Nonlinear multigrid methods of optimization in Bayesian tomographic image reconstruction

Charles Bouman
School of Electrical Engineering
Purdue University
West Lafayette, IN 47907

Ken Sauer
Laboratory for Image and Signal Analysis
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556

ABSTRACT

Bayesian estimation of transmission tomographic images presents formidable optimization tasks. Numerical solutions of this problem are limited in speed of convergence by the number of iterations required for the propagation of information across the grid. Edge-preserving prior models for tomographic images inject a nonlinear element into the Bayesian cost function, which limits the effectiveness of algorithms such as conjugate gradient, intended for linear problems. In this paper, we apply nonlinear multigrid optimization to Bayesian reconstruction of a two-dimensional function from integral projections. At each resolution, we apply Gauss-Seidel type iterations, which optimize locally with respect to individual pixel values. If the cost function is differentiable, the algorithm speeds convergence; if it is nonconvex and/or nondifferentiable, multigrid can yield improved estimates.

1. INTRODUCTION

Statistical methods of tomographic image reconstruction seek the solution which best matches the probabilistic behavior of the data. Maximum-likelihood (ML) estimation selects the reconstruction which most closely matches the data available, but may yield solutions which do not have many of the properties expected in the original function[1]. Maximum *a posteriori* (MAP) estimation allows the introduction of a prior distribution which reflects knowledge or beliefs concerning the types of images acceptable as estimates of the original cross-section. The Markov random field (MRF) has proven a useful image model in the tomographic setting, with relatively simple parameterization. Many variations from the common Gaussian MRF model have been proposed[2, 3, 4], with the aim of avoiding the excessive penalties extracted by the Gaussian's quadratic potential function, which tends to blur edges due to the high cost of abrupt transitions.

The generalized Gaussian MRF (GGMRF)[5] uses a potential function similar to the log of the generalized Gaussian noise density. It renders edges accurately without prior knowledge of their size, and it results in a convex optimization problem with no local minima. The application of these new MRFs as priors, however, poses difficult nonlinear optimization problems for MAP reconstruction. Instances of the GGMRF with local inter-pixel differences penalized by approximately the absolute value function appear to produce superior reconstructions, but at the cost of computational difficulty in finding optimal solutions. When applied to the transmission tomography problem, the absolute value potential function yields a nondifferentiable cost function for optimization.

The high computational cost of iterative reconstruction algorithms has limited their application to common tomographic reconstruction problems. Speeding convergence is essential in making statistical methods viable practical alternatives to deterministically-based algorithms. We have found that Gauss-Seidel(GS) type iterations move efficiently toward the MAP estimate by sequentially updating individual pixels[6]. But the convergence of GS is

Appeared in Proc. of *SPIE Conf. on Neural and Stochastic Methods in Image and Signal Processing*, vol. 1766, pp. 296-306, San Diego, California, July 19-24, 1992.

best at high spatial frequencies. Also, for the non-differentiable problem, the GS iterations may become "stuck" at non-differentiable points of the cost function. These problems can be addressed by including treatment of lower frequency information explicitly on lower resolution versions of the reconstruction. On coarser grids, spatial frequencies are effectively shifted upward, where they respond more favorably to the convergence patterns of GS iterations. Below, we explore the application of non-linear multigrid methods to the transmission tomographic estimation problem, in conjunction with GS iterations at each resolution. The multigrid allows much of the low frequency portion of the solution to be estimated on coarse grids, at reduced computational cost. For differentiable cost functions, this reduces numbers of iterations to convergence, while the improved initial condition for fine grid iterations produces qualitative and quantitative improvements in the non-differentiable case. In simulations, we apply multigrid optimization to a data set from a realistic synthetic tomographic phantom.

2. Nonlinear Multigrid

Solutions of physical problems commonly take the form

$$N(f) = 0.$$

where f is a function of one or more dimensions, and N is a nonlinear (or linear) functional. For example, all differential and integral equations have this form.

Such problems must be discretized to be solved numerically. To accomplish this the domain of f (normally $\mathbb{R}^n$) is approximated by a discrete grid of points. We will assume that this grid has spacing h, and the discrete approximation to f and N are denoted by $f^{(h)}$ and $N^{(h)}$ respectively. The discrete problem is then

$$N^{(h)}(f^{(h)}) = 0.$$

In practice, such discrete problems are often computationally difficult to solve when the grid spacing h is small. This is a result of the large number of points in the grid and the large amount of time required for information to propagate across the grid.

Multigrid techniques have been developed to address this difficulty since they combine the advantages of both fine and coarse grid spacings. The multigrid approach may be best understood by first considering the two grid problem. Imagine that a computationally efficient method exists to get a solution to the discrete problem with twice the grid spacing

$$N^{(2h)}(f^{(2h)}) = 0.$$

We can then use this coarse grid solution to more efficiently compute an *exact* fine grid solution. To do this we must first define linear matrices $I_{(h)}^{(2h)}$ and $I_{(2h)}^{(h)}$ for interpolating and decimating between different grid spacings.

$$I_{(2h)}^{(h)} f^{(2h)} = f^{(h)}$$

$$I_{(h)}^{(2h)} f^{(h)} = f^{(2h)}$$

The two grid solution is then computed in three steps:

1. Compute an approximate solution, $\tilde{f}^{(h)}$, to

$$N^{(h)}(f^{(h)}) = 0.$$

2. Compute the coarse grid solution, $\tilde{f}^{(2h)}$, to

$$N^{(2h)}(f^{(2h)}) - N^{(2h)}(I_{(h)}^{(2h)} \tilde{f}^{(h)}) = -I_{(h)}^{(2h)} N^{(h)}(\tilde{f}^{(h)}).$$

3. Compute an improved fine grid solution solution, $\hat{f}^{(h)}$

$$\hat{f}^{(h)} = \tilde{f}^{(h)} + I_{(2h)}^{(h)}(\tilde{f}^{(2h)} - I_{(h)}^{(2h)}\tilde{f}^{(h)})$$

4. Improve $\hat{f}^{(h)}$ using fine grid iterations

Notice that if the initial fine grid solution is correct, then the coarse grid corrections will be zero. This is important since solutions at coarser grids will be progressively worse approximations to the original problem. Therefore, coarse grid corrections should not perturb fine grid calculations when the initial answer is correct. This same strategy can also be used to compute the solution at the coarse grid spacing, $2h$. This forms a recursive algorithm which solves the fine grid problem by moving up and down through grid spacings.

3. Application to Transmission Tomographic Reconstruction

3.1. Formulation of Optimization Problem

Our objective is to apply the nonlinear multigrid techniques to the solution of the discrete tomographic reconstruction problem, where the discretized model for integral projections is expressed as

$$p = Af.$$

The vector p contains data from the forward projection of the image f, modeled by the operator A. The cost function resulting from formulating the problem as a MAP estimation, using the GGMRF as prior, and a quadratic approximation of the log conditional data density as a function of pixel values[6] is

$$C(f) = (p - Af)^t \mathbf{D}(p - Af) + \sum_{\{s,r\} \in U} \alpha_{sr} \lambda^q |f_s - f_r|^q. \tag{1}$$

The diagonal matrix D contains photon counts for each projection measurement, weighting more heavily those projections resulting from higher counts and therefore lower integral densities and lower noise. U is the set of all neighboring pairs of pixels. The GGMRF is modeled after the generalized Gaussian noise distribution, common in robust detection[7]. The parameters λ and q specify the weighting and character of the GGMRF prior, respectively. For $q = 2$, we have a Gaussian prior, which discourages sharp edges, while values of q nearer 1 do not. The value λ is inversely related to variance in the prior model, and corresponds to weighting of regularization. For the sake of compact notation and generality in the multigrid formulation, we temporarily replace the log prior density above with the function $\rho()$ of differences among neighboring pixels. We describe the case of a 4-pixel neighborhood, which can easily be generalized.

The problem we wish to solve can then be written as

$$\min_f \left\{ \sum_s \rho([\Delta_x f]_s) + \rho([\Delta_y f]_s) \quad + ||p - Af||_D^2 \right\} \tag{2}$$

where Δ_x and Δ_y are the forward difference operators

$$\begin{aligned} [\Delta_x f]_s &= f_{(s_1+1,s_2)} - f_{(s_1,s_2)} \\ [\Delta_y f]_s &= f_{(s_1,s_2+1)} - f_{(s_1,s_2)} \end{aligned}$$

and $\rho(\cdot)$ is some twice differentiable function. By differentiating (2) we find that the solution must also meet the following condition for all points $s = (s_1, s_2)$

$$0 = \rho'([\Delta_x f]_{(s_1-1,s_2)}) - \rho'([\Delta_x f]_{(s_1,s_2)}) + \rho'([\Delta_y f]_{(s_1,s_2-1)}) - \rho'([\Delta_y f]_{(s_1,s_2)}) + \frac{\partial}{\partial f_s}||p - Af||_D^2$$

where $\rho'(\cdot)$ is the derivative of $\rho(\cdot)$. The first part of this equation suggests a discrete approximation to a nonlinear differential operator. Since this operator is on a grid size with unit spacing, we denote it by $R^{(1)}(\cdot)$, and define it by the equivalent expression

$$R^{(1)}(f) = -\Delta_x^* \rho'(\Delta_x f) - \Delta_y^* \rho'(\Delta_y f) \tag{3}$$

where Δ_x^* and Δ_y^* are the backward difference operators

$$\begin{aligned} [\Delta_x^* f]_s &= f_{(s_1,s_2)} - f_{(s_1-1,s_2)} \\ [\Delta_y^* f]_s &= f_{(s_1,s_2)} - f_{(s_1,s_2-1)} \, . \end{aligned}$$

If we further define the linear operator and constant

$$\begin{aligned} L^{(1)} f &= 2A^t DAf \\ P^{(1)} &= 2A^t Dp \, , \end{aligned}$$

this results in the vector equation

$$R^{(1)}(f) + L^{(1)} f = P^{(1)} \, . \tag{4}$$

Note that this equation uniquely defines the solution if the functions $\rho(\cdot)$ are strictly convex. We will use both the optimization formulation of (2), and the equation formulation of (4) to solve the tomographic reconstruction problem. The optimization view is often useful when performing computation since it lends to algorithms which have guaranteed convergence properties. However, the equation formulation is more suitable to analysis using the tools of differential equations (i.e. frequency analysis and multigrid methods).

The MAP estimation problem is intrinsically discrete. However, multigrid techniques are based on the assumption that the discrete operators can be scaled at different resolutions. The basic assumption used in scaling differential operators is that the function is sufficiently smooth so that

$$I_{(1)}^{(2)} \Delta_x f^{(1)} \approx \frac{1}{2} \Delta_x I_{(1)}^{(2)} f^{(1)} \tag{5}$$

Using this approximation, we have that

$$\begin{aligned} R^{(k)}(f) &= -\frac{1}{2^k} \Delta_x^* \rho' \left(\frac{1}{2^k} \Delta_x f \right) - \frac{1}{2^k} \Delta_y^* \rho' \left(\frac{1}{2^k} \Delta_y f \right) \\ &= \frac{1}{2} R^{(k-1)} \left(\frac{1}{2} f \right) \end{aligned} \tag{6}$$

The discrete sum and constant terms scale somewhat differently.

$$\begin{aligned} L^{(k)}(f) &= 2 I_{(1)}^{(k)} A^t DA I_{(k)}^{(1)} f \\ P^{(k)} &= 2 I_{(1)}^{(k)} A^t Dp \end{aligned}$$

where

$$\begin{aligned} I_{(1)}^{(k)} &= \prod_{i=1}^{k-1} I_{(i)}^{(i+1)} \\ I_{(k)}^{(1)} &= \prod_{i=k-1}^{1} I_{(i+1)}^{(i)} \end{aligned}$$

A critical step in the multigrid approach is the solution of the coarse grid problem. Without loss of generality we assume that the grid spacings are 1 and 2. The basic equation for the coarse grid solution, $f^{(2)}$ is then

$$N^{(2)}(f^{(2)}) - N^{(2)}(I_{(1)}^{(2)} \tilde{f}^{(1)}) = -I_{(1)}^{(2)} N^{(1)}(\tilde{f}^{(1)}). \tag{7}$$

where

$$N^{(2)}(f) = R^{(2)}(f) + L^{(2)}f - P^{(2)} .$$

Equation (7) may be rewritten as

$$R^{(2)}(f^{(2)}) + L^{(2)}f^{(2)} = R^{(2)}(I^{(2)}_{(1)}\tilde{f}^{(1)}) + L^{(2)}I^{(2)}_{(1)}\tilde{f}^{(1)} - I^{(2)}_{(1)}R^{(1)}(\tilde{f}^{(1)}) - I^{(2)}_{(1)}L^{(1)}\tilde{f}^{(1)} + P^{(2)}$$

which yields the final equation

$$R^{(2)}(f^{(2)}) + L^{(2)}(f^{(2)}) = r^{(2)} + 2I^{(2)}_{(1)}A^t Dp^{(2)}$$

where $p^{(2)}$ and $r^{(2)}$ are defined by

$$\begin{aligned} p^{(2)} &= p + A\left(I^{(1)}_{(2)}I^{(2)}_{(1)}\tilde{f}^{(1)} - \tilde{f}^{(1)}\right) \\ r^{(2)} &= R^{(2)}(I^{(2)}_{(1)}\tilde{f}^{(1)}) - I^{(2)}_{(1)}R^{(1)}(\tilde{f}^{(1)}) \end{aligned}$$

In order to formulate this coarse resolution problem as the solution to an optimization criteria, we will make one further assumption about the choice of the interpolating/decimating matrices. We will assume that

$$I^{(2)}_{(1)} = \frac{1}{4}(I^{(1)}_{(2)})^t .$$

This is true for most common choices of matrices used on a 2 dimensional grid. We may then define the coarse grid matrix

$$A^{(2)} = AI^{(1)}_{(2)}$$

which results in the equation

$$4R^{(2)}(f^{(2)}) + 2(A^{(2)})^t DA^{(2)}f^{(2)} = 4r^{(2)} + 2(A^{(2)})^t Dp^{(2)}$$

Since this equations has a form very similar to the original fine grid solution, we know it is the dual to the optimization problem

$$\min_f \left\{ \sum_s 4\rho\left(\frac{1}{2}\lceil\Delta_x f\rceil_s\right) + 4\rho\left(\frac{1}{2}\lceil\Delta_y f\rceil_s\right) \quad - 4r^{(2)}f + ||p^{(2)} - A^{(2)}f||^2_D \right\} \tag{8}$$

This optimization problem can be computed efficiently using the Gauss-Seidel techniques previously developed. We assess computational costs at each level of resolution according to the expense of projecting and backprojecting data. This is because the computation is dominated by operations similar to projection and backprojection used in the updating of projection errors in the Gauss-Seidel algorithm. When reducing resolution by a factor of 2 in each variable, the number of pixels to be visited is reduced by a factor of four. But the average pixel will have twice as many intersecting rays to access, and the computational gain at level $k+1$ over level k is 2.

3.2. Recursive Formulation

The multigrid solution may be computed using a recursive subroutine call. In each case the main program has the form

main()
$f = 0$
$r = 0$
Repeat until converged
    MultigridV($f, r, p, k = 1$)

We use repeated applications of a modified version of the multigrid V algorithm[8]. This modification causes optimization to occur only during the coarse-to-fine operations. We iteratively repeat the MultigridV algorithm, until the solution is sufficiently converged.

MultigridV(f, r, p, k)

1. If k is the coarsest desired grid, return.
2. Compute the following:

$$\hat{f} = f_{int} = I_{(k)}^{(k+1)} f \tag{9}$$

$$\hat{p} = p + A^{(k)} \left(I_{(k+1)}^{(k)} I_{(k)}^{(k+1)} f - f \right) \tag{10}$$

$$\hat{r} = R^{(k+1)}(f_{int}) - I_{(k)}^{(k+1)} R^{(k)}(f) \tag{11}$$

3. MultigridV($\hat{f}, \hat{r}, \hat{p}, k+1$)
4. Perform coarse grid correction

$$f \leftarrow f + I_{(k+1)}^{(k)} (\hat{f} - f_{int})$$

5. Use the initial condition, f, to approximately solve the optimization

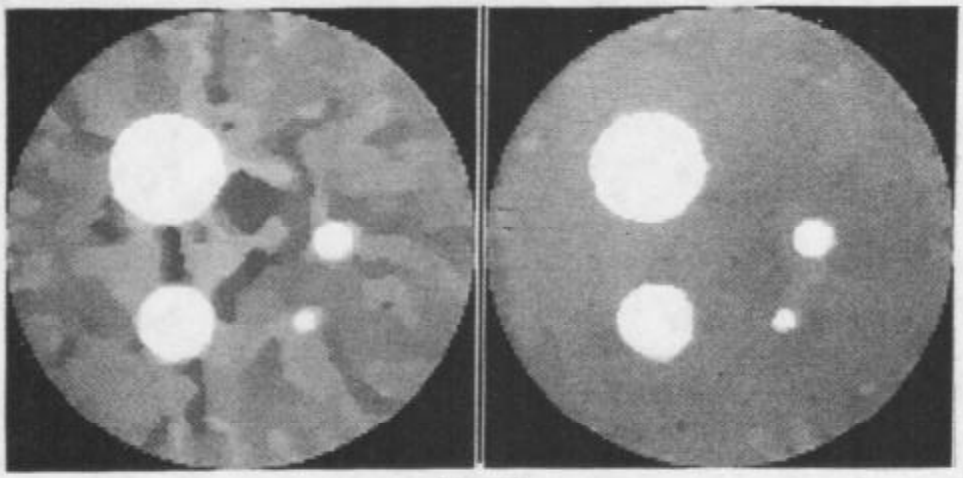
$$f \leftarrow \arg\min_f \left\{ \sum_s 4^k \rho \left(\frac{1}{2^k} [\Delta_x f]_s \right) + 4^k \rho \left(\frac{1}{2^k} [\Delta_y f]_s \right) \right. \quad \left. - 4rf + ||p - A^{(k)} f||_D^2 \right\}$$

Other variations are possible[8], and the form of the recursive structure may be adapted to specific optimization problems, and spatial frequency content in the initial error of the approximation.

4. SIMULATIONS

We have applied multigrid optimization techniques discussed in this paper to problems of tomographic reconstruction under limited dosage X-rays. In these cases, high noise levels in rays passing through very dense regions of the object under study yield artifacts corrupting the entire reconstruction. The data set was taken from the phantom in Fig. 1(a), which consists of four discs of density $0.48 cm^{-1}$ in a circular background of diameter 20 cm and density $0.2 cm^{-1}$. The simulated dosage was a rate of 2000 photons per ray (λ_T). At these dosages, the rays passing through the most dense regions are essentially blocked, making this reconstruction similar to a hollow projections problem. The 128 angles were equally spaced, and 128 rays were collected at each angle. Fig. 1(b) shows the CBP reconstruction from filtered backprojection, using a raised cosine rolloff filter. Though only the denser areas are too absorptive to render a good reconstruction, the entire image is of poor quality.

Improved reconstructions are possible using Bayesian estimation techniques. The character of the solution and the rate of convergence of GS iterations in MAP estimation depend heavily upon the prior model used. We apply the GGRMF as a probability density for the image ensemble, creating an optimization problem of the type given in (1). For the Gaussian case of the GGMRF ($q = 2$), we have a linear problem, and the advantage of a great deal of existing optimization theory and methodology. When $q < 2$, the solution is a nonlinear function of the data, and analytical tools are less developed. For the lower limit of the model of $q = 1$, the prior term is nondifferentiable in states where neighboring pixels are identically-valued, which leads to peculiar problems in optimization, in spite of the strict convexity of the cost function. Increasing the amount of regularization applied, by increasing λ, tends to speed convergence for all values of q.



a | b

Figure 6: (a) Reconstruction for $q = 1.0$ after 20 iterations by simple GS(left); (b) after equivalent cost of 20 iterations in multigrid optimization (right).

data correction up the pyramid. Though the difference in the objective function value in Fig. 7 is again small, image quality is significantly improved by multigrid.

5. CONCLUSION

Multigrid optimization offers improved convergence, and in some cases better solutions to Bayesian tomographic image estimation. In this paper, we have applied the method with the GGMRF as prior. Because for any $q > 1$, the resulting functional is convex and differentiable, most algorithms differ only in their rate of convergence to the unique local minimum cost approximation. The multigrid methods are also of interest in nonconvex and nondifferentiable problems. Efficient solution of MAP estimation with the absolute value potential function requires further investigation, including alternative correction terms for coarser grids.

6. ACKNOWLEDGEMENTS

This research was supported by an NEC Faculty Fellowship.

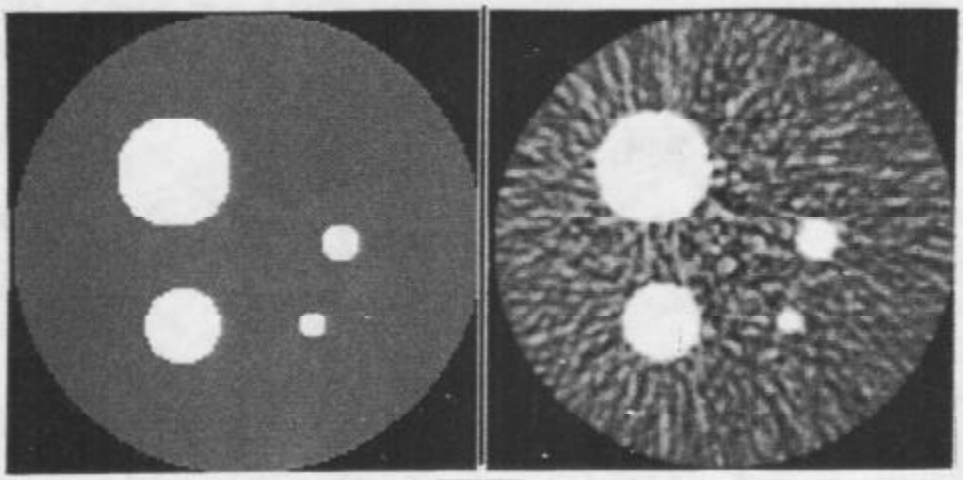
7. REFERENCES

1. D.L. Snyder and M.T. Miller, "The Use of Sieves to Stabilize Images Produced with the EM Algorithm for Emission Tomography," *IEEE Trans. Nucl. Sci.*, vol. NS-32, pp. 3864-3872, 1985.

2. S. Geman and D. McClure, "Bayesian Image Analysis: An Application to Single Photon Emission Tomography," in *Proc. Statist. Comput. Sect. Amer. Stat. Assoc.*, Washington, DC, pp. 12-18, 1985.

3. 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.

4. 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.



a | b

Figure 1: a) Original phantom (left); b) convolution backprojection reconstruction in low photon dosage with 128 projections at each of 128 angles (right); all synthetic phantom images are presented at a resolution of 128 x 128 pixels.

We implement a form of the multigrid which starts computation at the coarsest resolution and used repeated one-way passes down the multigrid pyramid (from coarse to fine). Note, though, that in returning from the fine grid to the coarse to commence another descent, we pass correction terms up between each pair of adjacent levels. Each of the cases presented includes 4 levels in the pyramid, an initial condition of $f = 0$, and $\lambda = 10$. The MAP estimate for $q = 1.5$ appears in Fig. 2. GS has essentially completely converged after fewer than 10 iterations at full resolution in this problem, but gain is available in performing early low frequency estimates at coarser resolutions. In Fig. 3, we plot the value of the objective function, the log *a posteriori* density. The plot for multigrid begins at the point when iterations first reach the finest level. Computational costs of coarse grid iterations are accumulated between visits to the finest, and registered in iteration counts proportionally to their relation to fine grid costs. Multigrid here improves convergence by about two iterations for equal values of log *a posteriori* density. This behavior is typical of values of q between approximately 1.3 and 2.0, where the MAP estimate is relatively smooth for the given parameters.

As q nears 1.0, the derivative of the function $\rho()$ in the GGMRF approaches a discontinuous function about the origin. The prior density's term adds to the cost function an element which increasingly resembles a convex polytope rather than a smooth function. This causes increasing costs for movement in individual pixel's values when they are near those of neighbors, with pronounced effects in convergence. In addition to rapidly achieving a first approximation on the finest grid, multigrid maintains a advantage over simple GS in terms of both objective function value, and image quality. The objective information appears in Fig. 5 for $q = 1.05$, while the reconstructions of Fig. 4 show appreciable difference in the solution after 20 iterations. Although on any level of the multigrid pyramid GS converges monotonically, we cannot guarantee monotonic convergence of the entire scheme. Corrections on coarse grids can be guaranteed to be zero if the solution is reached, but small perturbations may cause misrepresentation of the problem at coarse resolution. For these cases, the smoothness assumption used in arriving at (5) does not hold, but multigrid still offers improved convergence.

At $q = 1$, GS cannot be guaranteed to reach the MAP estimate, since despite the convexity of (1), a non-optimal solution may be optimal with respect to any single pixel's value relative to the current state. Fig. 6(a) shows the case of the GS estimate's being "stuck" in a nondifferentiable state. Figure 6(b) illustrates the effects of multigrid with GS, achieving a substantial advantage in image quality. Because the cost function is nondifferentiable for this prior density, we have removed the correction term corresponding to the prior, r, and passed only the projection

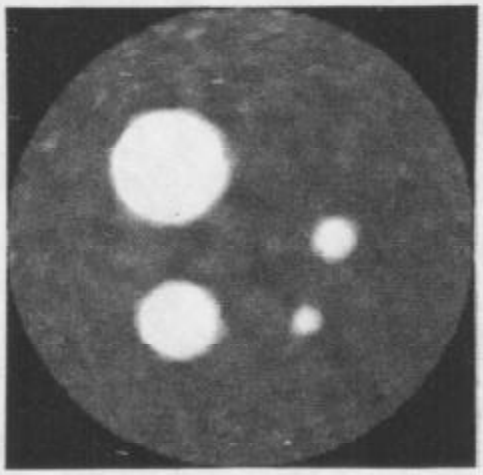


Figure 2: Reconstruction for $q = 1.5$ after 20 iterations by simple GS(left); the multigrid result at this point was not distinguishable from this image.

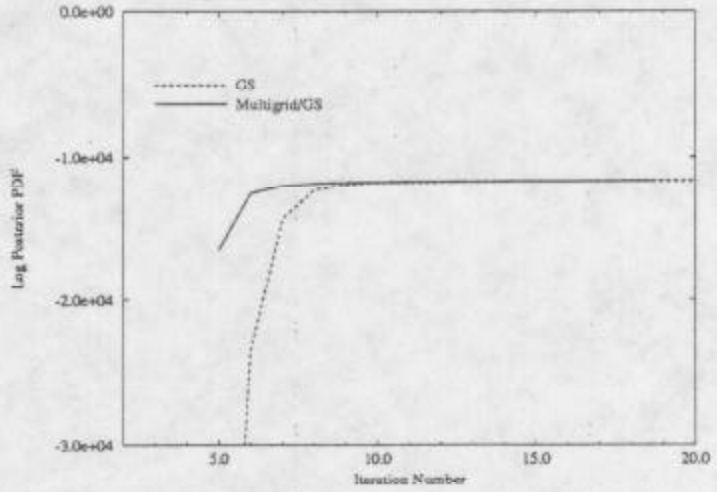


Figure 3: Convergence in terms of log *a posteriori* density for $q = 1.5$.

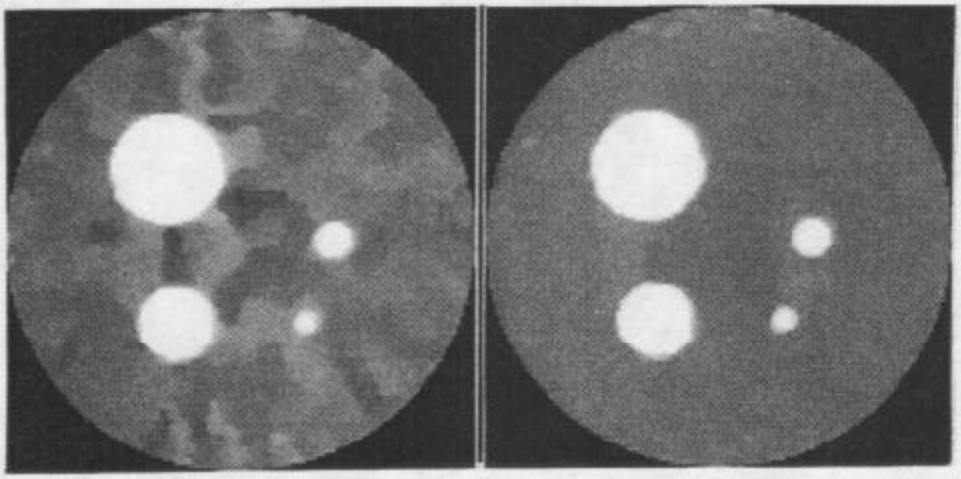


Figure 4: (a) Reconstruction for $q = 1.05$ after 20 iterations by simple GS(left); (b) after equivalent cost of 20 iterations in multigrid optimization (right).

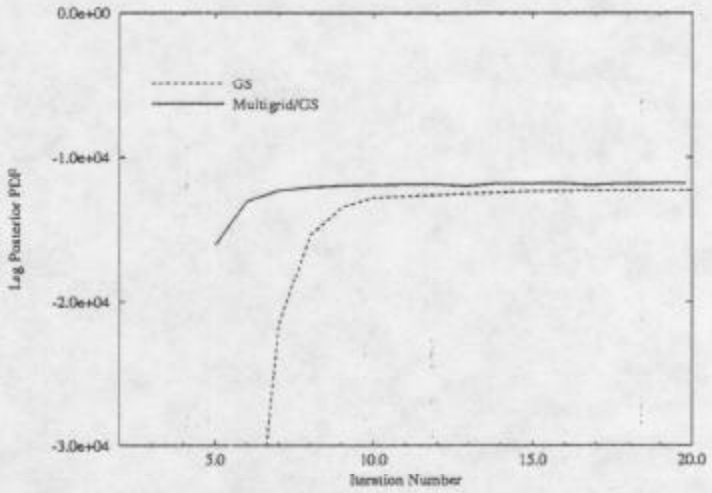
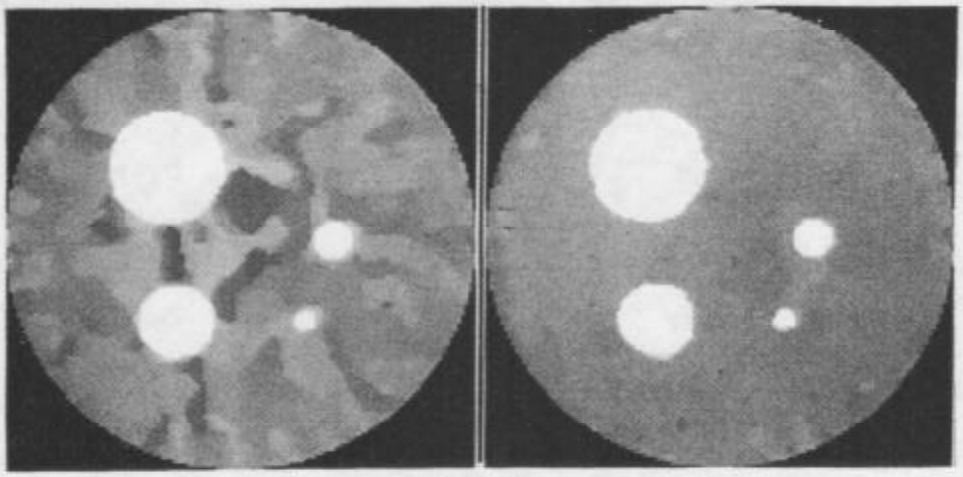


Figure 5: Convergence in terms of log *a posteriori* density for $q = 1.05$.



a | b

Figure 6: (a) Reconstruction for $q = 1.0$ after 20 iterations by simple GS(left); (b) after equivalent cost of 20 iterations in multigrid optimization (right).

data correction up the pyramid. Though the difference in the objective function value in Fig. 7 is again small, image quality is significantly improved by multigrid.

5. CONCLUSION

Multigrid optimization offers improved convergence, and in some cases better solutions to Bayesian tomographic image estimation. In this paper, we have applied the method with the GGMRF as prior. Because for any $q > 1$, the resulting functional is convex and differentiable, most algorithms differ only in their rate of convergence to the unique local minimum cost approximation. The multigrid methods are also of interest in nonconvex and nondifferentiable problems. Efficient solution of MAP estimation with the absolute value potential function requires further investigation, including alternative correction terms for coarser grids.

6. ACKNOWLEDGEMENTS

This research was supported by an NEC Faculty Fellowship.

7. REFERENCES

1. D.L. Snyder and M.T. Miller, "The Use of Sieves to Stabilize Images Produced with the EM Algorithm for Emission Tomography," *IEEE Trans. Nucl. Sci.*, vol. NS-32, pp. 3864-3872, 1985.

2. S. Geman and D. McClure, "Bayesian Image Analysis: An Application to Single Photon Emission Tomography," in *Proc. Statist. Comput. Sect. Amer. Stat. Assoc.*, Washington, DC, pp. 12-18, 1985.

3. 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.

4. 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.

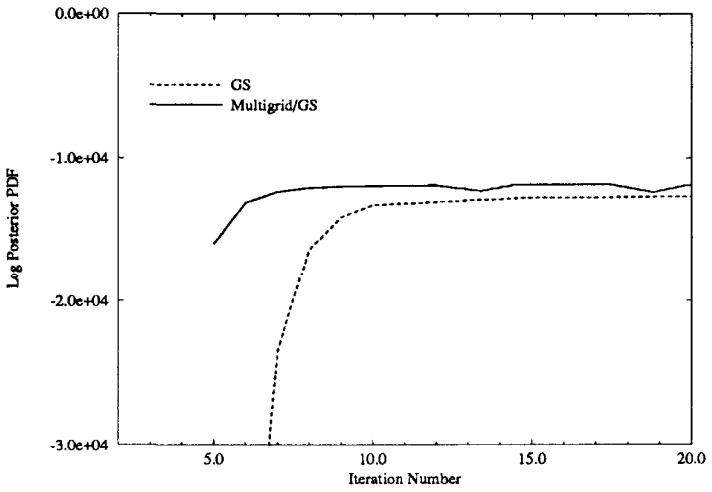


Figure 7: Convergence in terms of log *a posteriori* density for $q = 1.0$

5. C. Bouman and K. Sauer, "An Edge-Preserving Method for Image Reconstruction from Integral Projections," *Proc. Conf. on Info. Sci. and Syst.*, The Johns Hopkins University, Baltimore, MD, March 20-22, 1991, pp. 382-387.

6. K. Sauer and C. Bouman, "A Local Update Strategy for Iterative Reconstruction from Projections," submitted to the *IEEE Trans. Sig. Proc.*.

7. S.A. Kassam, *Signal Detection in Non-Gaussian Noise*, Springer-Verlag, New York, 1988.

8. W. L. Briggs, *A Multigrid Tutorial*, SIAM, Philadaelphia, PA, 1987.