A Multilevel Domain Decomposition Algorithm for Fast $O(N^2 \log N)$ Reprojection of Tomographic Images

Amir Boag, *Senior Member, IEEE*, Yoram Bresler, *Fellow, IEEE*, and Eric Michielssen, *Senior Member, IEEE*

***Abstract*—A novel algorithm for fast computation of tomographic image projections is presented. The method comprises a decomposition of an image into subimages followed by an aggregation of projections computed for the subimages. The multilevel domain decomposition algorithm is formulated as a recursive procedure. The computational cost of the proposed algorithm is comparable to that of FFT-based techniques while it appears to be more flexible than the latter. Numerical results demonstrate the effectiveness of the method.**

I. Introduction

THE reconstruction problem in two-dimensional (2-D) axial computed tomography (CT) is to recover an image from a set of its line-integral projections at different angles [1], [2]. The reverse operation, of computing a set of projections of a given 2-D image, is called reprojection, or computation of (angular samples of) the Radon transform. Reprojection is of interest in several applications. In X-ray CT, it is used [3] in iterative beam-hardening correction algorithms [4], in streak suppression algorithms [5], in algorithms for the removal of artifacts caused by the presence of metallic implants in the subject [6] or other high density objects [7], and in correcting for missing data [8], and partial volume effects [9]. In PET and SPECT imaging, reprojection has been used to compensate for attenuation [10]. Reprojection is also used in detection and compensation of various acquisition errors [11], [12] including those caused by patient motion [13], [14], and in accounting for Poisson noise statistics [15]. Reprojection is also a key component in iterative tomographic reconstruction algorithms, which offer various advantages over direct reconstruction methods in imaging modalities such as PET, SPECT, nondestructive testing, and more generally, in all situations of limited data (cf., [16], [17] and the references therein) and nonparallel-beam geometries [18]. Finally, reprojection is an important component in the efficient implementation of a wide variety of image processing algorithms [19], [20], including that for computing the well known Hough transform.

Direct algorithms for reprojection [12], [18], [21], [22] remain, in spite of various recent improvements, computationally very expensive, requiring $O(N^3)$ operations for generating N projections of an $N \times N$-pixel image. Likewise, the so-called "fast reprojection algorithms," including those based on backprojection and implemented on fast hardware backprojectors [23], and those implemented on other parallel architectures [19], [20], still require $O(N^3)$ operations. Their cost therefore scales as N^3, and they also offer little advantage in implementations on general-purpose computers. These various $O(N^3)$ algorithms, whether in software or hardware, are very expensive for typical state-of-the-art image sizes. For example, using one of these methods to compute the reprojection of a 4096×4096-pixel image requires $16^3 = 4096$ times the computation needed for a 256×256-pixel image.

There are, however, two families of reprojection algorithms with $O(N^2 \log N)$ cost. The first of these is the family of FFT-based algorithms, cf. [3], [24], and the references therein. These algorithms are based on the well-known Fourier Slice-Projection Theorem [1], [2] (see Property 3 in the next section). The primary problem with this approach is the required step of interpolation between the rectangular grid in Fourier space on which the transform of the image is computed using the FFT, and the polar grid on which the Fourier transform of the projections must be evaluated. The modern versions of these algorithms [24] overcome many of the associated difficulties, and offer good accuracy at a considerable speedup.

The second family of $O(N^2 \log N)$ reprojection algorithms [25]–[30] is based on a common principle. It uses a hierarchical decomposition of the line integral in a given direction into shorter line integral in the same direction called "segments" or "links." Adjacent segments are added up to create a segment of double length. Furthermore, line integrals at adjacent directions have some segments in common, or the segments of one such line integral can be accurately interpolated from those of the other. This reuse of segments in successive doubling steps accounts for the computational efficiency of these algorithms. Starting with segments corresponding to line integrals through two pixels, after $\log N$ such doubling steps, the complete line integrals are available. The earlier versions of these methods [25]–[28] used nearest-neighbor interpolation and computed the discrete Radon transform (partial sums of pixels whose center lies within a strip of predetermined width). However, versions that are more recent introduced linear [29] or

Manuscript received March 24, 1999; revised April 6, 2000. This work was supported in part by grants from AFOSR via the MURI program (F49620-96-1-0025) and by the NSF under Grant 95-02-138. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jeffrey A. Fessler.

A. Boag is with the Department of Electrical Engineering-Physical Electronics, Tel-Aviv University, Ramat-Aviv 69978, Tel-Aviv, Israel (e-mail: boag@eng.tau.ac.il).

Y. Bresler is with the Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: ybresler@uiuc.edu).

E. Michielssen is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: emichiel@uiuc.edu).

Publisher Item Identifier S 1057-7149(00)06977-3.

higher order [30] interpolation, providing better approximation to the continuous Radon transform. Some implementations of these methods have been demonstrated to provide considerable speedup over direct reprojection, with good accuracy. Related methods were proposed for the dual problem of backprojection, with nearest-neighbor interpolation in [31], linear interpolation in [32], [33], and linear or higher-order interpolation in [34].

In this paper, we present a different $O(N^2 \log N)$ algorithm based on multilevel domain decomposition and resembling our previous work in computational electromagnetics [35]. The method comprises a decomposition of an image into subimages followed by an aggregation of projections computed for the subimages. The multilevel computational sequence essentially reduces to the aggregation of projections of larger and larger images, starting with single pixel projections and ending with those of the whole image. The multilevel (M-L) algorithm attains computational complexity comparable to that of FFT-based techniques. The algorithm is formulated for a general image-discretization scheme using a series representation, of which the square pixel basis is a special case. Likewise, the formulation is for general interpolation schemes. In fact, both the accuracy and the computational burden involved are strongly dependent upon the specific interpolation schemes employed in the projection aggregation step. These issues are important to the efficiency of the method and will be considered elsewhere. For the sake of simplicity, here, we limit ourselves in the numerical experiments to the demonstration of the method using linear interpolation, and a cubic spline basis for the image.

The algorithm is developed in this paper to compute parallel-beam projections. Like the other fast algorithms, it can be used to obtain fan-beam projections by rebinning the parallel-beam projections produced by it.

We note that the method is most effective for the computation of one or more sets of projections at contiguous view angles (i.e., contiguous projections in one or more sectors). Thus, it will not be applicable to iterative reconstruction techniques that evaluate arbitrary subsets of projections at a time. However, it can still be applied in iterative techniques not only when all projections are evaluated simultaneously (as in gradient-based iterative algorithms), but also when one or more blocks of contiguous projections are evaluated at a time, or in limited view angle problems.

The proposed algorithm is compared to direct reprojection in computer simulations, confirming the $O(N^2 \log N)$ computational requirements. The results show good accuracy, both visually and numerically, with greater than 20-fold speedup for a 512×512 image. This compares favorably to speedups of about 10–15 reported for the FFT [24] and the length-doubling methods [33]. Because details of computer code and accuracy assessment can easily contribute factors of 2-3 to the speed of such algorithms, the merits of the new algorithm relative to these other families of fast algorithms will require careful further study, and may depend on the application.

II. Problem Statement

In this and the following sections, we use the term *essentially* in a well-defined sense when discussing the properties of various functions. A function is said to be essentially *supported* on a region G in its domain, if for all practical purposes its values outside G are negligible. For instance, all but a negligibly small fraction of its energy is contained in G, or its values outside G are bounded by some small number, and decay exponentially fast with distance from the boundary of G. For example, a time limited function can not be exactly bandlimited (i.e., have transform supported on an interval), but it can be essentially bandlimited (i.e., have transform essentially supported on an interval) [36].

Consider a 2-D image $f_c(x, y)$ vanishing outside the square domain $|x| \leq D/2, |y| \leq D/2$. We assume that the image can be approximated with sufficient accuracy by expansion with respect to a set of basis functions as

$$f_c(x, y) \approx \sum_{m,n=1}^{N} f(m, n) p\left[x - \left(m - \frac{N}{2} - \frac{1}{2}\right)\Delta, \; y - \left(n - \frac{N}{2} - \frac{1}{2}\right)\Delta\right] \quad (1)$$

where p is a pixel basis function and $\Delta = D/N$ is the discretization size. We refer to the $N \times N$ 2-D sequence $f(m, n)$ as the discrete version of $f_c(x, y)$, or simply as the discrete image. We assume that, in both x and y, the support of p is of $O(\Delta)$ in size and its essential bandwidth, denoted B, is of $O(\Delta^{-1})$. (Because p is spatially limited, it can only be essentially bandlimited.) For example, the Fourier transform of the size Δ square pixel has half-main-lobe width of Δ^{-1} along each of the frequency axes, or approximately bandwidth Δ^{-1}. Likewise, the tensor product of cubic B-splines used in our numerical experiments has support $4\Delta \times 4\Delta$ and bandwidth Δ^{-1}. Alternative choices of essentially bandlimited pixel-basis-functions for tomography are discussed in [37].

We consider a parallel-ray projection situation depicted in Fig. 1. The Radon transform of the image

$$\tilde{f}_c(r, \theta) = \Re f_c(r, \theta) = \int f_c(r \cos\theta - t \sin\theta, r \sin\theta + t \cos\theta)\, dt \quad (2)$$

produces parallel-ray projections of the image over a continuous range of angles $0 \leq \theta \leq \pi$. This set of projections is often called a sinogram. In many practical situations, the projections of f_c are measured only at a set of P angles $\{\theta_i\}_{i=1}^{P}$ (sometimes within a limited angular sector $|\theta| \leq \Theta$) and $O(N)$uniformly spaced radial points $\{r_k\}$. The required reprojection operation for a given image corresponds to the calculation specified by for all $\{\theta_i\}_{i=1}^{P}$ and $\{r_k\}$.

Let

$$\tilde{p}(r, \theta) = \Re p(r, \theta)$$

denote the projection of the pixel basis function. Clearly, it too has support of size $O(\Delta)$. Using the representation (1), it therefore follows that straightforward computation of $\tilde{f}(r_k, \theta_i)$ at one view angle and one radial position requires $O(N)$ operations. Thus, the operation count of the straightforward computation of projections at P view angles and N radial positions is of $O(PN^2)$, and for a typical case where $P = O(N)$, the com-

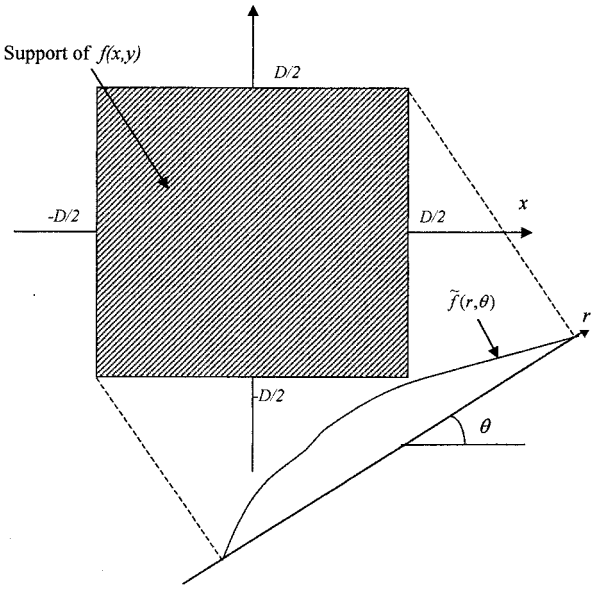


Fig. 1. Parallel ray projection.

putational cost is of $O(N^3)$. This unfavorable cost scaling limits the use of the algorithm for many applications, and is unacceptable in iterative image reconstruction (e.g., from limited data), which requires repeated reprojection of successive approximations of the image for angles $\{\theta_i\}$ for which the data was collected. Our goal is to develop an algorithm allowing reprojection over a given angular range at a significantly lower cost.

III. Multilevel Reprojection

As a preamble to the formulation of the algorithm, we present some observations regarding the properties of the Radon transform $\tilde{f}_c(r, \theta) = \Re f_c$ as a function of its two arguments. We will find these properties instrumental in the ensuing formulation of the proposed M-L algorithm.

The first two properties are straightforward to verify [1], [2] using the definition of the Radon transform.

Property 1: The Radon transform is a linear transformation.

Property 2: Shift Property of the Radon Transform: Let $f_{c,0}(x, y) = S_c[X_0, Y_0]f_c(x - x_0, y - y_0)$ denote an image shifted by the vector (x_0, y_0). Then its Radon transform is a radially shifted version of $\tilde{f}_c(r, \theta)$, i.e.,

$$\begin{aligned}\Re f_{c,0}(r, \theta) &= \tilde{S}_c[x_0, y_0]\tilde{f}_c(r, \theta) \\ &= \tilde{f}_c(r - x \sin\theta + y \cos\theta, \theta) \end{aligned} \tag{3}$$

where $\tilde{S}_c(x_0, y_0)$denotes a shift operator transforming projections to a new origin displaced by the vector (x_0, y_0) from the original one.

The next is a classical property of the Radon transform [1], [2].

Property 3: The Slice-Projection Theorem: Let $\tilde{F}_r(\omega_r, \theta)$ denote the Fourier transform of the projection $\tilde{f}(r, \theta)$ with respect to its first variable, and $F(\omega_x, \omega_y)$ the 2-D Fourier transform of the image $f_c(x, y)$. Then $\tilde{F}_r(\omega_r, \theta) = F(\omega_r \cos\theta, \omega_r \sin\theta)$, i.e., the radial Fourier transform of a projection at view angle θ is equal to a central slice at angle θ through the 2D Fourier transform of the image.

The fourth property follows from the so-called bowtie property [38], which addresses the essential support of the 2-D Fourier transform $\tilde{F}(\omega_r, \omega_\theta)$ of the sinogram $\tilde{f}_c(r, \theta)$ with respect to its two variables. [Note that because $\tilde{f}_c(r, \theta)$ is periodic with respect to θ, $\tilde{F}(\omega_r, \omega_\theta)$ will be a line spectrum in ω_θ, supported on integer values of ω_θ. Alternatively, a Fourier series with respect to θ can be used.] We use $\lfloor x \rfloor$ to denote the integer part of x.

Property 4: Angular Bandlimit of the Sinogram: If $\tilde{f}_c(r, \theta)$ is supported in $|r| < R$ and $\tilde{F}(\omega_r, \omega_\theta)$ essentially vanishes for $|\omega_r| > B_r$, then $\tilde{F}(\omega_r, \omega_\theta)$ essentially vanishes for $|\omega_\theta| > \lfloor RB_r \rfloor + 1$.

Rigorous estimates for the rate at which $\tilde{F}$ vanishes outside its essential support are given in [2]. These may be used to bound the error introduced by our use of Property 4, but such analysis is outside the scope of this paper.

The approximation of the image $f_c(x, y)$ employing basis functions bandlimited to a $B \times B$ square implies that f_c is also essentially bandlimited to a $B \times B$ square with respect to its variables. Hence, by the Slice-Projection Theorem (Property 3), the bandwidth of $\tilde{f}_c$ with respect to r is bounded by $\sqrt{2}B$. Now, clearly, if f_c is nonzero only in the domain $|x| \leq D/2$, $|y| \leq D/2$, the support of $\tilde{f}_c$ is limited to $|r| \leq D/\sqrt{2}$. Because $\tilde{f}_c$ is support limited and its transform essentially bandlimited in the radial variable, by Property 4 the bandwidth of $\tilde{f}_c$ with respect to θ is bounded by $\sqrt{2}BD/\sqrt{2} + 1 = DB + 1$. Next, because $\tilde{f}_c$ is periodic with respect to θ, it follows (by Fourier series expansion with respect to θ) that for each fixed $r = r_0$, $\tilde{f}_c(r_0, \theta)$ is a trigonometric polynomial in θ of order at most $DB + 1$. As such, it can be accurately interpolated from any set of its samples taken at $2DB + 2$ distinct angles [39]. Furthermore, because of the relationship $\tilde{f}_c(r, \theta) = \tilde{f}_c(-r, \theta + \pi)$, each projection provides two samples in the interval $[0, 2\pi)$[38]. Hence, $\tilde{f}_c(r, \theta)$ is fully determined by its values at any $DB + 1$ distinct view angles in $[0, \pi]$, and any additional projections can be interpolated from the projections at these view angles. Finally, recalling that $B = O(\Delta^{-1})$ so that $DB = O(D/\Delta) = O(N)$, we obtain the following.

Property 5: Number of Nonredundant Projections: The minimum number of nonredundant projections (and accordingly, the sampling rate) required to accurately interpolate $\tilde{f}_c(r, \theta)$ to arbitrary values $\{\theta_i\}$ within the range $|\theta| \leq \Theta$ is proportional to the image size N.

Straightforward reprojection would require cN^2P operations, where c is a constant. However, drawing on the above properties, we will show that the number of operations can be reduced by subdividing the image domain into a set of smaller subimages, computing their projections, and subsequently aggregating the results to obtain the desired projections of the whole image. In this two-level domain decomposition procedure, the image f is subdivided into M^2 subimages of size $(N/M) \times (N/M)$. The case of $M = 2$ is depicted in Fig. 2.

In the following derivation, it will be necessary to perform interpolations both in the angular and radial directions. To obtain the greatest speedup, these interpolations should be local, involving, say, L_θ and L_r terms for the angular and for the radial

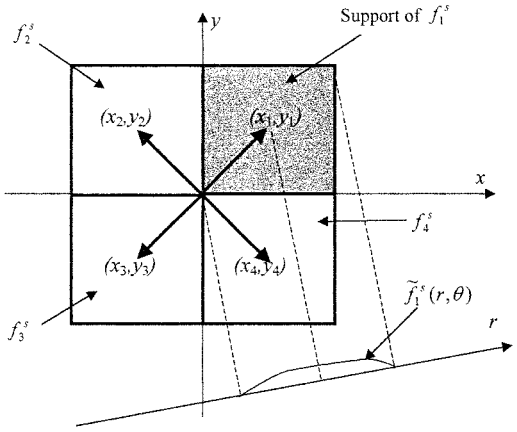


Fig. 2. Image and projection decomposition.

interpolations, respectively. In order to control the interpolation error with such local interpolation, we reduce the sampling intervals in θ and r by factors K_θ and K_r, respectively, from those required by the corresponding Nyquist rates. Thus, the algorithm is designed to produce $K_\theta P$ projections, each with up to $\sqrt{2}\ K_r N$ samples. Indeed, with oversampling by a factor $K > 1$, the interpolation error for a bandlimited signal decays exponentially, as $e^{-\pi L(k-1)/k}$, for an appropriately chosen interpolator of length L [40]. The amount of oversampling and the length of interpolators used will affect the tradeoff between computational cost and accuracy.

To derive the algorithm, we introduce the discrete-angle Radon transform $\Re_Q$, which produces projections with a continuous radial variable at Q view angles. We first describe how to compute this set of $Q = K_\theta P$ projections, $\tilde{f}_c = \Re_{K_\theta P} f_c$, using a two-level decomposition. Let $f_{c,i}^s$ denote a subimage of size $N\Delta/M$ with a local coordinate system at its center. Assume the center occupies the point (x_i, y_i)in the original image coordinates. It follows that the original image can be written as a sum of the M^2 shifted subimages

$$f_c = \sum_{i=1}^{M^2} S_c[x_i, y_i] f_{c,i}^s \tag{4}$$

where $S_c[x_i, y_i]$ denotes the image shift operator defined earlier. Applying $\Re_Q$ to both sides of, using the linearity Property 1 and shift Property 2, we obtain

$$\tilde{f}_c = \Re_{K_\theta P} f_c = \sum_{i=1}^{M^2} \tilde{S}_c[x_i, y_i]\, \Re_{K_\theta P} f_{c,i}^s \tag{5}$$

where $\tilde{S}_c[x_i, y_i]$ is the radial projection shift operator defined in Property 2, and $\Re_{K_\theta P} f_{c,i}^s$is the set of $K_\theta P$ projections of the subimage $f_{c,i}^s$, computed in *its own, local coordinate system*. Now, in view of Property 5, because the size of $f_{c,i}^s$ is $N\Delta/M$, $\Re_{K_\theta P} f_{c,i}^s$ can be interpolated from the projections $\tilde{f}_{c,i}^s = \Re_{K_\theta P/M} f_{c,i}^s$ of this subimage computed on a coarser angular grid with $K_\theta P/M$ view angles. This will maintain the same angular oversampling factor for the projections of f_i^s as for the entire image. Hence, we have

$$\Re_{K_\theta P} f_{c,i}^s = I_\theta^M \tilde{f}_{c,i}^s \tag{6}$$

where I_θ^M is the appropriate factor-M angular interpolation operator. Combining (5) and (6), we finally obtain

$$\tilde{f}_c = \sum_{i=1}^{M^2} \tilde{S}_c[x_i, y_i] I_\theta^M \tilde{f}_{c,i}^s. \tag{7}$$

In practice, each projection, which, as argued earlier, is bandlimited with respect to the radial variable r, is represented by its uniformly sampled version with respect to r. The goal of the algorithm is therefore to compute the set $\tilde{f}$ of $K_\theta P$ such sampled projections. The algorithm is then a mapping from the digital image f to the sampled projections $\tilde{f}$. Accordingly, applying radial sampling to both sides of (7) we obtain

$$\tilde{f} = \sum_{i=1}^{M^2} \tilde{S}[x_i, y_i] I_\theta^M \tilde{f}_i^s \tag{8}$$

where $\tilde{f}_i^s$are the radially sampled versions of $\tilde{f}_{c,i}^s$, and $\tilde{S}[a, b]$ denotes the discrete-to-discrete radial shift operator corresponding to $\tilde{S}_c[a, b]$. Because the shift may be by a noninteger multiple of the sampling interval, in general $\tilde{S}[a, b]$ involves interpolation with respect to r.

Equation (8) represents the entire two-level decomposition algorithm. The projections of each subimage are computed in its local coordinate system on an M-fold coarser angular grid, then interpolated to the original angular grid, radially shifted, and finally added up to produce the projections of the original image. The equalities in (8) hold essentially—to the extent that the out-of-band energy of the bandlimited signals can be ignored, and that exact interpolation is employed. As discussed earlier, accurate results can be obtained using local interpolation, by using appropriate oversampling.

Next, we evaluate the computational requirements of the proposed two-level algorithm. Each projection (with up to $\sqrt{2}K_r N/M$samples) of the N/M subimage f_i^s requires $cK_r(N/M)^2$ operations to compute, yielding a cost of $cK_\theta K_r P N^2/M^3$ operations for the set $\tilde{f}_i^s$ of $K_\theta P/M$ views. Next, the M-fold interpolation I_θ^M produces $K_\theta P$ projections with $\sqrt{2}K_r N/M$ samples each, at a cost of $c_\theta L_\theta K_\theta K_r PN/M$ operations, where L_θ is the length of the interpolator, and c_θ is a constant. The radial interpolation involved in $\tilde{S}(x_i, y_i)$ produces up to $\sqrt{2}K_r N/M$ samples each for the $K_\theta P$ projections in $I_\theta^M \tilde{f}_i^s$, at a cost of $c_r L_r K_\theta K_r PN/M$ operations, where L_r is the length of the radial interpolator, and c_r a constant. Finally, the addition of the shifted projections of each subimage requires $K_\theta K_r PN/M$ operations. The total cost of computing the projections becomes

$$C_1 = K_\theta K_r M^2 [cN^2 P/M^3 + (c_\theta L_\theta + c_r L_r + 1) PN/M]. \tag{9}$$

The asymptotically optimal[1] choice of $M = [c/(c_\theta L_\theta + c_r L_r + 1)]^{1/2} N^{1/2}$ implies that C_1 scales only as $O(PN^{1.5})$, as compared with the original cost of $O(PN^2)$. For $P = O(N)$ the cost of the domain decomposition scheme becomes $O(N^{2.5})$ as compared with $O(N^3)$.

The $O(N^{2.5})$ computational cost can be further reduced if instead of the single-level domain-decomposition scheme we adopt a multilevel procedure. If, for the sake of simplicity, we assume that the image size N is a power of two, the image f is subdivided into four subimages of equal size as shown in Fig. 2. This process of subdivision into smaller and smaller subimages can be repeated recursively, until the single pixel level is reached. Of course, at that level computation of projections becomes trivial, since projections of the pixel basis function $\tilde{p} = \Re p$ are (assumed to be) known analytically and only $O(1)$ operations are needed. Computation of projections is thus reduced to a multilevel aggregation of projections of subimages.

The proposed multilevel computational sequence can be most easily formulated as a recursive procedure. The following pseudo-algorithm describes the computation of sampled projections $\tilde{f}$ from the discrete representation of an image, f:

```
f̃ = Fast_Reprojection(f)
if size (f) = 1 × 1 (one pixel)
  f̃ = Direct_Reprojection(f)
else
  f̃ = 0
  for i = 1 to Number of subimages
    f̃_i^s = Fast_Reprojection(f_i^s)
    f̃ = f̃ + S̃[x_i, y_i] I_θ^M f̃_i^s
  end for
end if
```

The notation is as in the two-level algorithm. For the considered case of $M = 2$, the number of subimages is four, and the action of operator I_θ^M amounts to doubling the number of projections through interpolation.

Let the size of the subimage at stage i be $N2^{-i}$. The cost of aggregating its projections, which involves radial shifting and angular interpolation of its projections to $K_\theta P 2^{-i+1}$views followed by addition is

$$C_i = K_\theta K_r (c_\theta L_\theta + c_r L_r + 1) 2^{-2i+1} PN \tag{10}$$

similarly to the previous calculation. Because there are 4^i subimages at the ith level of the recursion, and there are $\log N$ such levels, the total cost becomes

$$2K_\theta K_r (c_\theta L_\theta + c_r L_r + 1) PN \log N. \tag{11}$$

Thus, the computational complexity of the multilevel domain decomposition algorithm is $O(NP \log N)$, or for $P = O(N)$, it becomes $O(N^2 \log N)$.

The foregoing calculation of the computational cost assumed constant interpolator lengths and oversampling factors, independent of N. Theoretically, because of accumulation of interpolation error over the $\log N$ levels of the algorithm, these interpolator lengths (and/or the oversampling factors) would need to be adjusted to maintain constant error for different N. To analyze these second-order effects, let $\varepsilon_\theta$and ε_r denote the relative interpolation errors at one level of the recursion. Then, to first order in the relative error, the worst-case relative error accumulated in the $\log N$ levels of the algorithm will be

$$E = (1 + \varepsilon_\theta + \varepsilon_r)^{\log N} - 1 \approx (\varepsilon_\theta + \varepsilon_r) \log N. \tag{12}$$

To simplify the calculation, let $K_\theta = K_r = K = 2$, and $L_\theta = L_r = L$. It follows [40] that $\varepsilon_\theta, \varepsilon_r < \varepsilon$, with $\varepsilon = e^{-\pi L(K-1)/K} = e^{-\pi L/2}$. Assuming that the total error E is to be independent of N, one obtains from (12) that a sufficient condition is $\varepsilon = E/(2 \log N)$, and hence

$$L = \frac{2}{\pi \log e} \left(\log \log N - \log \frac{E}{2} \right). \tag{13}$$

Substituting (13) into (11) yields this time a cost of $O(N^2 \log N \log \log N)$.

Thus, in theory, error accumulation in the multilevel algorithm may reduce the available speedup relative to the direct method asymptotically by a factor of up to $\log \log N$. However, as reported in the next section, we have not observed such error accumulation in our numerical experiments, suggesting that the worst-case analysis may be conservative in practice. The actual speedup observed is close to the one predicted by the $O(N^2 \log N)$dependence.

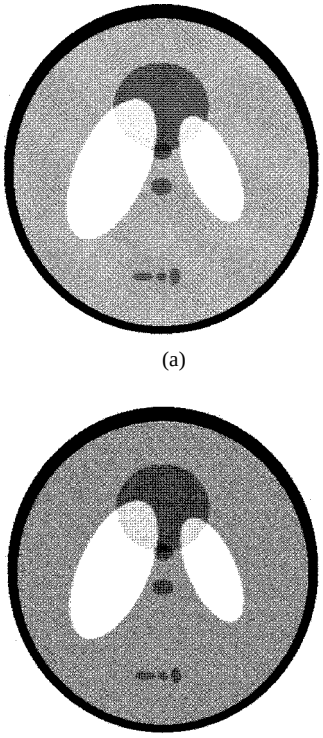
Needless to say, the theoretical predictions of computational cost do not account for details of the implementation, such as the architecture of a particular processor used. Issues such as memory size and bandwidth, caches, and parallel processors come into play, and could affect the actual scaling of computation time with problem size over a wide range of N.

IV. Numerical Examples

Prior to a numerical implementation of the algorithm described in the previous section, one has to make choices regarding the interpolation, selection of the basis function, and the amount of oversampling. These parameters are interrelated and they affect both the accuracy and complexity of the algorithm. A comprehensive study of the optimal parameter selection is beyond the scope of this paper. Choices facilitating easy implementation are adopted in the program that we have developed to demonstrate the accuracy of the proposed algorithm.

The basic algorithm involves two interpolation steps: angular interpolation in going from finer to coarser levels and radial interpolation to account for the radial shift. Of all the components of the algorithm, the issue of interpolation deserves most attention. As mentioned earlier, with appropriate oversampling, interpolators ensuring asymptotically exponential decay of the interpolation error with the interpolator length can be used [40]. However, for moderate image sizes and only approximately bandlimited basis functions the asymptotic behavior may not be reached, and the tradeoffs between angular and radial interpolator lengths, amounts of angular and radial oversampling, total computational cost, and nature of the errors require careful study. For the sake of simplicity, only linear interpolation is implemented in our program for both radial and angular variables.

[1]This choice minimizes C_1, and the approximation of M by the nearest integer will be close to the optimum for sufficiently large N.



(a) (b)

Fig. 3. FBP reconstructions of the Shepp–Logan head phantom for $N = 256$ from $P = 768$ projections computed using (a) the direct algorithm and (b) the fast algorithm with three-fold radial oversampling ($K_r = 3$) and no angular oversampling ($K_\theta = 1$). With these parameters, the fast reprojection was about 22 times faster than the direct one.

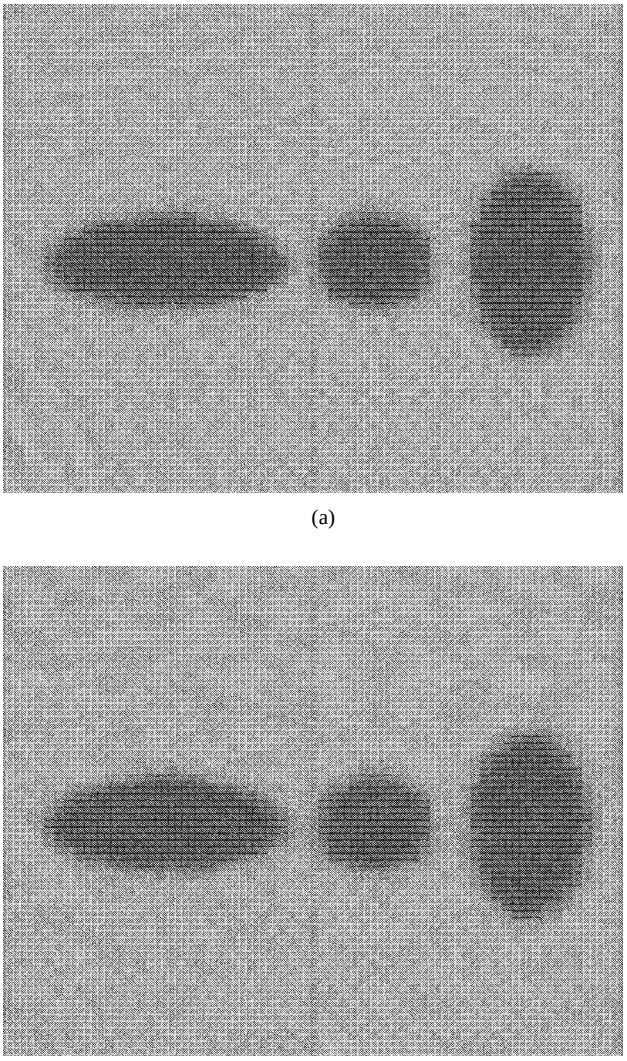
To ensure accurate linear interpolation, the basis function must be sufficiently smooth and some oversampling is necessary. A basis function comprising a tensor product of the cubic B-spline [7]

$$p(x, y) = b(x/\Delta)b(y/\Delta) \tag{14}$$

is chosen for its good approximation and smoothness properties. Recall that the cubic B-spline is given by a quadruple convolution of unit pulses, namely, $b = b_o * b_o * b_o * b_o$, where $b_o(t) = 1$ for $|t| \leq 1/2$ and zero otherwise.

The well-known Shepp–Logan phantom[2] [41], whose FBP reconstruction is shown in Fig. 3(a), is selected for this study. In all our experiments, we compute with the direct algorithm $P = 3N$ uniformly spaced projections (close to the standard angular sampling rule $P = \pi N$ [1,2,38]) in the interval $0 \leq \theta < \pi$ for an $N \times N$ image. Each projection is uniformly sampled in the radial variable at intervals equal to pixel size Δ. The results presented in Figs. 3–6 and Fig. 8 for the fast algorithm are for three-fold radial oversampling ($K_r = 3$), and no angular oversampling ($K_\theta = 1$) for all N.

In most applications, the results of reprojection are subsequently used for reconstructing an image. It is therefore important to assess the quality of the reprojection in terms of the quality of the image reconstructed from it. Therefore, in Fig. 3(a) and (b) we compare, for $N = 512$, images reconstructed from projections computed using the direct and fast algorithm, respectively. The reconstruction was performed using the FBP algorithm with the Shepp–Logan filter [41] and linear interpolation in the FBP, as implemented in the `iradon` function of Matlab 5. For increased detail, we show in Fig. 4 blow-up images of the region containing three small ellipses near the bottom of the reconstructions in Fig. 3. In both Figs. 3 and 4, the gray-level display has been truncated to the range $[0, 0.05]$ (with 1.00 corresponding to the skull) to improve viewing contrast and emphasize low-level errors. The results obtained by direct reprojection and by the proposed fast algorithm show little, if any, visible difference. Fast reprojection was, however, more than 20 times faster than direct reprojection.

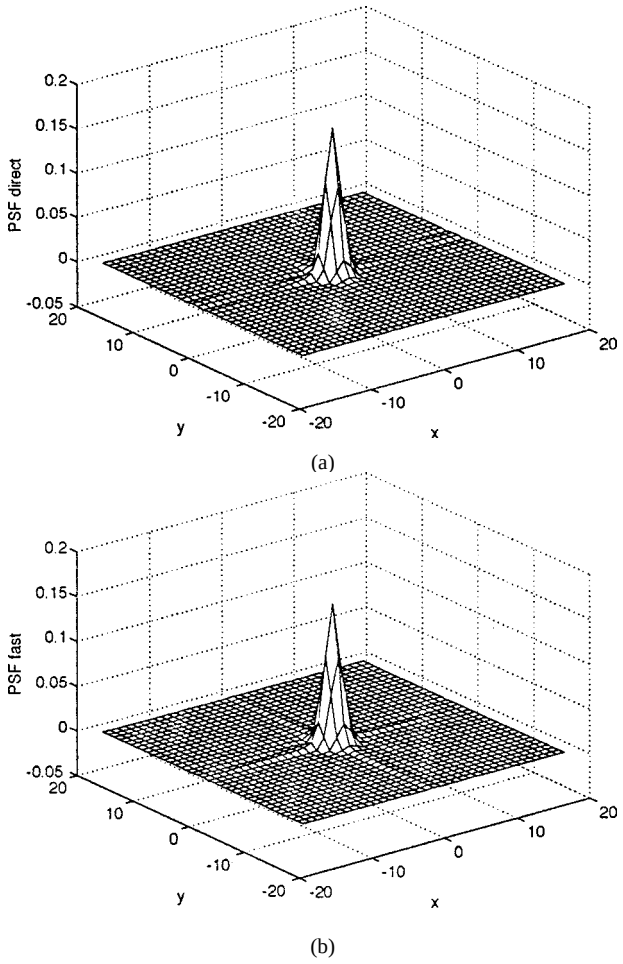


(a) (b)

Fig. 4. Blow-ups from Fig. 3: (a) direct reprojection and (b) fast reprojection.

As yet another way to compare the accuracy of the direct and fast reprojection algorithms, we computed the "impulse responses" or point-spread functions (PSFs) of the cascade of reprojection followed by FBP reconstruction for the two algorithms. Using the same parameters as before, except $N = 256$

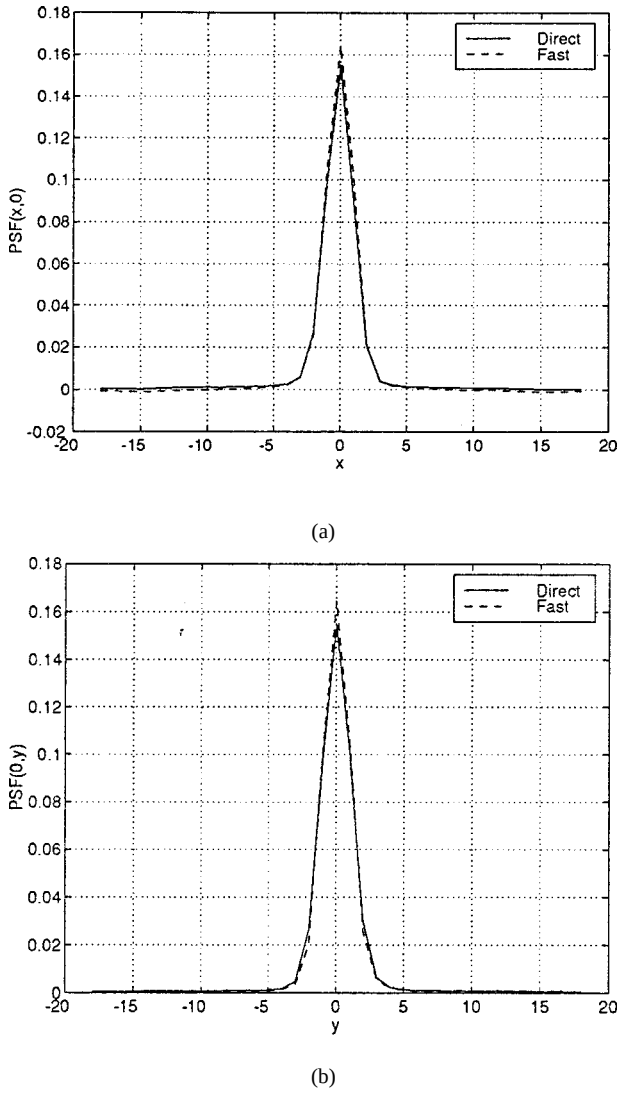
[2]Note that in disctinction from the so-called "modified Shepp–Logan phantom," the skull in this phantom has density 100 times that of interior features. Consequently, this phantom is considerably more challenging than its modified version: even small artifacts in the reconstruction of the skull propagate into the interior "brain" area, and are readily visualized on the relatively low-contrast brain.



(a)

(b)

Fig. 5. Impulse responses of the cascade of re-projection followed by FBP, for $N = 256$. Only the central 41×41 pixel region shown: (a) using the direct algorithm and (b) Using the fast algorithm with three-fold radial oversampling ($K_r = 3$) and no angular oversampling ($K_\theta = 1$).



(a)

(b)

Fig. 6. Slices along the x and y axes through the impulse responses shown in Fig. 5.

and $P = 768$, we computed the projections of an object with a single nonzero coefficient in the expansion (1), and then reconstructed the image using the FBP. The resulting "impulse responses" are shown in Fig. 5, with slices along the two axes compared in Fig. 6. The figures reveal nearly identical shapes for the main lobe of the PSF's. Fig. 5 suggests that the PSF for the fast algorithm has more noticeable "ribs" along the two axes away from the center, however this visual impression is qualified by Fig. 6. The ribs of the PSF's along the y axis virtually coincide, whereas those along the x axis are only slightly higher for the fast algorithm close to the center area. The x axis ribs in Fig. 6(a) have comparable magnitude away from the center, except that they are negative for the direct reprojection, and therefore do not show as well in the meshplot Fig. 5(a) at the viewpoint chosen. As another test, we repeated the experiment with the "impulse object" located at various positions throughout the input image. The responses at different positions were identical the two algorithms, confirming that both the direct and fast algorithms provide shift-invariant responses.

We conclude from these various comparisons that direct and the proposed fast reprojection should have very similar spatial resolution characteristics. It is interesting to note that authors of the length-doubling fast algorithms report a degradation of the resolution for their fast algorithms, with the worst-case degradation theoretically predicted to be proportional to $\log N$ [26], [33], [34].[3] The proposed algorithm does not appear to suffer such resolution degradation.

Finally, to quantify the accuracy of the fast reprojection relative to exact reprojection, we define the normalized rms error increment in the fast reprojection

$$\Delta_{\mathrm{rms}}(\tilde{f}) = \left(\frac{\sum_{ij} \left(\tilde{f}_{ij}^F - \tilde{f}_{ij}^D \right)^2}{\sum_{ij} \left(\tilde{f}_{ij}^D \right)^2} \right)^{1/2} \tag{15}$$

where $\tilde{f}_{ij}^F$ and $\tilde{f}_{ij}^D$ denote the projections computed by the proposed fast algorithm and by direct reprojection, respectively. Here, indices i and j correspond to the sampled angular and radial directions, respectively. As before, it often more meaningful

[3]In fact, a post-processing deconvolution step is proposed in [34] in an attempt to correct for the broadened impulse response.

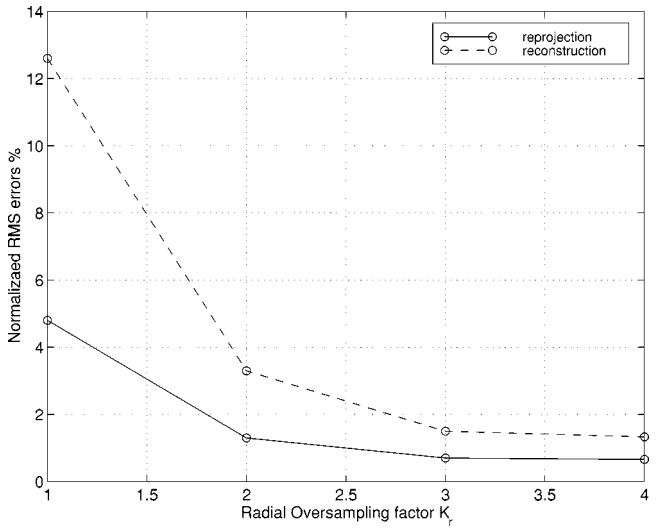


Fig. 7. Normalized RMS errors $\Delta_{\rm rms}(\bar{f})$ and $\Delta_{\rm rms}(f)$ of the fast reprojection and of the image reconstructed from it, respectively, versus the radial oversampling factor K_r. No angular oversampling was used $K_\theta = 1$.

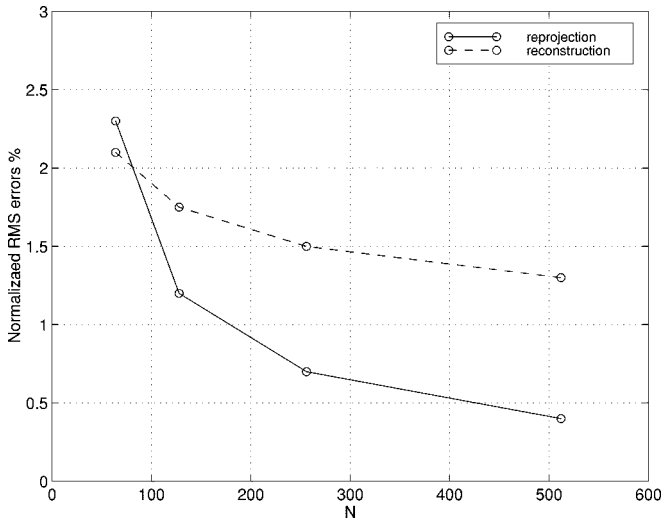


Fig. 8. Normalized RMS errors $\Delta_{\rm rms}(\bar{f})$ and $\Delta_{\rm rms}(f)$ of the fast reprojection and of the image reconstructed from it, respectively, versus image size, for $K_r = 3$ and $K_\theta = 1$.

to measure the error in the FBP reconstruction from the reprojection data. Denoting the approximate inverse Radon transform computed by the FBP by $\hat{R}^{-1}$, we define the normalized error increment in the reconstruction due to fast reprojection by

$$\Delta_{\rm rms}(f) = \left(\frac{\sum_{ij} \left(\hat{R}^{-1} \tilde{f}^D_{ij} - \hat{R}^{-1} \tilde{f}^F_{ij} \right)^2}{\sum_{ij} (f_{ij})^2} \right)^{1/2} . \qquad (16)$$

Small values of $\Delta_{\rm rms}(f)$ would indicate that the fast reprojection is comparable to direct reprojection, as far as the effect on reconstructions is concerned.

Fixing all other parameters as before (including $K_\theta = 1$), we calculated $\Delta_{\rm rms}(\bar{f})$ and $\Delta_{\rm rms}(f)$ for different radial oversampling factor K_r. The resulting normalized errors are shown in Fig. 7. The improvement in both error measures with increased oversampling is evident. The curves "saturate" at about $K_r = 3$, because other error sources begin to dominate. This behavior, and the relatively small error increment due to the fast reprojection at these parameter settings motivated our choice of $K_r = 3$ and $K_\theta = 1$ for all other experiments. We expect that better performance tradeoffs will be possible at reduced oversampling factors by employing optimal interpolation schemes [40].

The computational complexity is verified by considering images of various sizes. In Fig. 8, we present normalized CPU time versus image size N for computing $P = 3N$ uniformly sampled projections via the direct reprojection and the fast algorithm. The fast algorithm uses linear interpolation, radial oversampling by a factor $K_r = 3$, and no angular oversampling ($K_\theta = 1$) for all N. The computations for both algorithms were performed on the same DEC PC. The slope of the direct reprojection curve corresponds to the expected $O(N^3)$ behavior. In addition, the observed trend of the fast algorithm curve supports the estimated $O(N^2 \log N)$ complexity of the proposed algorithm. The crossover point between the two algorithms is between $N = 8$ and $N = 16$. Apparently, the overhead involved in the fast algorithm is small enough for it to be faster for $N = 16$ and higher.[4]

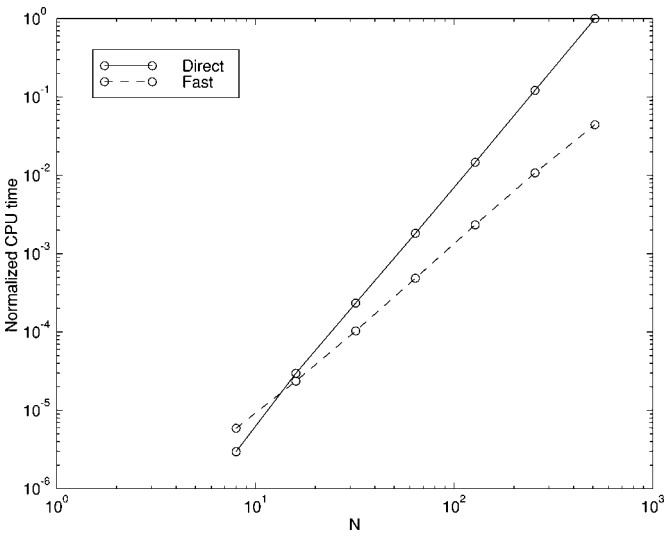


Fig. 9. Normalized CPU times versus image size N for computing projections via direct re-projection and the fast algorithm.

V. Conclusion

An algorithm for fast $O(N^2 \log N)$ reprojection of tomographic images via domain decomposition has been presented. The multilevel computational sequence essentially reduces to the aggregation of projections of larger and larger images, starting with single pixel projections and ending with those of the whole image. Numerical experiments have confirmed the $O(N^2 \log N)$ dependence of the computational requirements down to $N = 16$. The spatial resolution characteristics of the new algorithm, determined by studying the PSF of its cascade with FBP reconstruction, are very similar to those obtained

[4]However, recall that the *accuracy* comparison between the algorithms for N less than about 32 is not very meaningful in the case of Shepp–Logan phantom, because for those values both the reprojections and reconstructions (using either algorithm) are poor due to large discretization errors.

with direct reprojection. Furthermore, FBP reconstructions of a 512×512 Shepp–Logan head phantom from projections computed using the new algorithm produced results visually indistinguishable from direct reprojection, at a more than 20-fold speedup.

Further study is required, however, to optimize various parameters (including interpolation and oversampling) in the algorithm, and to compare the algorithm to other fast algorithms. Extension of the algorithm to three-dimensions will be addressed in forthcoming publications.

REFERENCES

[1] S. Deans, *The Radon Transform and Some of Its Applications*. New York: Wiley, 1983.

[2] F. Natterer, *The Mathematics of Computerized Tomography*. Chichester, U.K.: Wiley, 1986.

[3] C. R. Crawford, J. G. Colsher, N. J. Pelc, and A. H. R. Lonn, "High speed reprojection and its application," in *Proc. SPIE Int. Soc. Opt. Eng. Conf. Medical Imaging II*, vol. 914, Newport Beach, CA, 1988, pp. 311–318.

[4] P. M. Joseph and R. D. Spital, "A method for correcting bone induced artifacts in computed tomography scanners," *JCAT*, vol. 2, no. 1, pp. 100–108, 1978.

[5] G. Henrich, "A simple computational method for reducing streak artifacts in CT images," *Comput. Tomogr.*, vol. 4, pp. 67–71, 1980.

[6] G. Glover and N. Pelc, "An algorithm for the reduction of metal clip artifacts in CT reconstructions," *Med. Phys.*, vol. 8, no. 6, pp. 799–807, 1981.

[7] C. R. Crawford, "Method and apparatus for imaging in the presence of high density objects," U.S. Patent 4 709 333, 1987.

[8] H. Hu, C. R. Crawford, A. N. Pfoh, and J. Hsieh, "Tomographic image reconstruction using cross-plane rays," U.S. Patent 5 396 528, 1995.

[9] J. Hsieh, "Methods and apparatus for reducing partial volume artifacts," U.S. Patent 5 727 041, 1998.

[10] J. M. Ollinger, "Reconstruction-reprojection processing of transmission scans and the variance of PET images," *IEEE Trans. Nucl. Sci.*, vol. 39, p. 1122, Aug. 1992.

[11] S. C. Huang and D. C. Yu, "Capability evaluation of a sinogram error detection and correction method in computed tomography," *IEEE Trans. Nucl. Sci.*, vol. 39, pp. 1106–1110, Aug. 1992.

[12] E. C. Frey, Z. W. Ju, and B. M. W. Tsui, "A fast projector-backprojector pair modeling the asymmetric, spatially varying scatter response function for scatter compensation in SPECT imaging," *IEEE Trans. Nucl. Sci.*, vol. 40, pp. 1192–1197, Aug. 1993.

[13] B. M. W. Nowak, "View-to-view image correction for object motion," U.S. Patent 4 858 128, 1989.

[14] B. M. W. Arata, "Motion correction based on reprojetion data," U.S. Patent 5 552 605, 1996.

[15] J. M. Ollinger, "Iterative reconstruction-reprojection and the expectation-maximization algorithm," *IEEE Trans. Med. Imag.*, vol. 9, p. 94, Mar. 1990.

[16] A. H. Delaney and Y. Bresler, "A fast and accurate Fourier algorithm for iterative parallel-beam tomography," *IEEE Trans. Image Processing*, vol. 5, pp. 740–753, May 1996.

[17] A. H. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: An application to limited-angle tomography," *IEEE Trans. Image Processing*, vol. 7, pp. 204–221, Feb. 1998.

[18] G. L. Zeng, Y. L. Hsieh, and G. T. Gullberg, "A rotating and warping projector/backprojector for fan-beam and cone-beam iterative algorithm," *IEEE Trans. Nucl. Sci.*, vol. 41, pp. 2807–2811, Dec. 1994.

[19] J. L. C. Sanz and E. B. Hinkle, "Computing projections of digital images in image processing pipeline architectures," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 198–207, Feb. 1987.

[20] E. B. Hinkle, J. L. C. Sanz, A. K. Jain, and D. Petkovic, "P/sup 3/E: New life for projection-based image processing," *J. Parallel Distrib. Comput.*, vol. 4, pp. 45–78, Feb. 1987.

[21] P. M. Joseph, "An improved algorithm for reprojecting rays through pixel images," *IEEE Trans. Med. Imag.*, vol. MI-1, no. 3, pp. 192–198, 1982.

[22] D. C. Yu and S. C. Huang, "Study of reprojection methods in terms of their resolution loss and sampling errors," *IEEE Trans. Nucl. Sci.*, vol. 40, p. 1174, Aug. 1993.

[23] C. R. Craword, "Reprojection using a parallel backprojector," *Med. Phys.*, vol. 13, no. 4, pp. 480–483, 1986.

[24] H. Schomberg and J. Timmer, "The gridding method for image reconstruction by Fourier transformation," *IEEE Trans. Med. Imag.*, vol. 14, pp. 596–607, Sept. 1995.

[25] M. Brady and W. Yong, "Fast parallel discrete approximation algorithms for the Radon transform," in *Proc. ACM Symp. Parallel Algorithms Architectures*, July 1992, pp. 91–99.

[26] M. L. Brady, "A fast discrete approximation algorithm for the radon transform," *SIAM J. Comput.*, vol. 27, pp. 107–119, Feb. 1998.

[27] W. A. Gotz, "Ein Schnelle Diskrete Radon Transformation Basierend auf Rekursiv Definierten Digitalen Geraden," Ph.D. thesis, Univ. Innsbruck, Innsbruck, Austria, 1993.

[28] W. A. Gotz and H. G. Druckmuller, "A fast digital Radon transform—An efficient means for evaluating the Hough transform," *Pattern Recognit.*, vol. 28, pp. 711–718, 1996.

[29] P. E. Danielsson, "Iterative techniques for projection and back-projection," Dept. Elect. Eng., Linkoping Univ., Linkoping, Sweden, ISSN 1400-3902, June 10, 1997.

[30] A. Brandt and J. Dym, "Fast calculation of multiple line integrals," *SIAM J. Sci. Comput.*, vol. 20, pp. 1417–1429, 1999.

[31] S. Nillson, "Fast backprojection," Dept. Elect. Eng., Linkoping Univ., Linkoping, Sweden, ISSN 1400-3902, July 4, 1996.

[32] S. Nillson, "Application of fast backprojection techniques for some inverse problems of integral geometry," Dissertation 499, Dept. Math., Linkoping Univ., S-58 183 Linkoping, Sweden, Sept. 1997.

[33] M. Ingerhed, "Fast backprojection for computed tomography, implementation and evaluation," Dept. Elect. Eng., Linkoping Univ., SE-58 183 Linkoping, Sweden, Rep. LIU-TEK-LIC-1999:17, Apr. 1999.

[34] A. Brandt, J. Mann, M. Brodski, and M. Galun, "A fast and accurate inversion of the Radon transform," SIAM J. Appl. Math., to be published.

[35] E. Michielssen and A. Boag, "A multilevel matrix decomposition algorithm for analyzing scattering from large structures," *IEEE Trans. Antennas Propagat.*, vol. AP-44, no. 8, pp. 1086–1093, 1996.

[36] D. Slepian, "Some comments on Fourier analysis, uncertainty, and modeling," *SIAM Rev.*, vol. 25, pp. 379–393, July 1983.

[37] R. M. Lewitt, "Multidimensional digital image representations using generalized Kaiser–Bessel window functions," *J. Opt. Soc. Amer. A*, vol. 7, pp. 1834–1846, Oct. 1990.

[38] P. A. Rattey and A. G. Lindgren, "Sampling the 2-D Radon transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, no. 5, pp. 994–1002, 1981.

[39] A. A. G. Requicha, "The zeros of entire functions: Theory and engineering applications," *Proc. IEEE*, vol. 68, pp. 308–328, Mar. 1980.

[40] J. J. Knab, "Interpolation of band-limited functions using the approximate prolate series," *IEEE Trans. Inform. Theory*, vol. 25, no. 6, pp. 717–720, 1979.

[41] L. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," *IEEE Trans. Nucl. Sci.*, vol. NS-21, no. 1, pp. 21–43, 1974.

Amir Boag (S'89–M'91–SM'96) received the B.Sc. degree in electrical engineering and the B.A. degree in physics in 1983 (both summa cum laude), the M.Sc. degree in electrical engineering in 1985, and the Ph.D. degree in electrical engineering in 1991, all from Technion—Israel Institute of Technology, Haifa.

From 1991 to 1992, he was on the Faculty of the Department of Electrical Engineering, Technion. From 1992 to 1994, he was a Visiting Assistant Professor with the Electromagnetic Communication Laboratory, Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign. In 1994, he joined Israel Aircraft Industries as a Research Engineer and became a Manager of the Electromagnetics Department in 1997. Since 1999, he has been a Senior Lecturer in the Electrical Engineering-Physical Electronics Department, Tel Aviv University, Tel Aviv, Israel. His interests are in electromagnetic theory, wave scattering, and design of antennas and optical devices. He has published more than 30 journal articles and presented over 40 conference papers on electromagnetics and acoustics.

Yoram Bresler (S'82–M'85–SM'91–F'99) received the B.Sc. (cum laude) and M.Sc. degrees from the Technion—Israel Institute of Technology, Haifa, in 1974 and 1981, respectively, and the Ph.D. degree from Stanford University, Stanford, CA, in 1985, all in electrical engineering.

From 1974 to 1979, he was an Electronics Engineer in the Israeli Defense Force. From 1979 to 1981, he was a Consultant for the Flight Control Lab at the Technion, developing algorithms for autonomous TV aircraft guidance. From 1985 to 1987, he was a Research Associate at the Information Systems Laboratory, Stanford University, working on sensor array processing and medical imaging. In 1987, he joined the University of Illinois, Urbana-Champaign, where he is currently a Professor in the Department of Electrical and Computer Engineering and the Bioengineering Program, and Research Professor with the Coordinated Science Laboratory. His current research interests include multidimensional and statistical signal processing and their applications to inverse problems in imaging. He is on the editorial board of *Machine Vision and Applications*.

Dr. Bresler was an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING in 1992–1993, and a Member of the IEEE Image and Multidimensional Signal Processing Technical Committee in 1994–1998. In 1988 and 1989, he received the Senior Paper Awards from the IEEE Acoustics, Speech, and Signal Processing Society. He is the recipient of a 1991 NSF Presidential Young Investigator Award, the Technion Fellowship in 1995, and the Xerox Senior Award for Faculty Research in 1998. In 1999, he was named a University of Illinois Scholar.

Eric Michielssen (M'95–SM'99) received the M.S. degree in electrical engineering (summa cum laude) from the Katholieke Universiteit Leuven (KUL), Belgium, in 1987, and the Ph.D. in electrical engineering from the University of Illinois, Urbana-Champaign (UIUC), in 1992.

He was a Research and Teaching Assistant with the Microwaves and Lasers Laboratory, KUL, from 1987 to 1988, and the Electromagnetic Communication Laboratory, UIUC, from 1988 to 1992. He joined the Faculty of the Department of Electrical and Computer Engineering, UIUC, as a Visiting Assistant Professor in 1992, was appointed Assistant Professor of Electrical and Computer Engineering in 1993, and promoted to Associate Professor in 1998. Since 1995, he has served as Associate Director of the Center for Computational Electromagnetics, UIUC. He has authored or co-authored more than 70 journal papers and book chapters and more than 100 papers in conference proceedings. His research interests include all aspects of theoretical and applied computational electromagnetics. His principal research focus has been on the development of fast frequency and time domain integral-equation-based techniques for analyzing electromagnetic phenomena, and the development of robust, genetic algorithm driven optimizers for the synthesis of electromagnetic devices. He served as the Technical Chairman of the 1997 Applied Computational Electromagnetics Society (ACES) Symposium (Review of Progress in Applied Computational Electromagnetics, March 1997, Monterey, CA), and currently serves on the ACES Board of Directors and as ACES Vice-President. From 1997 to 1999, he was an Associate Editor for Radio Science.

Dr. Michielssen received a Belgian American Educational Foundation Fellowship in 1988 and a Schlumberger Fellowship in 1990. He was the recipient of a 1994 International Union of Radio Scientists (URSI) Young Scientist Fellowship, a 1995 National Science Foundation CAREER Award, and the 1998 Applied Computational Electromagnetics Society (ACES) Valued Service Award. Recently, he was named 1999 URSI—United States National Committee Henry G. Booker Fellow and selected as the recipient of the 1999 URSI Koga Gold Medal. Currently he is an Associate Editor for the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION. He is a member of URSI Commission B.