A FAST ITERATIVE TOMOGRAPHIC RECONSTRUCTION ALGORITHM

Alexander H. Delaney and Yoram Bresler

Coordinated Science Laboratory
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
1308 W. Main St. Urbana, IL 61801
E-mail: ybresler@uiuc.edu

ABSTRACT

We use a series-expansion approach and an operator framework to derive a new, fast and accurate, iterative tomographic reconstruction algorithm applicable for parallel-ray projections that have been collected at a finite number of arbitrary view angles and have been radially sampled at a rate high enough so that aliasing errors are small. We use the conjugate gradient algorithm to minimize a regularized least squares criterion, and we prove that the main step in each iteration is equivalent to a 2-D discrete convolution, which can be cheaply and exactly implemented via the FFT. The proposed algorithm requires $\mathcal{O}(N^2 \log N)$ multiplies per iteration to reconstruct an $N \times N$ image from P view angles, and requires the storage of half of a $2N \times 2N$ PSF

1. INTRODUCTION

In many applications of computerized tomography, such as in electron microscopy, astronomy, geophysical exploration, nondestructive evaluation, and others, it is not possible to collect a complete set of projections. Applying standard transform-based reconstruction algorithms, such as filtered backprojection (FBP) [1] to such limited-data problems results in poor reconstructions with severe artifacts. Iterative tomographic reconstruction algorithms, on the other hand, are capable of producing high-quality images from limited data since they can explicitly take into account the missing data, and they are able to incorporate *a priori* information about the solution to regularize the problem. The major drawback of these iterative algorithms is their computational cost, with just one iteration typically requiring more computation than is needed by the entire FBP algorithm.

For example, a common approach to iterative tomographic reconstruction is to formulate the problem as a large system of linear equations. That is, the reconstructed image is modeled as a weighted sum of shifted basis functions (usually square pixels), and the vector of discrete projection data $\mathbf{g}$ is then related to these weights via a large matrix equation $\mathbf{g} \approx \mathbf{A}\hat{\mathbf{f}}$, where $\mathbf{A}$ is the projection matrix and $\hat{\mathbf{f}}$ is the vector of weights. An optimization problem incorporating *a priori* information about the solution (and possibly the noise) is then established from this data model, and any of a number of iterative techniques is used to solve the problem. Typically, each iteration will require one or more matrix-vector multiplication involving $\mathbf{A}$ or $\mathbf{A}^T$. If we have $M \times P$ raysums (the result of collecting M samples from each of P projections), and we wish to reconstruct an image that consists of $N \times N$ basis functions, then $\mathbf{A}$ will be a $PM \times N^2$ matrix. Although $\mathbf{A}$ can be made sparse by using *local* basis functions (such as square pixels) to represent the reconstructed image, the matrix will still have about $2PN^2$ non-zero elements. Hence, $\mathcal{O}(PN^2)$ multiplies are needed for one iteration of a typical matrix-based iterative algorithm. (For comparison, the total number of multiplies required for the standard FBP algorithm is about PN^2). In addition the $2PN^2$ non-zero elements of the projection matrix $\mathbf{A}$ must be stored.

In [2] Medoff uses an operator framework to derive an iterative convolution backprojection algorithm that does not require the storage of a large projection matrix. Each iteration requires a forward reprojection of the estimate image, followed by a filtered-backprojection step. Unfortunately, the ideal reprojection and filtered-backprojection steps must be approximated, introducing errors into the algorithm that cause it to diverge if run too long. In addition, both the reprojection and filtered-backprojection steps require about PN^2 multiplies to compute.

Fourier-based iterative tomographic reconstruction algorithms (eg., [3]), are less general than their algebraic counterparts but are computationally less costly per iteration. These algorithms essentially convert the tomography problem to one of spectral extrapolation by assuming that parallel-ray projections are available over a continuous range of radial offsets and (usually) angles. Each iteration typically consists mainly of a 2-D filtering operation, which can be performed via the FFT, and requires much less computation than most other iterative algorithms ($\mathcal{O}(N^2 \log N)$ multiplies per iteration). In addition, there is no need to store a large projection matrix. Unfortunately previous Fourier-based algorithms have not properly modeled the discrete nature of the data. Besides approximating the continuous Fourier transform by the discrete Fourier transform, the projections are typically assumed to be available over a continuous range of angles, instead of at a discrete set of angles.

This work was supported in part by a National Science Foundation Grant MIP 91-57377 and by a Whitaker Foundation grant to A. Belmont

0-7803-2431-5/95 $4.00

In this paper we use a series-expansion approach [1] and an operator framework to derive a new, fast and accurate, iterative tomographic reconstruction algorithm that is applicable for parallel-ray projections that have been collected at a finite number of arbitrary view angles and have been radially sampled at a rate high enough so that aliasing errors are small. We use the conjugate gradient algorithm to minimize a regularized least-squares criterion, and we prove that the main step in each iteration is equivalent to a 2-D discrete convolution, which can be cheaply and exactly implemented via the FFT.

2. A REGULARIZED SERIES-EXPANSION FORMULATION OF TOMOGRAPHY

Our goal is to reconstruct an image $\hat{f} \in L^2(R^2)$ that is an estimate of $f \in L^2(R^2)$ from a set of noisy, sampled, bandlimited parallel-ray projections of f collected at a finite number of view angles, θ_p, $p = 0, 1, \cdots, P-1$. Using a series-expansion approach [1], we represent the estimate image $\hat{f}$ by a linear combination of shifted versions of a single generating function $b \in L^2(R^2)$. More specifically, we require that $\hat{f} = \mathcal{Q}\hat{f}_D$, where $\hat{f}_D \in R^{N\times N}$ is a discrete 2-D image, and the interpolation operator $\mathcal{Q} : R^{N\times N} \rightarrow L^2(R^2)$ is defined by

$$\mathcal{Q}\hat{f}_D(x,y) = \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} \hat{f}_D(m,n) b(x-m, y-n). \quad (1)$$

We model the set of projections $g_D \in \left\{\ell^2(Z)\right\}^P$ as $g_D \approx \mathcal{A}\hat{f}_D$, where the linear operator $\mathcal{A} : R^{N\times N} \rightarrow \left\{\ell^2(Z)\right\}^P$ is defined as $\mathcal{A} = \mathcal{S}_T\mathcal{R}_A\mathcal{Q}$. Here, $\mathcal{R}_A : L^2(R^2) \rightarrow \left\{L^2(R)\right\}^P$ is the discrete-angle Radon transform operator defined by

$$\mathcal{R}_A f(r,p) = \int f(r\cos\theta_p - t\sin\theta_p, r\sin\theta_p + t\cos\theta_p)\, dt, \quad (2)$$

and $\mathcal{S}_T : \left\{L^2(R)\right\}^P \rightarrow \left\{\ell^2(Z)\right\}^P$ is the radial generalized sampling operator defined by

$$\mathcal{S}_T g_A(m,p) = \int g_A(r,p)\,\psi(r-mT)\, dr, \quad (3)$$

where $\psi \in L^2(R^2)$ is the detector response function, $r \in R$, $\theta \in [0,\pi]^P$, $p = 0, 1, \cdots, P-1$, $m \in Z$, and T is the sampling period.

We then seek a solution to the following regularized, least-squares optimization problem:

$$\min_{\hat{f}_D}\left\{\frac{1}{2}\|g_D - \mathcal{A}\hat{f}_D\|^2_{\mathcal{W}} + \sum_i \lambda_i C_i \hat{f}_D\right\}, \quad (4)$$

where $C_i : R^{N\times N} \rightarrow R$ and $\lambda_i \in R$ are the i_{th} regularization functional and regularization parameter, respectively, and $\|\cdot\|_{\mathcal{W}}$ is a weighted norm defined by $\|g_D\|^2_{\mathcal{W}} = \sum_{p=0}^{P-1}\sum_{m=-\infty}^{\infty} g_D(m,p)\,\mathcal{W}g_D(m,p)$ with the positive definite spectral weighting operator $\mathcal{W} : \left\{\ell^2(Z)\right\}^P \rightarrow \left\{\ell^2(Z)\right\}^P$ being defined by

$$\mathcal{F}_D\mathcal{W}g_D(u,p) = W_p(u)\mathcal{F}_D g_D(u,p),$$

where $W_p(u) > 0$ for all $|u| \leq \pi$, and $p = 0, 1, \cdots, P-1$, and $\mathcal{F}_D : \left\{\ell^2(Z)\right\}^P \rightarrow \left\{\ell^2(Z)\right\}^P$ is the Discrete-Time Fourier transform (DTFT) operator by

$$\mathcal{F}_D g_D(u,p) = \sum_{m=-\infty}^{+\infty} g_D(m,p)\exp[-jum].$$

3. A NEW, EFFICIENT IMPLEMENTATION

Using the conjugate gradient algorithm to find an approximate solution to (4), it is easy to show that the main computational cost of finding f_D^k, the k-th iterate of the conjugate gradient algorithm, is in computing $\mathcal{A}^*\mathcal{W}\mathcal{A}f_D^{k-1}$ [4], where the adjoint of $\mathcal{A}$ is $\mathcal{A}^* = \mathcal{Q}^*\mathcal{R}_A^*\mathcal{S}_T^*$. It is straightforward to show that: $\mathcal{Q}^* : L^2(R^2) \rightarrow R^{N\times N}$ is defined by

$$\begin{aligned}\mathcal{Q}^* f(m,n) &= \int\int f(x,y) b(x-m, y-n)\, dx\, dy \\ &= \tilde{b} \star\star f(m,n), \end{aligned} \quad (5)$$

where $\tilde{b}(x,y) = b(-x,-y)$, and $\tilde{b} \star\star f$ represents the two-dimensional, continuous-index convolution of $\tilde{b}$ with f. Also, $\mathcal{R}_A^* = \mathcal{B}_A : \left\{L^2(R)\right\}^P \rightarrow L^2(R^2)$, the discrete-angle back-projection operator, is defined by

$$\mathcal{B}_A g_A(x,y) = \sum_{p=0}^{P-1} g_A(x\cos\theta_p + y\sin\theta_p, p), \quad (6)$$

and $\mathcal{S}_T^* : \left\{\ell^2(Z)\right\}^P \rightarrow \left\{L^2(R)\right\}^P$, is defined by

$$\mathcal{S}_T^* g_D(r,p) = \sum_{m=-\infty}^{\infty} g_D(m,p)\,\psi(r-mT).$$

We now prove that if the projections have been bandlimited before sampling so that no aliasing occurs, then we can exactly compute $\mathcal{A}^*\mathcal{W}\mathcal{A}f_D$ via a 2-D discrete convolution, and need only store half of a $2N \times 2N$ point spread function (PSF).

In what follows, the standard 1-D and 2-D Fourier transform operators are denoted by $\mathcal{F}$ and $\mathcal{F}_2$, respectively. In addition, we define the Fourier transform operator $\mathcal{F}_A$: $\left\{L^2(R)\right\}^P \rightarrow \left\{L^2(R)\right\}^P$ by

$$\mathcal{F}_A g_A(\omega,p) = \int g_A(x,p)\exp[-j\omega x]\, dx.$$

One-dimensional, continuous- and discrete-index convolution will be denoted by $d \star q$ and $f * h$, respectively, whereas two-dimensional, continuous- and discrete-index convolution will be denoted by $d \star\star q$ and $f ** h$, respectively. Furthermore, we define the one-dimensional continuous-index convolution between two functions $g_A, h_A \in \left\{L^2(R)\right\}^P$ by

$$g_A \tilde{\star} h_A(r,p) = \int g_A(t,p)\, h_A(r-t,p)\, dt.$$

In order to prove the main result we need the following two easily proved lemmas:

Lemma 1 (Shift lemma) *Let the shift operators* $\mathcal{Z}^{m,n} : L^2(R^2) \to L^2(R^2)$, *and* $\mathcal{Z}_A^{m,n} : \left\{L^2(R)\right\}^P \to \left\{L^2(R)\right\}^P$, *be defined by* $\mathcal{Z}^{m,n} f(x,y) = f(x-m, y-n)$ *and* $\mathcal{Z}_A^{m,n} h_A(r,p) = h_A(r-(m\cos\theta_p + n\sin\theta_p), p)$, *respectively, for* $m, n \in Z$. *Then (1)* $\mathcal{R}_A \mathcal{Z}^{m,n} = \mathcal{Z}_A^{m,n} \mathcal{R}_A$, *and (2)* $\mathcal{B}_A \mathcal{Z}_A^{m,n} = \mathcal{Z}^{m,n} \mathcal{B}_A$.

Proof: See [4].

Lemma 2 (Convolution lemma) *Let* $f \in L^2(R^2)$, $g_A \in \left\{L^2(R)\right\}^P$. *Then* $f \star\star \mathcal{B}_A g_A(x,y) = \mathcal{B}_A(g_A \tilde{\star} \mathcal{R}_A f)(x,y)$.

Proof: See [4].

Theorem 1 *If* $\psi \in L^2(R)$ *in (3) is chosen so that* $\mathcal{F}\psi(\omega) = 0$ *for* $|\omega| > \pi/T$, *then for any* $f_D \in R^{N\times N}$,

$$\mathcal{A}^*\mathcal{W}\mathcal{A} f_D(k,l) = \xi ** f_D(k,l), \tag{7}$$

for $0 \le k, l < N$, *where*

$$\xi(m,n) = \sum_{p=0}^{P-1} \rho_A(m\cos\theta_p + n\sin\theta_p, p), \tag{8}$$

and $\rho_A \in \left\{L^2(R)\right\}^P$ *is defined by*

$$\mathcal{F}_A \rho_A(\omega, p) = \frac{1}{T} W_p(\omega T) \left|\mathcal{F}\psi(\omega)\, \mathcal{F}_2 b(\omega\cos\theta_p, \omega\sin\theta_p)\right|^2. \tag{9}$$

Proof:

$$\begin{aligned} &\mathcal{A}^*\mathcal{W}\mathcal{A} f_D(k,l) \\ &= \mathcal{Q}^*\mathcal{B}_A\mathcal{S}_T^*\mathcal{W}\mathcal{S}_T\mathcal{R}_A\mathcal{Q} f_D(k,l) \\ &= \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} f_D(m,n)\, \eta(m,n,k,l), \end{aligned} \tag{10}$$

where

$$\eta(m,n,k,l) \triangleq \mathcal{Q}^*\mathcal{B}_A\mathcal{S}_T^*\mathcal{W}\mathcal{S}_T\mathcal{R}_A\mathcal{Z}^{m,n} b(k,l). \tag{11}$$

Our goal is to reduce (10) to a convolution, by showing that $\eta(m,n,k,l) = \xi(k-m, l-n)$.

Consider first the operator $\mathcal{M} \triangleq \mathcal{S}_T^*\mathcal{W}\mathcal{S}_T$, and note that $\mathcal{M}$ acting on any $g_A \in \left\{L^2(R)\right\}^P$ is equivalent to the following steps: 1) convolving each of the P 1-D continuous-index functions of g_A with ψ; 2) sampling each of the resulting P functions with a sampling period of T; 3) digitally filtering the p_{th} sampled function with W_p, for $p = 0, 1, \cdots, P-1$; and 4) reconstructing P continuous-index functions by interpolating each of the resulting P discrete-index functions with ψ. Hence, if ψ is chosen so that $\mathcal{F}\psi(\omega) = 0$ for $|\omega| \ge \pi/T$, so that there is no aliasing after the sampling in Step 2, then it follows from the Shannon sampling theorem that

$$\mathcal{F}_A \mathcal{M} g_A(\omega, p) = \frac{1}{T} W_p(\omega T) \left|\mathcal{F}\psi(\omega)\right|^2 \mathcal{F}_A g_A(\omega, p). \tag{12}$$

Note that $\mathcal{M}$ is a linear shift-invariant filter in the radial variable r.

Now, using (5), Lemma 1, and the fact that $\mathcal{M}\mathcal{Z}_A^{m,n} = \mathcal{Z}_A^{m,n}\mathcal{M}$, we have:

$$\begin{aligned} \eta(m,n,k,l) &= \tilde{b} \star\star \mathcal{B}_A \mathcal{M} \mathcal{R}_A \mathcal{Z}^{m,n} b(k,l) && (13)\\ &= \tilde{b} \star\star \mathcal{B}_A \mathcal{Z}_A^{m,n} \mathcal{M} \mathcal{R}_A b(k,l) && (14)\\ &= \tilde{b} \star\star \mathcal{Z}^{m,n} \mathcal{B}_A \mathcal{M} \mathcal{R}_A b(k,l) && (15)\\ &= \tilde{b} \star\star \mathcal{B}_A \mathcal{M} \mathcal{R}_A b(k-m, l-n) && (16)\\ &= \xi(k-m, l-n), && (17) \end{aligned}$$

where $\xi(k,l) = \tilde{b} \star\star \mathcal{B}_A \mathcal{M} \mathcal{R}_A b(k,l)$. Furthermore, from lemma 2 it follows that

$$\begin{aligned} \xi(k,l) &= \mathcal{B}_A((\mathcal{M}\mathcal{R}_A b) \tilde{\star} \mathcal{R}_A \tilde{b})(k,l) && (18)\\ &= \sum_{p=0}^{P-1} \rho_A(k\cos\theta_p + l\sin\theta_p, p), && (19) \end{aligned}$$

where $\rho_A(r,p) = (\mathcal{M}\mathcal{R}_A b) \tilde{\star} \mathcal{R}_A \tilde{b}$. Using the projection-slice theorem, $\mathcal{F}_A \mathcal{R}_A b(\omega, p) = \mathcal{F}_2 b(\omega\cos\theta_p, \omega\sin\theta_p)$, and $\mathcal{F}_A \mathcal{R}_A \tilde{b}(\omega, p) = \overline{\mathcal{F}_2 b}(\omega\cos\theta_p, \omega\sin\theta_p)$, where $\overline{\mathcal{F}_2 b}$ denotes the complex conjugate of $\mathcal{F}_2 b$, so that with (12) we obtain

$$\mathcal{F}_A \rho_A(\omega, p) = \frac{1}{T} W_p(\omega T) \left|\mathcal{F}\psi(\omega)\, \mathcal{F}_2 b(\omega\cos\theta_p, \omega\sin\theta_p)\right|^2. \tag{20}$$

□

Note that although the support of ξ will be over all of Z^2, we only need to store $\xi(k,l)$ for $0 \le k, l < 2N$ because we need to compute $\xi ** f_D(k,l)$ only for $0 \le k, l < N$, and $f_D(m,n)$ is defined only for $0 \le m, n < N$. Furthermore, because $\xi(k,l) = \xi(-k,-l)$, only half of the $2N \times 2N$ elements of ξ actually need to be stored. In addition, we can compute the convolution in (7) quickly and exactly by first zero padding f_D to be of size $2N \times 2N$, taking the 2-D FFT, multiplying the result pointwise with the 2-D FFT of ξ, taking the inverse 2-D FFT, and truncating the result to be of size $N \times N$. Using this method, the number of multiplies needed to compute $\mathcal{A}^*\mathcal{W}\mathcal{A}\hat{f}_D$, and hence (approximately) one iteration of the conjugate gradient algorithm, is about $24(1+\log N)N^2$, where N^2 is the number of elements in $\hat{f}_D$. Note that the computational cost of the proposed method is independent of the number of view angles and the choice of b used to represent $\hat{f}$ (cf. Eq. (1)), whereas in the standard algorithms the cost increases linearly with both the number of view angles and the radius of support of b.

In addition to calculating $\mathcal{A}^*\mathcal{W}\mathcal{A} f_D^k$ at every iteration, we must also calculate $q_D = \mathcal{A}^*\mathcal{W} g_D = \mathcal{Q}^*\mathcal{B}_A\mathcal{S}_T^*\mathcal{W} g_D$ once as part of the initialization of the conjugate gradient algorithm [4]. It is straightforward to show, in a manner similar to the proof of Theorem 1, that

$$q_D(m,n) = \sum_{p=0}^{P-1} \gamma_A(m\cos\theta_p + n\sin\theta_p, p), \tag{21}$$

where $\gamma_A \in \left\{L^2(R)\right\}^P$ is defined by

$$\begin{aligned} &\mathcal{F}_A \gamma_A(\omega, p) = \\ &\quad \mathcal{F}_D g_D(\omega T)\, W_p(\omega T)\, \overline{\mathcal{F}\psi}(\omega)\, \overline{\mathcal{F}_2 b}(\omega\cos\theta_p, \omega\sin\theta_p). \end{aligned} \tag{22}$$

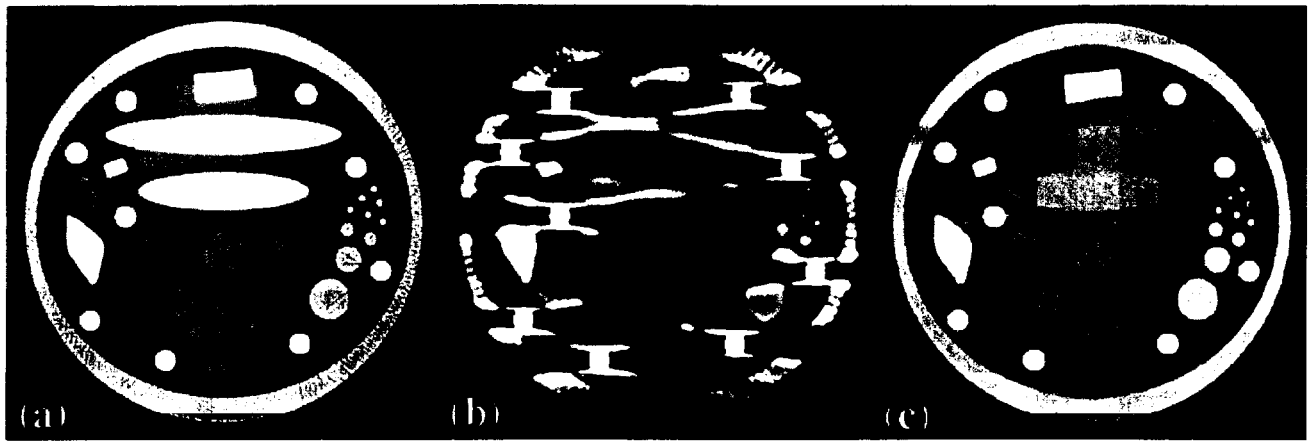

Figure 1: (a) Full-angle FBP reconstruction. (b) Limited-angle FBP reconstruction. (c): Limited-angle reconstruction using the proposed algorithm.

Note that both (8)-(9) and (21)-(22) are similar to the standard FBP algorithm, where the filtering now varies with projection angle.

Although explicit formulas for the evaluation of the PSF ξ and the discrete image q_D are given in (8)-(9) and (21)-(22), respectively, computing their exact values is usually not practicable. The difficulty arises in evaluating ρ_A or γ_A at the exact points required by (8) or (21), respectively. Fortunately, very close approximations to ξ and q_D can be easily calculated by using linear interpolation on finely sampled versions of ρ_A and γ_A, respectively. By creating a sampled version of $\mathcal{F}_A\rho_A$ (or $\mathcal{F}_A\gamma_A$), padding with zeros and then taking the inverse FFT, a close approximation to a sampled version of ρ_A (or γ_A) can be cheaply calculated. Using a large amount of zero padding and a small sampling period will increase the accuracy in computing ξ (or q_D) at the expense of increased computation. This increase in computation is small compared to the overall cost of the iterative algorithm, making it reasonable to create very accurate approximations to ξ and q_D. In addition, for many scenarios multiple reconstructions will be done using the same projection geometry, so that ξ need only be computed once and stored for future use.

4. NUMERICAL EXAMPLES

To demonstrate the effectiveness of the proposed algorithm we reconstruct a phantom from simulated parallel-ray projections collected by unity-spaced detectors that integrate evenly over a width of one. (Hence, the projections are slightly aliased.) The true detector response, ψ, was assumed to be unknown, and the ideal bandlimiting filter was used for ψ instead. The spectral weights W_p, for $p = 1, 2, \cdots, P-1$, (used in the weighted norm of (4)) were chosen to be

$$W_p(u) = \frac{1}{(2\pi)}|u|H(u), \quad p = 1, 2, \cdots, P-1,$$

where $H(u)$ is a Hanning window. The basis function b, used to represent the estimate image $\hat{f}$, was chosen to be the tensor product of the cubic B-spline, and the spacing between coefficients was normalized to one in each coordinate and was equal to the spacing between the radial samples of the parallel-ray projections. The displayed images are sampled on a 240×240 grid with unity spacing in each coordinate.

Figs. 1 (b) and (c) compare limited-angle reconstructions using the standard FBP algorithm and the proposed algorithm (after 80 iterations), for 112 noiseless projections, evenly spaced from -79^o to 79^o. A full-angle (256 projections evenly spaced over 180^o) FBP reconstruction is shown in Fig. 1 (a). The regularization functionals used in the iterative algorithm penalize image values that exceed a minimum and maximum value, as well as the sum of absolute image gradients. (Although for small gradients, the square of the gradient is used instead of the absolute value.) Note that the reconstruction from the proposed algorithm has none of the severe artifacts seen in the FBP reconstruction, and is very close to the full-angle reconstruction except for the blurring of the long, horizontal ellipses.

5. REFERENCES

[1] G. T. Herman, *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography*, New York: Academic Press, Inc., 1987.

[2] B. P. Medoff, "Image reconstruction from limited data: Theory and applications in computerized tomography," in *Image Recovery: Theory and Application*, (H. Stark, ed.), ch. 9, pp. 321-368, San Diego: Academic Press, Inc., 1987.

[3] M. I. Sezan and H. Stark, "Applications of Convex projection theory to image recovery in tomography and related areas," in *Image Recovery: Theory and Application*, (H. Stark, ed.), ch. 11, pp. 415-462, San Diego: Academic Press, Inc., 1987.

[4] A. H. Delaney and Y. Bresler, "A fast and accurate iterative reconstruction algorithm for parallel-beam tomography," Submitted to IEEE Trans. Image Processing, 1994.