisson models for transmission and emission tomography, the right-hand side of the inequality takes the form

$$\sum_i \sum_{j\in S}\frac{A_{ij}}{W_{S,i}}\left[p_{S,i}(x^n,x_j) - y_i \log(p_{S,i}(x^n,x_j))\right]$$

for emission and

$$\sum_i \sum_{j\in S}\frac{A_{ij}}{W_{S,i}}\left[y_T \exp\{-p_{S,i}(x^n,x_j)\} + y_i p_{S,i}(x^n,x_j)\right]$$

for transmission, where y_T is the dosage parameter.

The log likelihood functions in both emission and transmission cases can be approximated by a second-order Taylor series expansion in x

$$\log \mathcal{P}(Y=y|X=x) \approx -1/2(p-Ax)^t D(p-Ax) + c(y),$$

with p the vector of measured integral projections, $c(y)$ a constant relative to x, and D a diagonal matrix with entries being the photon counts $\{y_i\}$ (transmission) or $\{y_i^{-1}\}$ (emission)[6]. This approximation is quite accurate for most common transmission problems, and in both cases, lends both analytical leverage and qualitative understanding to studies of optimization techniques and their convergence behavior.

Under this approximation, we may apply the result of (3) to obtain the following.

$$\begin{aligned}&\sum_i f_{S,i}\left(\sum_{j\in S} A_{ij}x_j\right) \\ &\leq \sum_i \sum_{j\in S}\frac{A_{ij}D_{ii}}{W_{S,i}}(e_i^n - W_{S,i}\Delta_j^{n+1})^2.\end{aligned} \quad (4)$$

where

$$\begin{aligned} e_i^n &= p_i - \sum_j A_{ij}x_j^n \\ \Delta_j^{n+1} &= x_j - x_j^n \end{aligned}$$

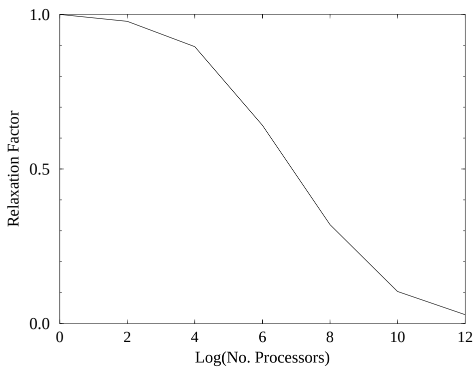

Figure 1: Value of the under-relaxation factor for the quadratic approximation of the log-likelihood. The number of processors ranges from one for the entire image on the left, to one for each pixel on the right. The latter case corresponds to completely parallel updates.

Here e_i^n is the error state vector in the projection data after iteration n, and Δ_j^{n+1} is the current change in pixel j. Minimization as a function of Δ_j^{n+1} yields

$$\Delta_j^{n+1} = \frac{\sum_i D_{ii} A_{ij} e_i^n}{\sum_i D_{ii} A_{ij} W_{S,i}} \tag{5}$$

This is the same form as the updates derived in [4], except that this formulation calls for under-relaxation of the greedy updates by the factor

$$\frac{\sum_i D_{ii} A_{ij}^2}{\sum_i D_{ii} A_{ij} W_{S,i}},$$

which reduces to the local update of [4] when S contains only one pixel. The value of this factor for the center pixel of a 128×128 reconstruction from 128×128 uniformly spaced projections is shown in Figure 1.

3 Numerical Results

Our first results consist of trials using a phantom with attenuation values similar to human tissue in a low dosage transmission tomography simulation. The data are 128×128 projections, and the reconstruction is computed at a resolution of 128×128 pixels. We solve the maximum *a posteriori* (MAP) reconstruction with the generalized Gaussian Markov random field (GGMRF)[7]. These prior densities add to the optimization the log prior $\sum_{i,j} \frac{b_{ij}}{q} \left(\frac{x_i - x_j}{\sigma} \right)^q$, which is convex for our choices of q.

The parallel computation assumes that after each processor has updated a pixel, the state of the projection error vector e^n can be updated and shared among all processors. We show results for a number but practical matters must also be addressed. Questions do remain on the effects, both short-term and asymtotic, of over/under-relaxation in this framework. The linear speed-up shown here is based on synchronous update of the projection state vector for all processors at the end of a cycle, which may in practice be limited by memory and communication speeds. Future work on this topic will include implementation of the Gauss-Seidel approach on parallel computing machinery.

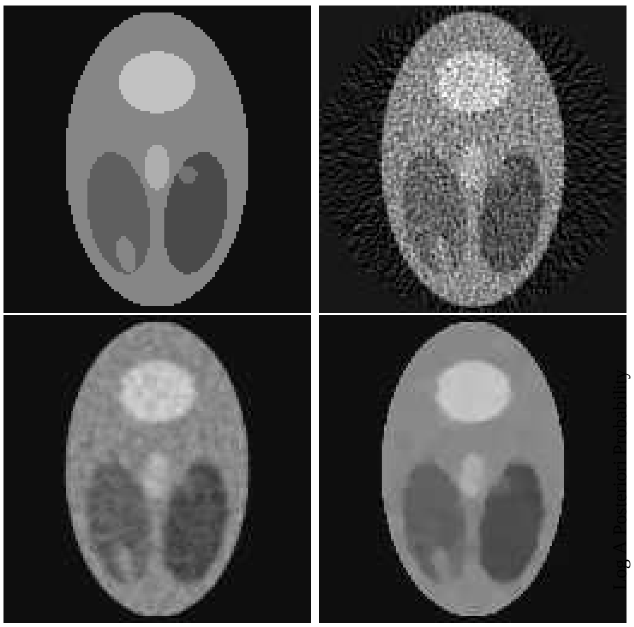

Figure 3: Above: Emission brain phantom, from which approximately 3×10^6 total photons are counted. Upper right: FBP reconstruction. Lower left: MAP estimate with Gaussian image model, $\sigma = 0.5$. Lower right: MAP estimate with GGMRF model, $q = 1.1$, $\sigma = 0.2$.

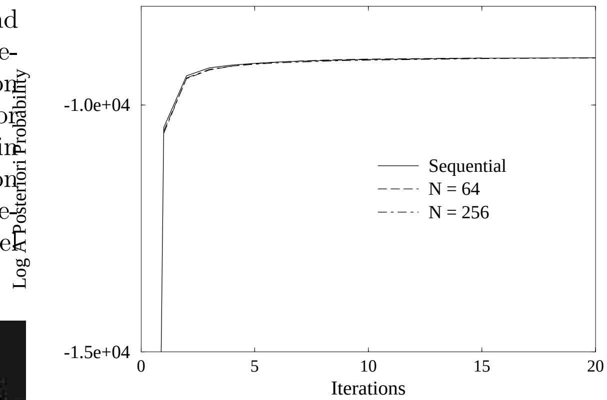

Figure 4: Convergence of parallel and purely sequentially computed pixel updates for emission phantom reconstruction under the Gaussian prior and no under-relaxation.

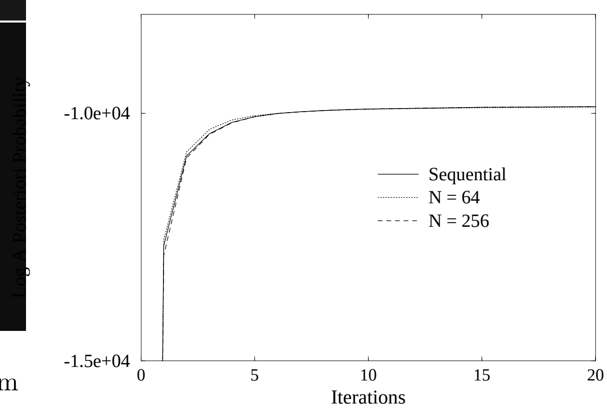

Figure 5: Convergence comparison with GGMRF prior and $q = 1.1.$, and no under-relaxation.

References

[1] T. Hebert and R. Leahy, "A Generalized EM Algorithm for 3-D Bayesian Reconstruction from Poisson data Using Gibbs Priors," *IEEE Trans. Med. Im.*, vol. 8, no. 2, pp. 194-202, June 1989.

[2] P. J. Green "Bayesian Reconstructions from Emission Tomography Data Using a Modified EM Algorithm," *IEEE Trans. Med. Im.*, vol. 9, no. 1, pp. 84-93, March 1990.

[3] A.R. De Pierro, "A Modified Expectation Maximization Algorithm for Penalized Likelihood Estimation in Emission Tomography," *IEEE Trans. on Med. Im.*, vol. 14, no. 1, pp. 132-137, March 1995.

[4] K. Sauer and C. Bouman, "A Local Update Strategy for Iterative Reconstruction from Projections," *IEEE Trans. on Sig. Proc.*, vol. 41, no. 2, pp. 533-548, Feb. 1993.

[5] J.A. Fessler, "Penalized Weighted Least-Squares Image Reconstruction for Positron Emission Tomography," *IEEE Trans. Med. Im.*, vol. 13, no. 2, pp. 290-300, June 1994.

[6] C. Bouman and K. Sauer, "Fast Numerical Methods for Emission and Transmission Tomographic Reconstruction," *Proceedings of the Twenty-seventh Annual Conference on Information Sciences and Systems*, pp. 611-616, The Johns Hopkins University, Baltimore, MD, March, 1993.

[7] C. Bouman and K. Sauer, "A Generalized Gaussian Image Model for Edge-Preserving MAP Estimation," *IEEE Trans. Image Proc.*, vol. 2, no. 3, pp. 296-310, July 1993.