$O(N^3 \log N)$ Backprojection Algorithm for the 3-D Radon Transform

Samit Basu, *Member, IEEE*, and Yoram Bresler*, *Fellow, IEEE*

***Abstract*—We present a novel backprojection algorithm for three-dimensional (3-D) radon transform data that requires $O(N^3 \log_2 N)$ operations for reconstruction of an $N \times N \times N$ volume from $O(N^2)$ plane-integral projections. Our algorithm uses a hierarchical decomposition of the 3-D radon transform to recursively decompose the backprojection operation. Simulations are presented demonstrating reconstruction quality comparable to the standard filtered backprojection, which requires $O(N^5)$ computations under the same circumstances.**

***Index Terms*—Backprojection, 3-D radon transform, fast algorithm, hierarchical, cone beam tomography.**

I. Introduction

RECONSTRUCTION of three-dimensional (3-D) objects from their plane-integral projections has been a problem of great interest for well over a decade. Data in the form of plane-integral projections, also known as samples of the 3-D radon transform [1] (distinct from the line-integral data of the X-ray transform) arises in several important applications. First, in some imaging modalities, such as magnetic resonance imaging (MRI) or synthetic aperature radar (SAR), it is possible to sample the 3-D radon transform directly [1], [2]. In fact, in MRI, data acquisition modes corresponding to samples of the 3-D radon transform, known as 3-D-projection modes, may offer advantages in terms of robustness to motion artifacts (e.g., see [3]). Second, much current research has focused on the derivation of reconstruction algorithms for the cone-beam acquisition geometry, which assumes a point source and a two-dimensional (2-D) detector. An important class of inversion techniques, such as Grangeat's method or Smith's method [4], [5], process the cone-beam projections to produce samples of the first derivative or Hilbert transform of the 3-D radon transform. A method for inverting the 3-D radon transform can then be used to perform the reconstruction.[1]

In either case, whether the 3-D radon transform data arises by cone beam data collection followed by a transformation, or by direct acquisition, the reconstruction problem requires inversion of the 3-D radon transform. Unfortunately, in three dimensions, the standard filtered backprojection (FBP) algorithm is computationally very expensive, requiring $O(N^5)$ operations for recovery of an N^3 volume from N^3 samples of the radon transform. The standard fast algorithm for reconstruction is the well-known algorithm of Marr *et al.* [7]. This algorithm is based on a factorization of the 3-D radon transform into 2-D radon transforms, and leads ultimately to an $O(N^4)$ backprojection algorithm.

Recently, however, a number of even faster algorithms have been proposed to perform reconstruction from 3-D radon transform data in $O(N^3 \log_2 N)$ operations [7]–[9]. These fast methods are generally extensions of preexisting fast methods for the 2-D radon transform, and are either reformulations of 2-D techniques for the 3-D problem, or utilize the Marr factorization of the 3-D radon transform into a pair of 2-D radon transforms along with fast 2-D algorithms. To these $O(N^3 \log_2 N)$ algorithms we might add one that can be constructed by combining the Marr decomposition of [6] with our own fast hierarchical backprojection (FHBP) algorithm for the 2-D radon transform [10].

In this paper, we present an extension of our 2-D FHBP algorithm to a fast backprojection algorithm for the 3-D radon transform. The resulting algorithm, which we dub the 3-D FHBP, is a new method for fast backprojection with a cost of $O(N^3 \log_2 N)$ operations for reconstructing a N^3 volume from N^3 samples of the 3-D radon transform. The new algorithm is a "native" 3-D algorithm, and does not rely on the factorization of the 3-D radon transform into a pair of 2-D radon transforms, as are used in [6] and [7].

This difference is important because the factorization of the 3-D radon transform relies on a special (angularly separable) sampling pattern of the parameter space, on which data may not be available. For example, using the Grangeat transformation from cone-beam to (derivative of) 3-D radon, data are obtained at particular angles determined by the source trajectory. To apply the Marr decomposition, they need to be angularly

Manuscript received February 17, 2000; revised December 14, 2001. This work was supported in part by the National Science Foundation (NSF) under Grant CCR 99-72980 and under Research Infrastructure Grant CDA 96-24396, in part by a Fellowship from the Joint Services Electronics Program, in part by the Mac VanValkenburg Graduate Fellowship, and in part by the National Computational Science Alliance under Grant ASC9900024N using the NCSA SGI/CRAY Origin2000. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was C. Crawford. *Asterisk indicates corresponding author.*

S. Basu was with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. He is now with General Electric Global Research Center, CT Laboratory, Niskayuna, NY 12309 USA.

*Y. Bresler is with the Coordinated Science Laboratory and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 1308 W. Main Street, Urbana, IL 61801 USA (e-mail: ybresler@uiuc.edu).

Publisher Item Identifier S 0278-0062(02)02939-7.

[1]Although the 3-D radon inversion may be performed by a FBP-type algorithm, the resulting cone-beam inversion algorithms are usually classified as radon inversion algorithms. The term FBP or (FDK-type) cone-beam algorithm is reserved for direct methods that perform *cone-beam* (and not 3-D radon) backprojection.

resampled (interpolated) to the separable sample set. In contrast, our algorithm may be used directly with "native" sampling patterns, eliminating the need for such resampling. Alternatively, our algorithm may be used with a more efficient uniformly spaced grid on the unit sphere[2]. Data may be directly acquired on this grid (e.g., in MRI), or resampled to it.

The proposed 3-D-FHBP algorithm consists of the following steps. The original FBP reconstruction calls for the backprojection of N^2 (filtered) plane-integral projections onto a N^3 volume. Each projection is a 1D function corresponding to integrals over parallel planes. However, using properties of the backprojection, we decompose this into a sum of eight backprojections, each of N^2 plane-integral projections onto $N^3/8$ volumes, where each volume represents one octant of the reconstruction, centered at the origin. We then note that these backprojections are "oversampled", and that the number of plane-integral projections needed to characterize each volume can be reduced by a factor of four. This "scaling" rule is based on a sampling result of [1], and states that for a fixed image bandwidth, the number of projections needed to reconstruct a volume of size N^3 is proportional to N^2, provided that the projection angles satisfy some technical conditions (see Appendix B). Thus, each of the eight backprojections can be replaced by a backprojection of $N^2/4$ plane-integral projections onto $N^3/8$ volumes. This results in a factor of four savings in computation over a direct backprojection. If the algorithm is then applied recursively to each of the octant backprojections, the resulting algorithm has $O(N^3 \log_2 N)$ cost. Thus, reconstruction is possible in a very small fraction of the time necessary for a regular FBP.

The paper is organized along lines similar to [10]. In Section II, we first derive an exact hierarchical decomposition of the backprojection operation. We then use heuristic arguments in Section III to construct the fast (and approximate) algorithm. In Section IV, we demonstrate the performance of our algorithm on a standard 3-D phantom. Conclusions and proposals for future research are given in Section V.

II. Hierarchical Decomposition of the 3-D Backprojection

In this section, we will derive a hierarchical decomposition of the 3-D backprojection. In order to construct an accurate decomposition, we will need to explicitly incorporate the various discretizations that are commonly used in the formulation of the reconstruction problem. Furthermore, to simplify the manipulations as much as possible, we will use an operator formulation, and closely follow the notation of [10]. Hence, operators between discrete spaces will be denoted using a bold Roman font, and operators between continuous spaces, or between continuous and discrete spaces will be denoted using a script font. The Hilbert spaces that will arise in the formulation are $\ell_2(\mathbb{Z}^3)$, the space of 3-D square-summable sequences indexed by $\mathbb{Z}^3$, $L_2(\mathbb{R}^3)$, the space of square-integrable 3-D functions on $\mathbb{R}^3$, $\ell_2^{P\times P}(\mathbb{Z})$, the $P \times P$-wise Cartesian product of $\ell_2(\mathbb{Z})$, and $L_2^{P\times P}(\mathbb{R})$, the $P \times P$-wise Cartesian product of $L_2(\mathbb{R})$. A typical element of $L_2(\mathbb{R}^3)$ is a continuous-space 3-D object (density), while a typical element of $\ell_2(\mathbb{Z}^3)$ is a discrete/sampled 3-D object. On the other hand, a typical element of $L_2^{P\times P}(\mathbb{R})$ is a collection of P^2 continuous-time 1D functions $g_c(r, m, n)$, which represent projections of a continuous 3-D object, and are indexed by $m, n \in \{0, 1, \ldots, P-1\}$, and $r \in \mathbb{R}$. A typical element of the space $\ell_2^{P\times P}(\mathbb{Z})$ is a collection of P^2 discrete-time sequences $g(k, m, n)$, as might be obtained by sampling an element of $L_2^{P\times P}(\mathbb{R})$ in the "radial" direction, and indexed by $m, n \in \{0, 1, \ldots, P-1\}$, and $k \in \mathbb{Z}$.

We will assume that the data collected are samples of the 3-D radon transform, defined by $\mathcal{R}_P : L_2(\mathbb{R}^3) \mapsto L_2^{P\times P}(\mathbb{R})$

$$g_c(s, m, n) = \mathcal{R}_P f_c(s, m, n) = \int_{\mathbb{R}^3} f_c(\boldsymbol{x}) \delta(\boldsymbol{x} \cdot \boldsymbol{\omega}_{m,n} - s)\, d\boldsymbol{x}$$

where δ is the Dirac delta distribution, $\boldsymbol{\omega}_{m,n}$ are the unit vectors in the direction of the angular sample locations, and $\cdot$ denotes the standard dot product on $\mathbb{R}^3$. For fixed m, n, the 1D function $h_{m,n}(s) = \mathcal{R}_P f_c(s, m, n)$ will be called a projection of f_c in the direction $\boldsymbol{\omega}_{m,n}$. While the 3-D-FHBP algorithm can work with a variety of sampling patterns, for the sake of simplicity, we will work in this paper with the separable sampling pattern[3] defined in spherical coordinates by

$$\boldsymbol{\omega}_{m,n} = \begin{bmatrix} \sin((m+0.5)\Delta\phi)\cos(n\Delta\theta) \\ \sin((m+0.5)\Delta\phi)\sin(n\Delta\theta) \\ \cos((m+0.5)\Delta\phi) \end{bmatrix} \quad m, n \in \{0, \ldots, P-1\} \quad (1)$$

where[4] $\Delta\phi = \Delta\theta = (2\pi/P)$. Note that ϕ and θ are the azimuth and elevation angles, respectively.

We start with the well known 3-D FBP reconstruction for the 3-D radon transform, which estimates the object $f_c \in L_2(\mathbb{R}^3)$ as

$$\hat{f}_c = \mathcal{B}_P \mathcal{F} g_c \quad (2)$$

where $\mathcal{B}_P : L_2^{P\times P}(\mathbb{R}) \mapsto L_2(\mathbb{R}^3)$ is the weighted, radially continuous backprojection, defined as

$$\mathcal{B}_P g_c(\boldsymbol{x}) = \sum_{m=0}^{P-1} \sum_{n=0}^{P-1} \sin((m+0.5)\Delta\phi) \times g_c(\boldsymbol{x} \cdot \boldsymbol{\omega}_{m,n}, m, n)\Delta\phi\Delta\theta \quad (3)$$

[2]In [6] it was claimed that the separable sampling pattern is suboptimal, because the samples are not uniformly (in the sense of samples per unit surface area) distributed on the unit sphere. A similar conclusion follows from our discussion in Appendix B of Natterer's results on resolution [1], becuase the separable pattern is redundant, containing about twice the number of point sufficient to ensure that a general sampling pattern is m-resolving.

[3]We recall that if data with this sampling pattern is available, then the elegant method of [6] can be used to obtain reconstructions in $O(N^4)$ operations.

[4]Our somewhat unusual choice of $\phi_m = (m + 0.5)\Delta\phi$ ensures that the samples at the "north and south poles", i.e., $\phi \in \{0, \pi\}$ are excluded. Furthermore, because of the inherent symmetry of the spherical coordinate system, $\boldsymbol{\omega}_{m,n} = \boldsymbol{\omega}_{P-m-1,P/2+n}$ and $\boldsymbol{\omega}_{m,n} = -\boldsymbol{\omega}_{P/2-m-1,P/2+n}$ for $m, n \in \{0, \ldots P/2 - 1\}$. The radon transform therefore has similar symmetries, $g_c(s, m, n) = g_c(s, P-m-1, P/2+n) = g_c(-s, P/2-m-1, P/2+n)$, implying that the set of projections $\{g_c(s, m, n)\}_{m,n=0}^{P-1}$ has a fourfold redundancy. Although we use this fact in our implementation, we omit this nuance in the sequel to simplify the presentation.

and $\mathcal{F} : L_2^{P\times P}(\mathbb{R}) \mapsto L_2^{P\times P}(\mathbb{R})$ is a radial convolution operator, defined by

$$\mathcal{F}g_c(s,m,n) = \int \beta(s-t)g_c(t,m,n)\,dt \tag{4}$$

where β is a suitably chosen apodized second derivative filter $\beta(s)$.

In practice it is impossible to implement (2) for two reasons. First, we do not have access to the continuous projections $g_c(s,m,n)$, but rather to a sampled set of projections $g(l,m,n) = \mathcal{R}_P f_c(lT,m,n)$. We can handle this by applying standard DSP techniques to the sampled data, and approximating the continuous convolution with a discrete one. Second, the left hand side of (2) is a continuous reconstruction that must be discretized to be represented on a computer. Hence, any implementation of (2) is ideally composed of the following four steps.

1) Radial interpolation of the projection data to continuous projections.
2) Backprojection via (3) to a continuous reconstruction.
3) Sampling of the reconstruction on a uniform grid to obtain a discrete reconstruction.
4) Truncation of the discrete reconstruction to the size of interest.

We will define the operator which implements these steps as the *discrete backprojection operator*, and denote it $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}$.

The operator $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}$ maps radially sampled radon transform data back to a reconstruction with $N\times N\times N$ support. The operator $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}$ depends on a matrix of parameters $\boldsymbol{\tau} \in [-.5,.5]^{P\times P}$, which we will describe shortly. For notational brevity, we will occasionally write $\mathbf{B}_{P,N}$, when the specific value for $\boldsymbol{\tau}$ is unimportant to the discussion.

Based on the steps given above, the discrete backprojection is defined as

$$\mathbf{B}^{\boldsymbol{\tau}}_{P,N} = \mathbf{K}_N \mathcal{S}\mathcal{B}_P\mathcal{I}_{\boldsymbol{\tau}} \tag{5}$$

where operator $\mathcal{I}_{\boldsymbol{\tau}} : \ell_2(\mathbb{Z}^3) \mapsto L_2(\mathbb{R}^3)$ is the radial interpolation operator (Step 1)

$$\mathcal{I}_{\boldsymbol{\tau}}g(s,m,n) = \sum_{k\in\mathbb{Z}} g(k,m,n)\psi(T(k+\tau_{m,n}) - s). \tag{6}$$

The function ψ is the radial interpolation kernel. For the case of linear radial interpolation, we would choose $\psi(t) = \Lambda(t)$, where

$$\Lambda(t) = \begin{cases} 1 - \frac{|t|}{T}, & |t| \le T \\ 0, & \text{else.} \end{cases}$$

The parameter $\boldsymbol{\tau} \in [-.5,.5]^{P\times P}$ is a matrix whose m,nth term is the phase of the radial sampling locations for projection direction $\boldsymbol{\omega}_{m,n}$. With $\tau_{m,n} = 0$, the radial samples for the m,nth projection are located at $(\ldots,-2T,-T,0,T,2T,\ldots)$. With $\tau_{m,n} \neq 0$, these sample locations will be shifted by $\tau_{m,n}T$. The effect of $\tau_{m,n}$ on the sample locations is illustrated in Fig.1. The role of $\boldsymbol{\tau}$ will be discussed later.

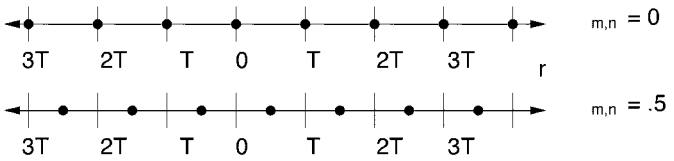

Fig. 1. The effect of $\tau_{m,n}$ on the sample locations for the radial coordinate of the projection at direction $\boldsymbol{\omega}_{m,n}$.

The operator $\boldsymbol{S} : L_2(\mathbb{R}^3) \mapsto \ell_2(\mathbb{Z}^3)$ is the generalized sampling, or mollification, operator that represents the sampling of the continuous reconstruction on a discrete grid (Step 3)

$$\boldsymbol{S}g_c(\boldsymbol{z}) = \int_{\mathbb{R}^3} g_c(\boldsymbol{x})b(\boldsymbol{x}-\boldsymbol{z})\,d\boldsymbol{x}.$$

The function $b \in L_2(\mathbb{R}^3)$ is usually taken to be the indicator function for a unit-size voxel, or the indicator function for a sphere, so that the samples represent local averages of the reconstruction. Smoother choices of b may be more amenable to estimating derivatives from the reconstruction, but for simplicity, we will choose b to be the indicator for a radius-$1/2$ sphere.

The final operator $\mathbf{K}_N : \ell_2(\mathbb{Z}^3) \mapsto \ell_2(\mathbb{Z}^3)$ is the discrete truncation operator (Step 4) that zeros all samples outside a volume of size $N\times N\times N$

$$\mathbf{K}_N g(\boldsymbol{z}) = \begin{cases} g(\boldsymbol{z}), & -\frac{N}{2} < z_1,z_2,z_3 \le \frac{N}{2} \\ 0 & \text{else} \end{cases}.$$

We can write an explicit expression for $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}$ by combining the relevant operator definitions. It is simple to show that

$$\mathbf{B}^{\boldsymbol{\tau}}_{P,N}g(\boldsymbol{z}) = \begin{cases} \sum_m \sum_n \sum_k g(k,m,n)\eta_{\boldsymbol{z},k,m,n}, & -\frac{N}{2} < z_1,z_2,z_3 \le \frac{N}{2} \\ 0, & \text{else} \end{cases} \tag{7}$$

where

$$\eta_{\boldsymbol{z},k,m,n} = \sin[(m+1)\Delta\phi]\int_{\boldsymbol{x}} b(\boldsymbol{x}-\boldsymbol{z})\psi[T(k+\tau_{m,n}) - \boldsymbol{x}\cdot\boldsymbol{\omega}_{m,n}]\,d\boldsymbol{x}. \tag{8}$$

With $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}$ so defined, we note that the cost of computing $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}g$ (which dominates the FBP) is $O(P^2N^3)$ operations, assuming that ψ (and thus η) has small support. If $P = O(N)$, then we recover the $O(N^5)$ cost for the standard backprojection. Our goal in this paper is to derive a fast approximation to $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}$ by constructing a hierarchical decomposition for it. We begin by constructing an exact, but slow $O(N^5)$ decomposition for it. The result is a decomposition of $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}$ as a sum of backprojections onto smaller volumes, which are then shifted and added to form the larger volume.

In order to construct a valid decomposition of the backprojection operation, we need a crucial property of the backprojection operation: *backprojection is local, i.e., each point in the backprojected volume is affected only by projection data for planes passing through a neighborhood of that point.* As a result, when backprojecting onto a portion of the final volume, we will be able to ignore all projection data that does not affect points in

that portion of the volume. To see the argument formally, we first consider the operator $\mathbf{B}^{\boldsymbol{\tau}}_{P,M}$, which can be written explicitly (by expanding (5) with $N = M$) as

$$\mathbf{B}^{\boldsymbol{\tau}}_{P,M} g(\boldsymbol{z}) = \sum_k \sum_m \sum_n g(k,m,n) \sin((m+1)\Delta\phi) \times \Delta\phi\Delta\theta \int_{\mathbb{R}^3} b(\boldsymbol{x}-\boldsymbol{z}) \psi[T(k+\tau_{m,n}) - \boldsymbol{x}\cdot\boldsymbol{\omega}_{m,n}] d\boldsymbol{x} \quad \|\boldsymbol{z}\|_\infty \le M/2 \tag{9}$$

where $\|\cdot\|_\infty$ denotes the standard infinity norm. Suppose, now, that b has radius of support R_1, and ψ has a radius of support R_2. Then it is simple to show through geometric arguments and algebraic simplification, that the integrand of (9) vanishes for all $k, m, n, \boldsymbol{z}$ satisfying

$$|\boldsymbol{\omega}_{m,n}\cdot\boldsymbol{z} - T(k+\tau_{m,n})| \ge R_1 + R_2. \tag{10}$$

Thus, for a given M, the projection sample $g(k,m,n)$ is not used in (9) for any $\boldsymbol{z}$ if and only if

$$\min_{\|\boldsymbol{z}\|_\infty \le M/2} |\boldsymbol{\omega}_{m,n}\cdot\boldsymbol{z} - T(k+\tau_{m,n})| \ge R_1 + R_2. \tag{11}$$

We can simplify (11) by settling for a looser, sufficient condition on k, by applying the inequality $|a-b| \ge ||a|-|b||$ to yield the condition

$$|k+\tau_{m,n}| \ge \frac{R_1 + R_2 + M\sqrt{3}/2}{T} \tag{12}$$

or, by an application of the triangle inequality and the fact that $|\tau_{m,n}| \le (1/2)$

$$|k| \ge \frac{R_1 + R_2 + M\sqrt{3}/2}{T} + \frac{1}{2}. \tag{13}$$

Hence, we define a radial truncation operator $\hat{\mathbf{K}}_M : \ell_2^{P\times P}(\mathbb{Z}) \mapsto \ell_2^{P\times P}(\mathbb{Z})$ which keeps only those points that satisfy one of (11), (12) or (13). For simplicity we will use (13) in the definition

$$\hat{\mathbf{K}}_M g(k,m,n) = \begin{cases} g(k,m,n), & |k| \le \frac{R_1+R_2+M\sqrt{3}/2}{T} + \frac{1}{2} \\ 0, & \text{else} \end{cases}.$$

The following property is a direct consequence of our construction of $\hat{\mathbf{K}}$.

Property 1 (Truncation of $\mathbf{B}_{P,N}$ Input):

$$\mathbf{B}^{\boldsymbol{\tau}}_{P,N} = \mathbf{B}^{\boldsymbol{\tau}}_{P,N}\hat{\mathbf{K}}_N. \tag{14}$$

We also need four shifting operators. The first is a discrete object shifter $\mathbf{Z}_{\boldsymbol{\delta}} : \ell_2(\mathbb{Z}^3) \mapsto \ell_2(\mathbb{Z}^3)$, for $\boldsymbol{\delta} \in \mathbb{Z}^3$, defined by

$$\mathbf{Z}_{\boldsymbol{\delta}} f(\boldsymbol{z}) = f(\boldsymbol{z} - \boldsymbol{\delta}).$$

The second is a continuous object shifter $\mathcal{M}_{\boldsymbol{\delta}} : L_2(\mathbb{R}^3) \mapsto L_2(\mathbb{R}^3)$, for $\boldsymbol{\delta} \in \mathbb{R}^3$, defined by

$$\mathcal{M}_{\boldsymbol{\delta}} f_c(\boldsymbol{x}) = f_c(\boldsymbol{x} - \boldsymbol{\delta}).$$

We will also need shifting operations in the radon transform domain. The third shifter is a discrete projection radial shifter $\hat{\mathbf{Z}}_{\boldsymbol{\delta}} : \ell_2^{P\times P}(\mathbb{Z}) \mapsto \ell_2^{P\times P}(\mathbb{Z})$ for $\boldsymbol{\delta} \in \mathbb{Z}^{P\times P}$, defined by

$$\hat{\mathbf{Z}}_{\boldsymbol{\delta}} g(k,m,n) = g(k - \delta_{m,n}, m, n)$$

and the fourth and final shifter is a continuous projection radial shifter $\hat{\mathcal{M}}_{\boldsymbol{\delta}} : L_2^{P\times P}(\mathbb{R}) \mapsto L_2^{P\times P}(\mathbb{R})$ for $\boldsymbol{\delta} \in \mathbb{R}^{P\times P}$, defined by

$$\hat{\mathcal{M}}_{\boldsymbol{\delta}} g_c(r,m,n) = g_c(r - \delta_{m,n}, m, n).$$

The shift and truncation properties from [10] hold for these operators, with minor changes due to notation. We will restate them without proof.

Property 2 (Shift Property of $\mathcal{B}_P$):

$$\mathcal{M}_{\boldsymbol{\delta}}\mathcal{B}_P g_c(\boldsymbol{x}) = \mathcal{B}_P \hat{\mathcal{M}}_{\boldsymbol{\delta}'} g_c(\boldsymbol{x}) \tag{15}$$

where $\delta'_{m,n} = \boldsymbol{\delta}\cdot\boldsymbol{\omega}_{m,n}$.

Property 3 (Shift Property of $\mathcal{S}$):

$$\mathbf{Z}_{\boldsymbol{\delta}}\mathcal{S} f_c(\boldsymbol{z}) = \mathcal{S}\mathcal{M}_{\boldsymbol{\delta}} f_c(\boldsymbol{z}) \tag{16}$$

for $\boldsymbol{\delta} \in \mathbb{Z}^3$.

Property 4 (Shift Property of $\mathcal{I}_{\boldsymbol{\tau}}$):

$$\hat{\mathcal{M}}_{\boldsymbol{\delta}}\mathcal{I}_{\boldsymbol{\tau}} g(r,m,n) = \mathcal{I}_{\langle \boldsymbol{\tau}+\boldsymbol{\delta}/T\rangle} \hat{\mathbf{Z}}_{-[\boldsymbol{\tau}+\boldsymbol{\delta}/T]} g(r,m,n) \tag{17}$$

where $[\boldsymbol{x}]_{i,j}$ is $\boldsymbol{x}_{i,j}$ rounded to the the nearest integer and $\langle \boldsymbol{x}\rangle_{i,j} = \boldsymbol{x}_{i,j} - [\boldsymbol{x}]_{i,j}$.

Remark: The need for Property 4 is the motivation behind the introduction of $\boldsymbol{\tau}$ into (6), and thus into $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}$. If we assumed that all of the projections were radially sampled at the same radial locations, so that effectively $\boldsymbol{\tau} = 0$, then continuous shifts of the projections that are not exact integers cannot be represented in terms of exact manipulations of the discrete projections (i.e., in general radial shifts and $\mathcal{I}_{\mathbf{0}}$ do not commute). However, by adding the additional degrees of freedom represented by $\boldsymbol{\tau}$, these operations do commute. Furthermore, Property 4 is necessary for the construction of the slow hierarchical decomposition (see Theorem 1 below, and the proof in Appendix A).

With the properties, the following decomposition can be constructed.

Theorem 1: For $N = 2^K$, $K \in \mathbb{Z}, K \ge 1$

$$\mathbf{B}^{\boldsymbol{\tau}}_{P,N} = \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{B}^{\langle \boldsymbol{\nu}_i\rangle}_{P,N/2} \hat{\mathbf{K}}_{N/2} \hat{\mathbf{Z}}_{-[\boldsymbol{\nu}_i]} \tag{18}$$

where $(\boldsymbol{\nu}_i)_{m,n} = \tau_{m,n} - \boldsymbol{\delta}_i\cdot\boldsymbol{\omega}_{m,n}$ and

$$\begin{aligned} &\boldsymbol{\delta}_1 = [-N/4, -N/4, -N/4]^T \quad \boldsymbol{\delta}_2 = [-N/4, -N/4, N/4]^T \\ &\boldsymbol{\delta}_3 = [-N/4, N/4, -N/4]^T \quad \boldsymbol{\delta}_4 = [-N/4, N/4, N/4]^T \\ &\boldsymbol{\delta}_5 = [N/4, -N/4, -N/4]^T \quad \boldsymbol{\delta}_6 = [N/4, -N/4, N/4]^T \\ &\boldsymbol{\delta}_7 = [N/4, N/4, -N/4]^T \quad \boldsymbol{\delta}_8 = [N/4, N/4, N/4]^T \end{aligned} \tag{19}$$

for $N \ge 4$, and

$$\begin{aligned} &\boldsymbol{\delta}_1 = [1,1,1]^T \quad \boldsymbol{\delta}_2 = [1,1,0]^T \\ &\boldsymbol{\delta}_3 = [1,0,1]^T \quad \boldsymbol{\delta}_4 = [1,0,0]^T \\ &\boldsymbol{\delta}_5 = [0,1,1]^T \quad \boldsymbol{\delta}_6 = [0,1,0]^T \\ &\boldsymbol{\delta}_7 = [0,0,1]^T \quad \boldsymbol{\delta}_8 = [0,0,0]^T \end{aligned} \tag{20}$$

for $N = 2$.

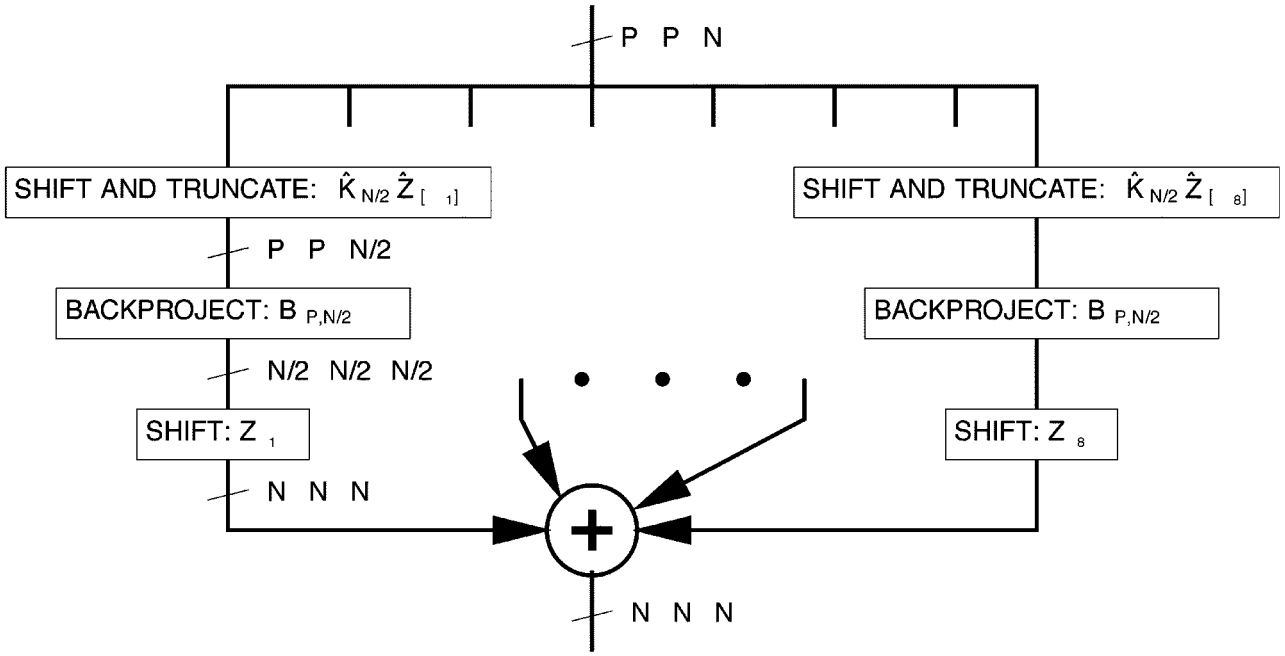

Fig. 2. Diagram of the decomposition of Theorem 1.

The proof of this theorem is given in Appendix A.

The interpretation of Theorem 1 as a procedure for backprojecting P^2 projections onto an N^3 volume is as follows.

1) *Exact Subdivision:* Subdivide the projection data into eight sets, one for each octant of the original reconstruction, by computing $g_i' = \hat{\mathbf{K}}_{N/2}\hat{\mathbf{Z}}_{-[\boldsymbol{\nu}_i]}g(\boldsymbol{z})$. This corresponds to a discrete radial shifting, and truncation operation.
2) *Backprojection:* Backproject the P^2 projections in each g_i' to form f_i, the ith octant of the reconstruction (size $(N/2)^3$), centered at the origin.
3) *Aggregation:* Shift and add the eight octants to form the final reconstruction volume.

The process is exact, i.e., no approximations are involved in the construction of the decomposition. Furthermore, the process can be applied recursively to the backprojections of the octants, a point we will return to later. The decomposition for a single application of Theorem 1 is illustrated in Fig. 2.

To determine the complexity requirements of computing $\mathbf{B}_{P,N}$ via Theorem 1, let us assume that direct computation using Eq. (9) requires a total of $c_1P^2N^3$ operations. Then the subdivision step (Step 1) requires, for the nonzero points of g_i', $c_2P^2N/2$ operations for each $i \in \{1,2,\ldots,8\}$, for a total of $4c_2P^2N$. By our assumption, each backprojection step (Step 2) requires $c_1P^2(N/2)^3$ operations per octant, for a total of $c_1P^2N^3$ operations. The aggregation step (Step 3) requires c_3N^3 operations. The total operation count is thus $c_1P^2N^3 + 4c_2P^2N + c_3N^3$. The exact hierarchical decomposition of Theorem 1 leads to a slow $O(N^5)$ algorithm, as claimed. However, it is the basis of the fast algorithm we will develop in the next section.

III. Fast Hierarchical Backprojection

A. Sampling the 3-D Radon Transform

To construct the fast algorithm, we note that in the decomposition of Theorem 1, the backprojections of the octants are of the form $\mathbf{B}_{P,N/2}$ of P^2 projections onto $N/2\times N/2\times N/2$ volumes, while the original backprojection was $\mathbf{B}_{P,N}$, i.e., P^2 projections onto $N\times N\times N$ volumes. A similar situation occurs in the 2-D case. There, the backprojection of P projections onto an $N\times N$ image is decomposed into a sum of four backprojections of P projections onto $N/2\times N/2$ sized images, each centered at the origin. Next, we recall the sampling result of [11], which states that the number of angular samples needed to characterize the radon transform is proportional to the linear size of the image. Using this result, we argued in [10] that the sinograms could be angularly decimated by a factor of two prior to backprojection.

To make a similar argument in the 3-D case, we require a result analagous to the sampling result of [11]. In particular, we require a "scaling rule" that states that the angular sampling interval $\Delta\phi = \Delta\theta = 2\pi/P$ is inversely proportional to the size of the object. Some recent work has established the spectral support of the radon transform, providing a "bow-tie" like result for the 3-D problem [12], [13]. Sampling results for essentially bandlimited functions have also been available in [1]. In Appendix B, we use these sampling results to derive the required scaling rule: that the angular sampling interval for the separable sampling pattern is inversely proportional to the object radius. Also, note that the scaling rule is significantly more general. The total number of projections required depends on the object radius, regardless of the sampling pattern (provided it satisfies certain technical conditions). Thus, for a fixed object bandwidth, the sampling intervals are inversely proportional to the object size, and hence $P = O(N)$ samples are sufficient to characterize the 3-D radon transform of a N sized volume.

B. Single Stage

Using $P = O(N)$, it follows that, because the backprojections of the octants in Theorem 1 have the form $\mathbf{B}_{P,N/2}$, the corresponding projections can be angularly decimated by a factor of four (a factor of two in ϕ and a factor of two in θ) prior to backprojection, introducing only a small error. Hence, to construct the fast algorithm, we will replace the operator $\mathbf{B}^{\boldsymbol{\tau}}_{P,N/2}$ in Theorem 1 by $\mathbf{B}^{\boldsymbol{\tau}}_{P/2,N/2}\mathbf{O}$, where $\mathbf{O}$ is an angular smoothing and downsampling operator. It maps a set of P^2 radially discrete plane integral projections to a set $(P/2)^2$ radially discrete projections. Thus, *internally*, $\mathbf{B}^{\boldsymbol{\tau}}_{P/2,N/2}\mathbf{O}$ will have additional steps not appearing in the fours-step definition of $\mathbf{B}^{\boldsymbol{\tau}}_{P,N/2}$, although *functionally* it will be equivalent to it (to a good approximation).

Note that the argument about angular sampling (and decimation) of the 3-D radon transform applies to the radially continuous rather than radially discrete data. Therefore, the angular decimation step must be "sandwiched" between radial interpolation and resampling steps. Thus, we construct $\mathbf{O}$ using the following sequence of steps.

1) The radially discrete projections $g(k,m,n)$ are interpolated to radially continuous projections $g_c(r,m,n)$ with $0 \le m,n < P$.
2) The radially continuous projections are angularly decimated by a factor of two with respect to both ϕ and θ to form $g'_c(r,m,n)$ with $0 \le m,n < (P/2)$.
3) The radially continuous projections $g'_c(r,m,n)$ are resampled for input to $\mathbf{B}_{P/2,N/2}$.

The first step is accomplished, just as in the construction of $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}$, by $\mathcal{I}_{\boldsymbol{\tau}}$. For the second step, we define an angular decimation operator $\mathcal{D} : L_2^{P\times P}(\mathbb{R}) \mapsto L_2^{P/2\times P/2}(\mathbb{R})$, which is a convolutional angular smoothing followed by angular downsampling. The angular smoothing is given by

$$\hat{g}_c(r,m,n) = \sum_i \sum_j \rho(i,j) g_c[r,(m-i) \bmod P,(n-j) \bmod P] \quad m,n \in \{0,\ldots,P-1\} \tag{21}$$

where the function ρ is the smoothing kernel that reduces the aliasing errors in the angular decimation step, and the $\bmod P$ accounts for the 2π periodicity of g_c with respect to the spherical angles. The summation over i and j extends over the support of ρ. Angular downsampling retains only the even-indexed samples $\hat{g}_c(r,2m,2n)$, so that the decimation operator is defined by

$$\mathcal{D}g_c(r,m,n) = \sum_i \sum_j \rho(i,j) g_c[r,(2m-i) \bmod P,(2n-i) \bmod P] \quad m,n \in \left\{0,\ldots,\frac{P}{2}-1\right\}. \tag{22}$$

The final step, which converts the radially continuous projections back to radially discrete projections for input to $\mathbf{B}^{\boldsymbol{\tau}}_{P/2,N/2}$, is accomplished by simply resampling the projections radially using a resampling operator $\mathcal{G} : L_2^{P/2\times P/2}(\mathbb{R}) \mapsto \ell_2^{P/2\times P/2}(\mathbb{Z})$, defined by

$$\mathcal{G}_{\boldsymbol{\tau}} g_c(k,m,n) = g_c((k+\tau_{2m,2n})T,m,n).$$

Combining $\mathcal{G}$, $\mathcal{D}$, and $\mathcal{I}$, we arrive at the following definition of $\mathbf{O} : \ell_2^{P\times P}(\mathbb{Z}) \mapsto \ell_2^{P/2\times P/2}(\mathbb{Z})$

$$\mathbf{O}g(k,m,n) = \mathcal{G}_{\boldsymbol{\tau}}\mathcal{D}\mathcal{I}_{\boldsymbol{\tau}} g(k,m,n)$$

where $\mathbf{O}$ is defined by

$$\mathbf{O}g(k,m,n) = \sum_i \sum_j \rho(i,j) \sum_l g[j,(2m-i) \bmod P,(2n-j) \bmod P] \cdot \psi((k-l+\tau_{2m,2n}-\tau_{2m-i,2n-j})T). \tag{23}$$

The expression simplifies further in the special case that ρ is a 3×3 smoothing kernel, so that the summation only extends over $i,j \in \{-1,0,1\}$. In this case the $\bmod P$ operation is inactive, and can be dropped.

In interpreting the algorithm, it is important to note the role of the smoothing step in the decimation operator $\mathcal{D}$. In addition to preventing aliasing in the downsampling step, it also ensures that information from all the projections is retained in those kept after downsampling. Otherwise, the performance of the algorithm in the presence of noise could suffer. Indeed, the ideal anti-aliasing smoothing filter for a factor 2 decimation in each coordinate is an ideal 2-D low-pass filter with cutoffs at radian frequencies of $\pi/2$. Assuming that the projections are corrupted with white noise, the noise variance in each projection after decimation is reduced by a factor of $\pi^2/(\pi/2)^2 = 4$. Because the number of projections has been reduced by the same factor, the signal to noise ratio is not affected.

To demonstrate how $\mathbf{O}$ is incorporated into the hierarchical decomposition of the backprojection, we first recall the decomposition of Theorem 1

$$\mathbf{B}^{\boldsymbol{\tau}}_{P,N} g(\boldsymbol{z}) = \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{B}^{\langle \boldsymbol{\nu}_i \rangle}_{P,N/2} \hat{\mathbf{K}}_{N/2} \hat{\mathbf{Z}}_{-[\boldsymbol{\nu}_i]} g(\boldsymbol{z}). \tag{24}$$

We now make the approximation

$$\mathbf{B}^{\langle \boldsymbol{\nu}_i \rangle}_{P,N/2} \approx \mathbf{B}^{\langle \boldsymbol{\nu}_i \rangle \downarrow 2}_{P/2,N/2} \mathbf{O} \tag{25}$$

where for $\boldsymbol{x} \in \mathbb{R}^{P\times P}$, $\boldsymbol{x} \downarrow 2 \in \mathbb{R}^{P/2\times P/2}$ is $\boldsymbol{x}$ downsampled by two, i.e., $(\boldsymbol{x} \downarrow 2)_{p,q} = x_{2p+1,2q+1}$. Substituting (25) into (24) yields

$$\mathbf{B}^{\boldsymbol{\tau}}_{P,N} \approx \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{B}^{\langle \boldsymbol{\nu}_i \rangle \downarrow 2}_{P/2,N/2} \mathbf{O} \hat{\mathbf{K}}_{N/2} \hat{\mathbf{Z}}_{-[\boldsymbol{\nu}_i]}.$$

One final application of Property 1 yields

$$\mathbf{B}^{\boldsymbol{\tau}}_{P,N} \approx \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{B}^{\langle \boldsymbol{\nu}_i \rangle \downarrow 2}_{P/2,N/2} \hat{\mathbf{K}}_{N/2} \mathbf{O} \hat{\mathbf{K}}_{N/2} \hat{\mathbf{Z}}_{-[\boldsymbol{\nu}_i]}. \tag{26}$$

Defining the operator $\mathbf{A}^i_{P,N} : \ell_2^{P\times P}(\mathbb{Z}) \mapsto \ell_2^{P/2\times P/2}(\mathbb{Z})$ as

$$\mathbf{A}^i_{P,N} = \hat{\mathbf{K}}_{N/2} \mathbf{O} \hat{\mathbf{K}}_{N/2} \hat{\mathbf{Z}}_{-[\boldsymbol{\nu}_i]}$$

the decomposition is compactly rewritten as

$$\mathbf{B}^{\boldsymbol{\tau}}_{P,N} \approx \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{B}^{\langle \boldsymbol{\nu}_i \rangle \downarrow 2}_{P/2,N/2} \mathbf{A}^i_{P,N}. \tag{27}$$

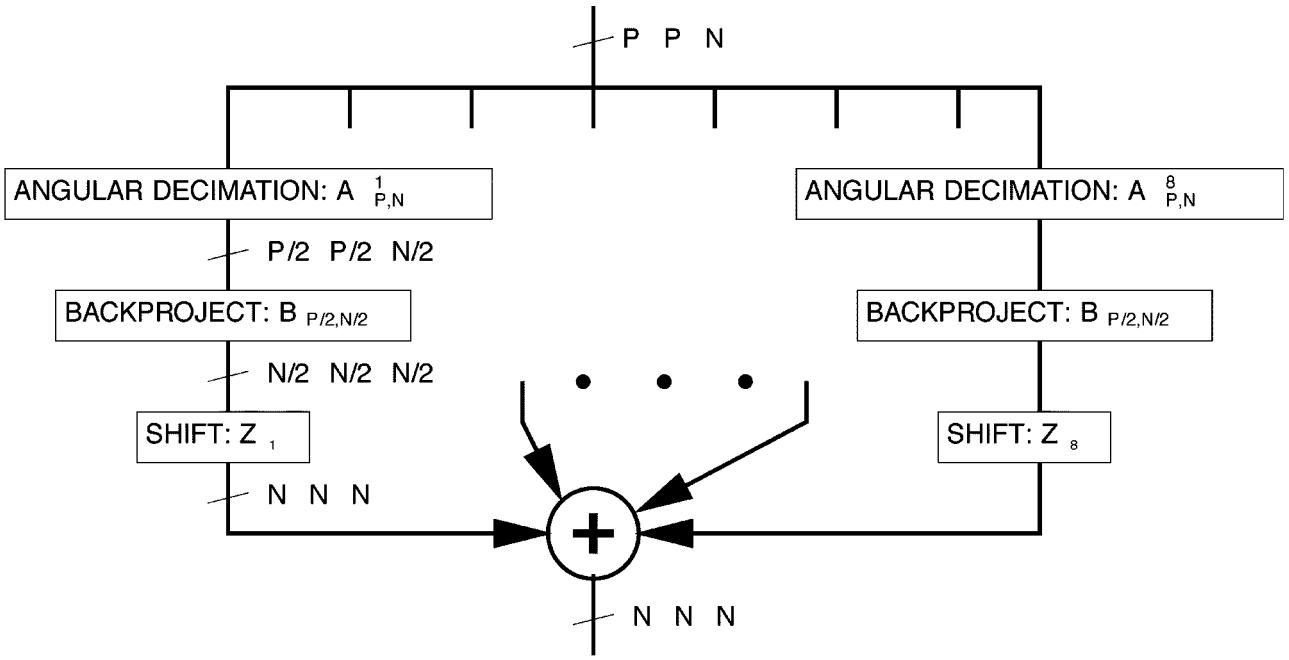

Fig. 3. Diagram of the decomposition of (26).

The decomposition of (26) can be described as the following sequence of steps.

1) *Approximate Subdivision:* Subdivide the projection data into eight sets, one for each octant of the original reconstruction, by computing $g_i' = \mathbf{A}_i g$. This corresponds to discrete radial shifting, truncation, radial interpolation and resampling, angular smoothing and downsampling.
2) *Backprojection:* Backproject $P^2/4$ projections contained in each g_i' to form f_i, the reconstruction of the ith octant.
3) *Aggregation:* Shift f_i to the appropriate octant of the final reconstruction, and add into place.

This sequence is diagrammed in Fig. 3. Comparison with Fig. 2 reveals the addition of the angular resampling steps, which form the basis of the approximation.

To determine the computational cost of using (26) to compute $\mathbf{B}^{\boldsymbol{\tau}}_{P,N}$, we make the same assumption as we did in the previous section, i.e., that $\mathbf{B}^{\boldsymbol{\tau}}_{P,N} g$ requires $c_1 P^2 N^3$ operations. The approximate subdivision step (Step 1) requires $c_4(P/2)^2(N/2)$ operations per octant, where c_4 depends on the choice of ρ and ψ in (23), as well as the choice of (11), (12) or (13) in the definition of $\hat{\mathbf{K}}_M$. We arrive at this figure by noting that $A^i_{P,N} g$ is nonzero for $O(P^2N/8)$ points, and that each point (within indexing), requires a constant number of operations for a fixed choice of algorithm parameters (ρ, ψ, b, T, etc.). Thus, the total cost for Step 1 for all octants is $c_4 P^2 N$. The cost of the backprojections (Step 2) is $c_1(P/2)^2(N/2)^3$ operations per octant for a total of $c_1 P^2 N^3/4$. Finally, the cost of the aggregation step (Step 3) is still $c_3 N^3$ operations. The total cost for (26) is thus $c_1 P^2 N^3/4 + c_4 P^2 N + c_3 N^3$ operations, for roughly a factor of four speedup over (9) for large N, P.

C. Recursive Form

The savings from the use of (26) for a single stage of the decomposition as described in the previous subsection, is a factor of four in computation. To realize the $N^2/\log_2 N$ speedup, we must apply (26) in a recursive manner, i.e., replace the backprojections in Step 2 with (26). Equivalently, we define a fast backprojection operator in a recursive manner with the following pair of equations[5]

$$\tilde{\mathbf{B}}^{\boldsymbol{\tau}}_{P,N} = \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \tilde{\mathbf{B}}^{\langle \boldsymbol{\nu}_i \rangle \downarrow 2}_{P/2,N/2} \mathbf{A}^i_{P,N}$$
$$\tilde{\mathbf{B}}^{\boldsymbol{\tau}}_{P,1} = \mathbf{B}^{\boldsymbol{\tau}}_{P,1}. \tag{28}$$

Fig. 4 shows the application of (28) for the case $N = 4$.

To determine the computational cost associated with (28), we assume that the backprojection $\mathbf{B}_{P,1}$ of $P = O(N)$ projections into a single voxel, requires constant time c_5. The cost of the algorithm is then dominated by the computation of $\mathbf{A}_i$ and $\mathbf{Z}_{\delta_i}$ at each level of the decomposition. Again, assuming that $P = O(N)$, these steps require $c_6 N_k^3$ operations, where N_k is the size of an octant at the kth level in the decomposition. Now, there will be a total of 8^k octants at level k in the decomposition, each of size $N/2^k$. Thus, the total operation count per level of the reconstruction is a constant $c_6 8^k (N/2^k)^3 = c_6 N^3$ operations. There are a total of $\log_2 N$ levels, so the total cost of the algorithm is simply $c_6 N^3 \log_2 N = O(N^3 \log_2 N)$, where c_6 includes implementation-dependent factors.

The storage requirements of the fast algorithm can also be computed. The FBP algorithm requires N^3 space to store the reconstruction volume and a single projection. A straightforward implementation of the fast algorithm requires $15/7N^3$ space, as the reconstruction volume computations can be done in place (thereby eliminating the need to explicitly shift the octants and add them back in place), and the projection data manipulations require at most $15/7N^3$ space. To explain this last fact, note that

[5]This type of recursive definition occurs frequently in signal processing applications, and is one way to approach the construction of the fast Fourier transform (see, [14]).

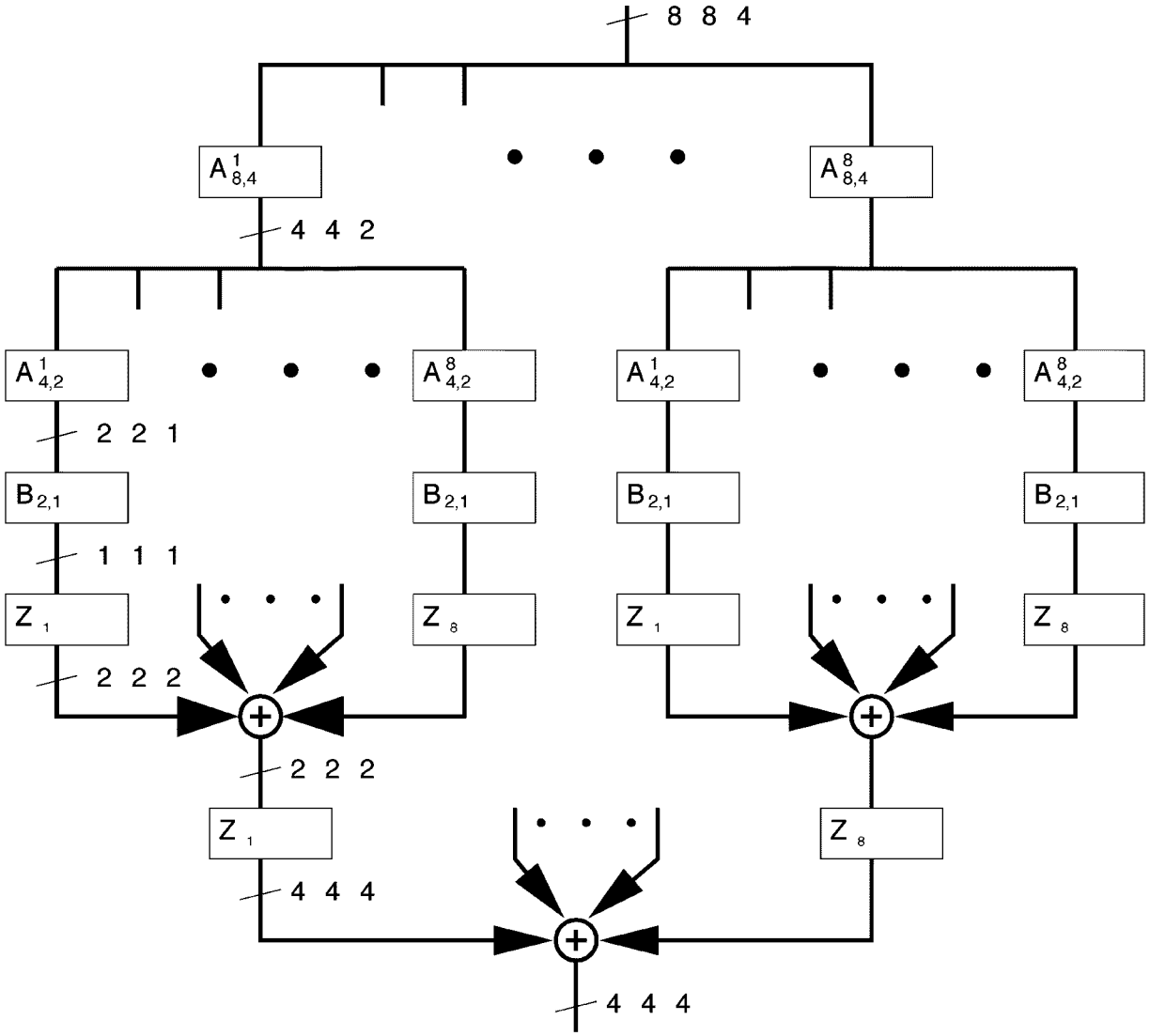

Fig. 4. Diagram of the computation of $\bar{\mathbf{B}}\mathbf{T}_{P,N}$ for $N = 4$ and $P = 8$.

the original projection data requires N^3 space. For the decomposition, space is required to store the projection data for an octant (i.e., $\hat{\mathbf{K}}\hat{\mathbf{Z}}g$). This is at most $N^3/8$. Due to the recursion, an additional $(N/2^k)^3$ space will be required at each level of the recursion, but all computations at a given level can be done using the same space (in the form of a stack). Thus, the total space requirements in terms of the projection data are at most (for N approaching infinity) $N^3(1+1/8+1/64+\cdots) = 8/7N^3$. Thus, in principle, the fast algorithm requires only slightly more space than a straightforward FBP algorithm.

D. Upsampling

While the recursive application of (26) leads to an $O(N^3 \log_2 N)$ algorithm, it has been our experience in 2-D that artificially increasing the radial and angular sampling rates improves the accuracy of the reconstruction, at the cost of a higher operation count. Following our approach in [10], we propose to increase the radial sampling rate by simply resampling the projections with a higher sampling rate by means of the interpolation kernel. Thus, we define a radial upsample-by-C operator $\mathbf{U} : \ell_2^{P\times P}(\mathbb{Z}) \mapsto \ell_2^{P\times P}(\mathbb{Z})$ by

$$\mathbf{U}g(k,m,n) = \sum_l g(l,m,n)\psi(kT/C - lT).$$

The radial upsampling is performed prior to calling the fast reconstruction algorithm, with T replaced by T/C. Although our complexity analysis ignored the exact influence of T on the cost, its trivial to show that the cost of the algorithm increases linearly with C. The space requirements also increase linearly in C.

For angular upsampling, however, we take a different approach. If we simply angularly upsample the data by factor C in both the ϕ and θ directions, the storage requirements of the algorithm increase *quadratically* in C. Thus, if we wish to angularly upsample by a factor of $C = 4$ in each direction, the space requirements increase by a factor of 16. This is inefficient, in light of the following, more elegant approach.

To achieve angular upsampling, let us recall Theorem 1. There, a backprojection of an N^3 object using P^2 projections

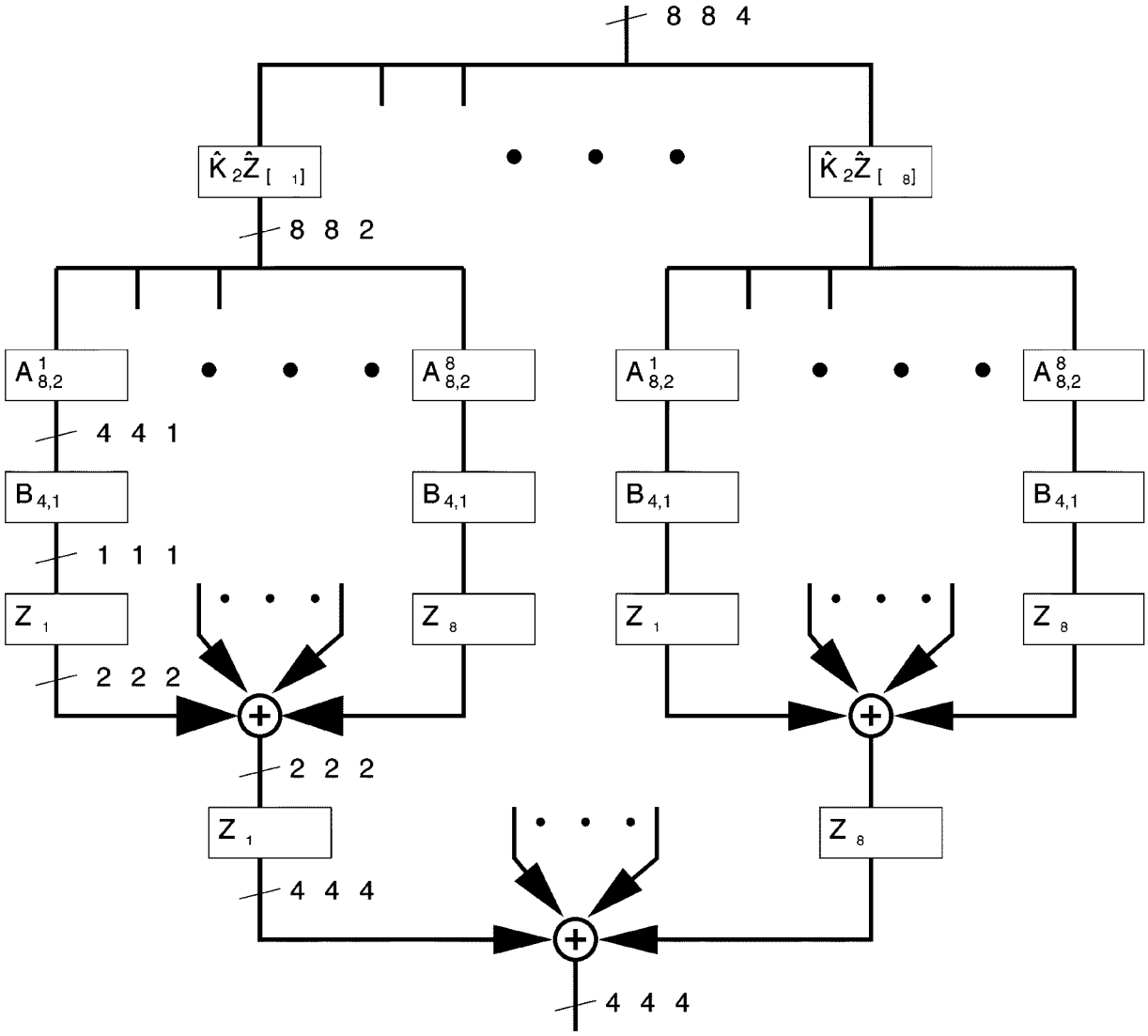

Fig. 5. A simplified diagram of the recursive decomposition when one stages of Theorem 1 is used prior to decomposing via (27).

was decomposed into 8 backprojections of $N^3/8$ objects using the same number of projections. Hence, these smaller backprojections are effectively angularly upsampled by a factor of 2 with respect to both ϕ and θ. By applying Theorem 1 for a total of Q times, we decompose $\mathbf{B}_{P,N}$ into 8^Q backprojections, each of the form $\mathbf{B}_{P,N/2^Q}$. Thus, we can increase the angular sampling rate by $C = 2^Q$ with respect to both ϕ and θ. Fig. 5 illustrates this process for $N = 4$ when Theorem 1 is used once, and (27) is used for the remaining subdivisions. Comparing Figs. 5 and 4, we see that the number of projections used in the single voxel backprojections has increased fourfold by a single application of Theorem 1. Furthermore, an analysis of the space requirements reveals that the space requirements remain $15/7N^3$. So in terms of space requirements, this is a very efficient way of boosting the angular sampling rate. A flowchart of the algorithm is given in Fig. 6.

Of course, the cost requirements increase dramatically with Q. For Q small, relative to N, the operations in Theorem 1 are negligible, and the dominating cost is the evaluation of 8^Q backprojections, each of P^2 projections onto $(N/2^Q)^3$ volumes. If these backprojections are performed using the recursive fast algorithm, then the dominant term in the total cost is $O(4^Q P^2 N \log_2 N) = O(4^Q N^3 \log_2 N)$. Of course for Q fixed, the factor of 4^Q is negligible as N becomes large. Note that 4^Q is the same complexity factor incurred by simply upsampling the data prior to invoking the algorithm.

There is a degree of flexibility in terms of the ordering of the various stages of the decomposition. For example, we have described (and implemented) the scheme in which Q applications of Theorem 1 are followed by $\log_2 N - Q$ applications of (26). But the ordering of the steps is arbitrary. Further research is necessary to determine the optimal placement of the Q exact decompositions among the $\log_2 N$ stages.

IV. Simulations

In this section, we present some simulation studies that we have performed on our fast backprojection algorithm, using the Shepp-Logan 3-D head phantom [8]. To use our fast backprojection algorithm for reconstruction, it is first necessary to radially filter the projections with an approximate second-order derivative kernel. We have used the standard second order difference

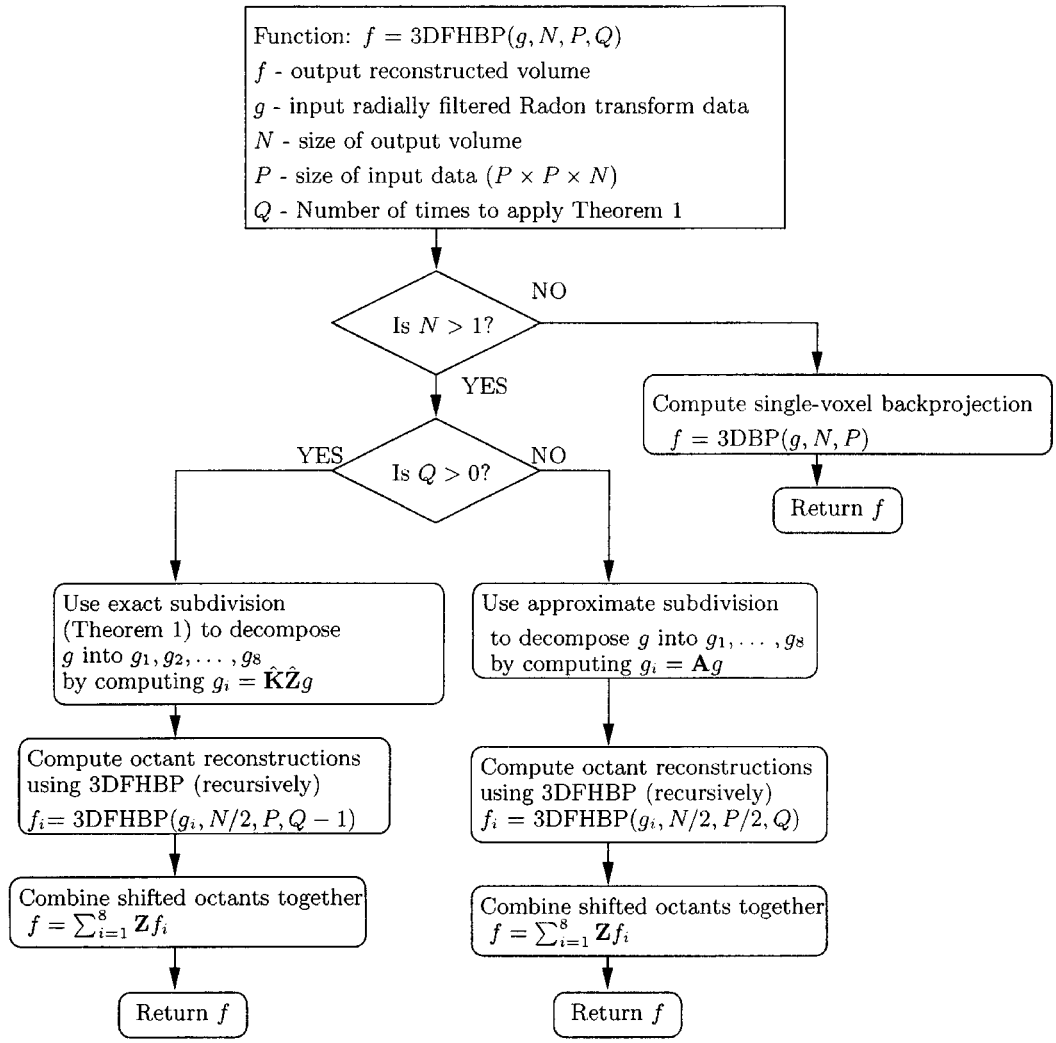

Fig. 6. A flow chart of the proposed fast 3-D backprojection algorithm. The routine $f = 3DBP(g, N, P)$ is a single-voxel backprojection routine. The operator subscripts have been suppressed for readability.

TABLE I
PARAMETERS USED IN THE 3-D-FHBP-AND FBP-BASED RECONSTRUCTIONS

Parameter	Meaning	Value used
β	Filter in FBP	[-1 2 -1]
N	Volume size	256
P	Number of projections in each spherical direction	256
T	Radial sampling interval	0.5
Q	Number of times Theorem 1 was applied	2
$b(\boldsymbol{x})$	Volume sampling function	Indicator for $\|\boldsymbol{x}\|_2 \leq \frac{1}{2}$
ρ	Smoothing filter for angular downsampling	$\begin{bmatrix} .25 & .5 & .25 \\ .5 & 1 & .5 \\ .25 & .5 & .25 \end{bmatrix}$
C	Radial upsampling factor	2

kernel, as described in [6]. Synthetic plane-integral projections were computed for $P = 256$, and the reconstruction volume was size $N = 256$. The detector spacing was set to $T = 0.5$. The choices for the various parameters are summarized in Table I.

In Fig. 7, a montage of slices from the original phantom, the FBP reconstruction and the 3-D-FHBP reconstruction are shown. The left hand column of Fig. 7 contains slices from the Shepp-Logan 3-D head phantom. The top slice is in the xy plane at $z = -0.125$, the middle slice is in the xz plane with $y = 0.25$, and the bottom image is in the yz plane with $x = 0$. The center column consists of the same slices from the FBP reconstruction. The rightmost column consists of slices from the reconstruction method proposed in this paper. The algorithm was run with a radial oversampling factor of 2, and $Q = 2$ levels of the exact decomposition. For all images, the grayscale display was set to the range $[1, 1.04]$ to show the small features of the phantom.

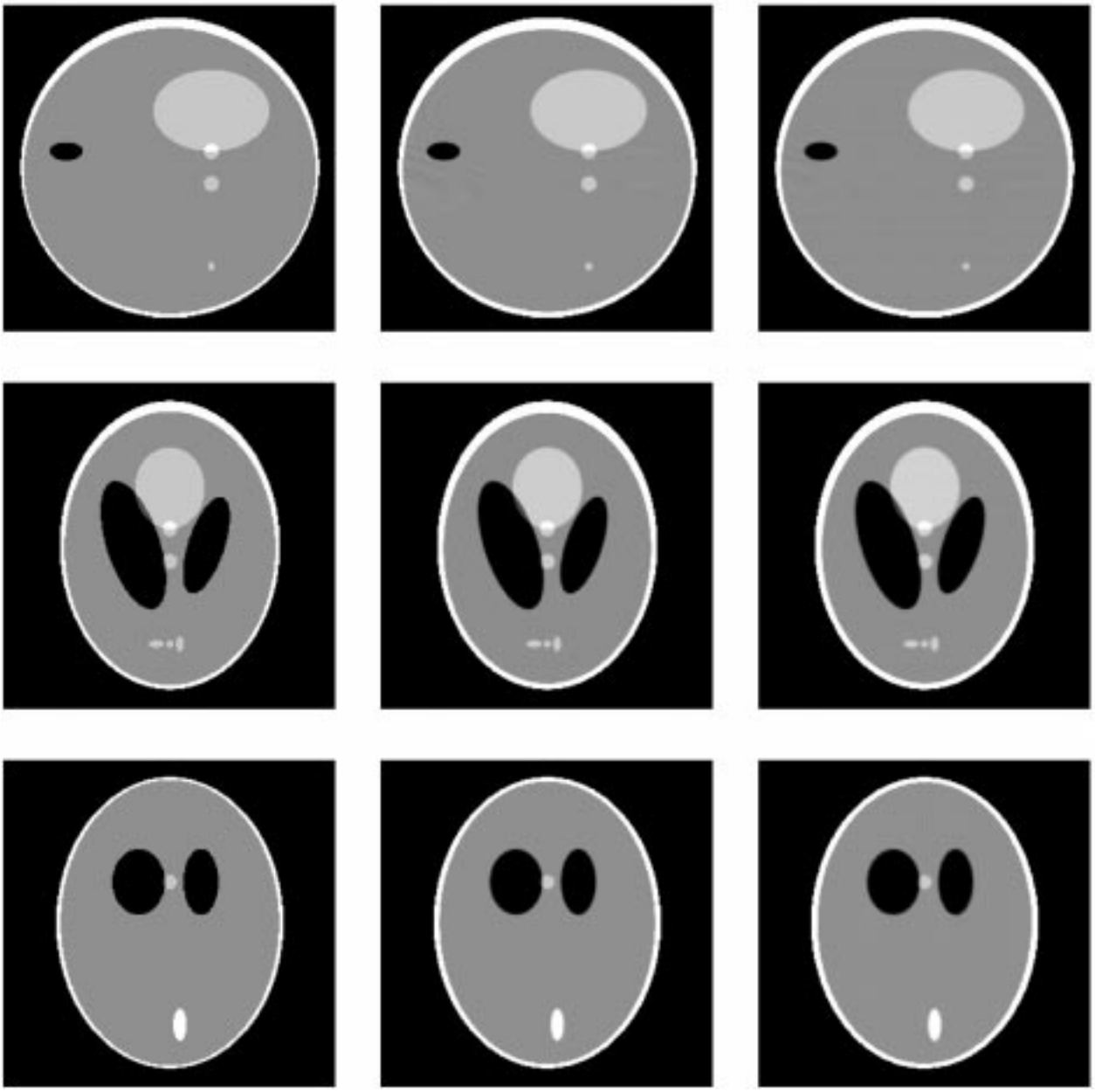

Fig. 7. A montage of slices through the phantom (lefthand column), FBP reconstruction (center column) and 3-D-FHBP recosntruction (right column). The top row is an xz slice through the reconstructions, the middle row is a yz slice, and the bottom row is a xy slice. The 3-D-FHBP reconstruction algorithm was 200 times fast than the FBP in this instance.

In Fig. 8, plots of rows and columns of these images are shown for comparison. The top plot is through column 160 of all three xy images. The middle plot is through row 206 of all three xz images, and the bottom plot is through column 96 of all three images. Note that both the FBP and the 3-D-FHBP reconstructions accurately recover the small details present, including the sharp edges.

Based on run times for these simulations, we estimate that a complete FBP reconstruction of the 256^3 volume would take approximately[6] 3000 hours on a single central processin unit (CPU) from an SGI Origin 2000. With the stated parameters, the 3-D-FHBP reconstructions took 15 hours on a single CPU of this machine. We expect the speedups to become more significant for progressively larger reconstruction volumes.

To provide a more quantitative comparison of the 3-D-FHBP and FBP reconstructions, we computed point spread functions (PSFs) for both algorithms using the same parameters as above. Fig. 9 depicts PSFthe spherically averaged PSF for a point at the isocenter. Note that even though the test point was at the isocenter, this is not a special point for the FHBP algorithm. In the process of performing the backprojection, the octant containing the isocenter is repeatedly shifted so that the center of that octant is properly centered. As a result, the sinogram undergoes significant shifting and nontrivial angular decimations. The good match between the psfs provides additional evidence

[6]This figure is an estimate for two reasons. First, our implementation was a parallel implementation. Second, we only reconstructed the specific slices of the FBP reconstruction of interest, instead of the entire volume.

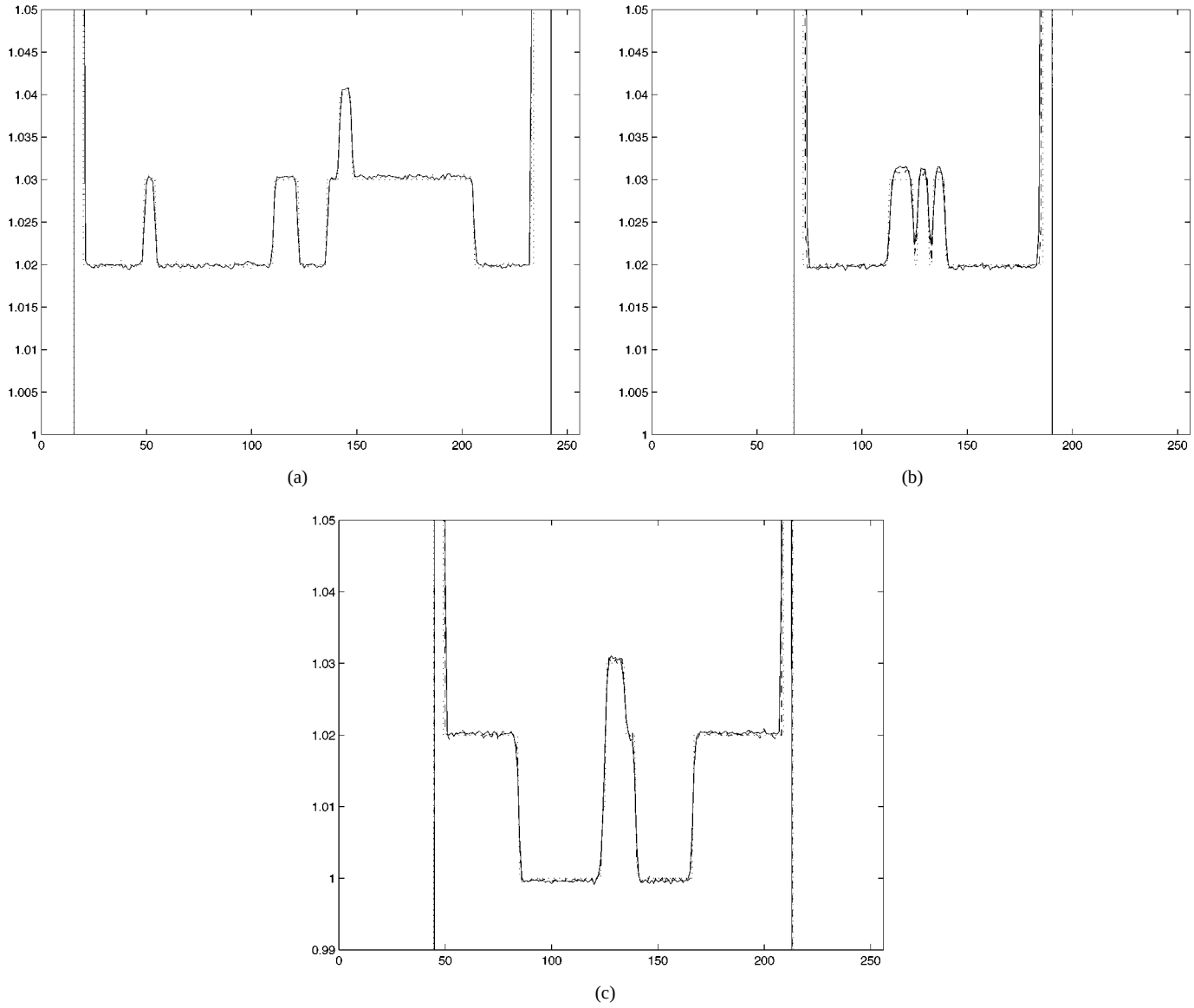

Fig. 8. Cuts through rows and columns of the images in Fig. 7. (a) column 160 of the top row of images in Fig. 7. (b) row 220 of the middle row of images in Fig. 7. (c) row 186 of the bottom row of images in Fig. 7. In each plot, the dotted line represents the phantom, the broken line the FBP reconstruction, and the solid line the 3-D-FHBP-based reconstruction.

that the FHBP algorithm provides high-quality reconstructions with a significant speedup over direct reconstruction.

V. Conclusion

We have demonstrated a novel 3-D backprojection algorithm with $O(N^3 \log_2 N)$ cost that uses a hierarchical, recursive decomposition of the backprojection operation to achieve significant speedups over direct computations. Experiments with a head phantom show good performance, orders of magnitude faster than the direct FBP. The algorithm may be used to obtain an $O(N^3 \log_2 N)$ reconstruction algorithm for cone-beam tomogrphy, by combining it with a fast hierarchical 2-D reprojection algorithm [15] to compute the 2-D radon transforms arising in a a Granget or Smith transformation from cone-beam to 3-D radon data. Future research will address the evaluation of the resulting cone-beam algorithm, as well as the issue of direct comparisons with the methods of [6]–[8].

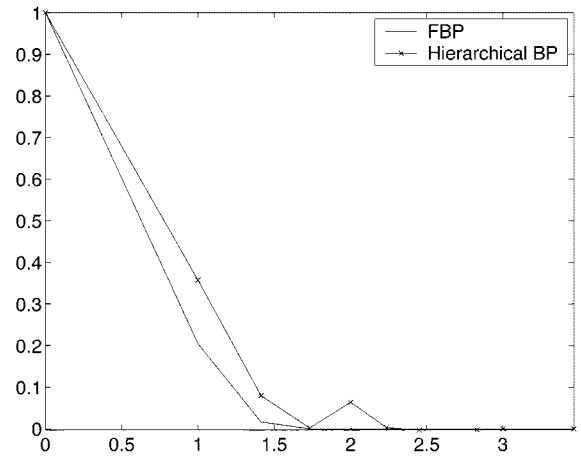

Fig. 9. A comparison of PSF for the FHBP and FBP algorithms for the isocenter point. The PSFs were spherically averaged to produce these plots.

APPENDIX A
PROOF OF THEOREM 1

Proof: For $f = \mathbf{B}_{P,N} g$

$$
\begin{aligned}
f &= \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{K}_{N/2} \mathbf{Z}_{-\boldsymbol{\delta}_i} f \\
&= \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{K}_{N/2} \mathbf{Z}_{-\boldsymbol{\delta}_i} \mathbf{B}_{P,N} g \\
&= \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{K}_{N/2} \mathbf{Z}_{-\boldsymbol{\delta}_i} \mathbf{K}_N \mathcal{S} \mathcal{B}_P \mathcal{I}_{\boldsymbol{\tau}} g \\
&= \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{K}_{N/2} \mathcal{S} \mathcal{M}_{-\boldsymbol{\delta}_i} \mathcal{B}_P \mathcal{I}_{\boldsymbol{\tau}} g \quad \text{Property 3} \\
&= \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{K}_{N/2} \mathcal{S} \mathcal{B}_p \hat{\mathcal{M}}_{-\boldsymbol{\delta}'_i} \mathcal{I}_{\boldsymbol{\tau}} g \\
&= \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{K}_{N/2} \mathcal{S} \mathcal{B}_P \mathcal{I}_{\langle \boldsymbol{\nu}_i \rangle} \hat{\mathbf{Z}}_{-[\boldsymbol{\nu}_i]} g \quad \text{Property 2 and 4} \\
&= \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{B}_{P,N/2} \hat{\mathbf{Z}}_{-[\boldsymbol{\nu}_i]} g \\
&= \sum_{i=1}^{8} \mathbf{Z}_{\boldsymbol{\delta}_i} \mathbf{B}_{P,N/2} \hat{\mathbf{K}}_{N/2} \hat{\mathbf{Z}}_{-[\boldsymbol{\nu}_i]} g \quad \text{Property 1}
\end{aligned}
$$

where $(\boldsymbol{\delta}'_i)_j = \boldsymbol{\delta}_i \cdot \boldsymbol{\omega}_j$. □

APPENDIX B
SAMPLING THE 3-D RADON TRANSFORM

In this appendix, we consider the problem of sampling the 3-D radon transform. Our goal is to find a "scaling" rule that relates the angular sampling interval to the object radius (provided the bandwidth of the object is fixed). The rule can be easily derived from [1, Theorem 2.4, Ch. III]. Let f be an essentially bandlimited function with essential bandlimit b (meaning that the L_1 norm of the spectrum of f is negligible outside a ball of radius b), and assume f is supported on the unit ball (i.e., the radius of support of f is unity). Then it follows from [1, Theorem 2.4] that f can be reliably recovered from samples of its 3-D radon transform at a set of angles A_m provided it satisfies the following two conditions.

1) A_m is m-resolving, meaning that no nontrivial trigonometric polynomials of a special form vanish on A_m (see [1, page 69]).
2) $m > b$.

For the 3-D radon transform, condition 1) is satisfied for almost all sets of $(m+1)(m+2)/2$ angles. In particular, for the separable sampling pattern with $(m+1)^2$ directions defined by $\Delta\phi = \Delta\theta = \pi/(m+1)$, it can be shown that the resulting set A_m is m resolving. Condition 2) provides the scaling rule. To see how, let f be an essentially bandlimited function of unit bandwidth, with radius R. If we let $f'(x) = f(x/R)$, then f' is an essentially bandlimited function of bandwidth R with unit radius. Furthermore, if $g'_k(s)$ is a plane-integral projection of f' at direction (θ_k, ϕ_k), then by rescaling we can recover $g_k(s)$, the plane-integral projection of f in the same direction. Thus, if we can recover f' from its projections $g'_k(s)$, then we can recover f from from its projections $g_k(s)$. The scaling rule is completed by applying condition 2), so that $m > R$ is sufficient to guarantee recovery of both f and f'.

REFERENCES

[1] F. Natterer, *The Mathematics of Computerized Tomography*. New York: Wiley, 1986.

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

[3] R. V. D. Walle, T. Voet, and I. Lemahien, "The motion sensitivity of magnetic resonance imaging: A comparison between fourier transform imaging and projection reconstruction," in *Proc. 1995 IEEE Engineering in Medicine and Biology 17th Annu. Conf.*, vol. 1, 1997, pp. 457–458.

[4] P. Grangeat, "Mathematical framework of cone bean 3-D reconstruction via the first derivative of the Radon transform," in *Lecture Notes in Mathematics*, G. Herman, A. Louis, and F. Natterer, Eds. New York: Springer, 1991, vol. 1497, Mathematical Methods in Tomography, pp. 66–97.

[5] B. Smith and C. Peck, "Implementations, comparisons, and an investigation of hueristic techniques for cone-beam tomography," *IEEE Trans. Med. Imag.*, vol. 15, pp. 519–531, Aug. 1996.

[6] R. Marr, C. Chen, and P. Lauterbur, "On two approaches to 3-D reconstruction in NMR zeugmatography," in *Lecture Notes in Medical Informatics*. Berlin, Germany: Springer, 1981, Mathematical Apects of Computerized Tomography.

[7] C. Axelsson and P. Danielsson, "Three-dimensional reconstruction from cone-beam data in $O(N^3 \log N)$ time," *Phys. Med., Biol.*, vol. 39, pp. 477–491, Mar. 1994.

[8] S. Schaller, T. Flohr, and P. Steffen, "An efficient fourier method for 3-D Radon inversion in exact cone-beam CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 17, pp. 244–250, Apr. 1998.

[9] H. Turbell, "Three-dimensional image reconstruction in circular and helical computed tomography," Ph.D. dissertation, Dept. Elect. Eng., Linkoping Univ., Linkoping, Sweden, 1999.

[10] S. Basu and Y. Bresler, "$O(N^2 \log_2 N)$ filtered backprojection reconstruction algorithm for tomography," *IEEE Trans. Image Processing*, vol. 9, pp. 1760–1773, Oct. 2000.

[11] P. A. Rattey and A. G. Lindgren, "Sampling the 2-D Radon transform," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-29, pp. 994–1002, May 1981.

[12] V. Palamodov, "Localization of harmonic decomposition of the Radon transform," *Inverse Prob.*, vol. 11, pp. 1025–1030, Oct. 1995.

[13] X. Pan, "Quasi-bandlimited properties of Radon transforms and their implications for increasing angular sampling densities," *IEEE Trans. Med. Imag.*, vol. 17, pp. 395–406, June 1998.

[14] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[15] A. Boag, Y. Bresler, and E. Michielssen, "A multilevel domain decomposition algorithm for fast $O(N^2 \log N)$ reprojection of tomographic images," *IEEE Trans. Image Processing*, vol. 9, pp. 1573–1582, Sept. 2000.