9-1-1993

Wavelet Localization of the Radon Transform

Tim Olson
Dartmouth College

Joe DeStefano
Dartmouth College

WAVELET LOCALIZATION OF THE RADON TRANSFORM

Tim Olson and Joe DeStefano

Technical Report PCS-TR93-196

9/93

WAVELET LOCALIZATION OF THE RADON TRANSFORM *

Tim Olson and Joe DeStefano
Department of Mathematics and Computer Science
Dartmouth College, Hanover, NH 03755
(603) 646 - 3173

ABSTRACT

In this paper we develop an algorithm which significantly reduces radiation exposure in x-ray tomography, when a local region of the body is to be imaged. The algorithm uses the properties of wavelets to essentially localize the Radon transform. The algorithm differs from previous algorithms for doing local tomography because it recovers an approximation to the original image, not the image modulo the nullspace of the local tomography operator, or the Lambda transform of the image. This is possible because we do not truly invert the interior Radon transform, but rather sample the Radon transform sparsely away from the local region of interest. Much attention in the field has been directed towards localized tomography. We believe that this technique represents a significant contribution towards this effort.

I. INTRODUCTION

The Radon transform was first introduced by J. Radon in 1917. Little computational attention was given to it until the advent of computers enabled the fast evaluation of Fourier transforms and their corresponding convolutions. The Radon transform is now a mainstay of medical imaging as well as many other remote imaging sciences.

One problem with the Radon transform is that in two dimensions, where most medical imaging is conducted, the inversion formula is globally dependent upon the line integrals of the object function f [6]. In many situations, a physician may only be interested in images of a very local area of the body. One would prefer to expose only that local portion of the patient's body to whatever radiation is being used. The non-locality of the Radon transform, however, forces one to expose to radiation all of the two-dimensional slices of the body which intersect with the region of interest.

There have been a number of attempts to alleviate this problem. One approach is to do truly three dimensional tomography, but this requires integrals over full two-dimensional hyperplanes which we are trying to avoid. A recent approach which has shown great promise involves reconstructing not the density function f, but rather another function $\Lambda f - \mu\Lambda^{-1} f$ which is similar to the original function in many ways [7, 8]. This approach is well adapted for edge detection but is it not designed to recover the original density of the image.

Although local values are not even uniquely determined by local integrals over hyperplanes, it has been proven [25, 6, 9], that the elements which are not uniquely determined do not vary wildly well within the region of interest. While this is true, the result is asymptotic in nature, and doesn't really apply to a very localized region. We will present examples which show that when the region of interest is truly local, the nullspace varies significantly within the region of interest, and therefore is significant to the reconstruction process.

Reconstructions modulo the nullspace of the interior Radon transform have been obtained and perform well as long as the nullspace of the operator does not contribute significantly to the image [14]. In contrast, our algorithm is not affected adversely by the elements of the nullspace, but rather recovers these elements in the region of interest.

It has been noted [6] that a rough estimate of the global properties of the Radon transform would be sufficient to obtain a reliable reconstruction of a local region. This is the foundation for our algorithm. We construct a rough estimate of the global properties of the Radon transform through efficient sparse sampling. This rough estimate allows us to produce a true reconstruction in the region of interest.

We will present an algorithm which uses the time-frequency localization properties of wavelets, as well as many of the properties of the Radon transform, to exactly quantify that portion of the non-local information which must be sampled for a reliable reconstruction. We will also present error bounds which specify the amount of non-local information that is necessary in order to recover an approximation to the image with a given accuracy.

In [13] it was suggested that wavelets be used for

*This work was supported in part by DARPA as administered by the AFOSR under contract AFOSR-90-0292.

the localization. We will be show that the properties which a basis must satisfy to localize the Radon transform via the filtered backprojection algorithm specify that wavelets or wavelet type time-frequency localized bases be used. A short summary of the following work may be found in [4].

Organization

In the next section we outline the goals of this paper. In section III we present standard definitions and background from computerized tomography. Section IV outlines the reasons why standard filtered backprojection is dependent on non-local values of the Radon transform. In section V the properties which make wavelets the natural basis functions with which to construct an algorithm for reconstruction from essentially localized data are presented. The algorithm is presented in section VI. In section VII we conduct an error and stability analysis of the algorithm. Section VIII presents the background to Lamda-tomography, and examines whether an algorithm similar to ours can be generated from Lambda-tomography. We present numerical examples of the algorithm, and a comparison of similar algorithms in section IX. Section X presents a modification of the algorithm for edge detection. Section XI examines the relative merits of several standard wavelets for our algorithm. Finally, we compare the qualities of various wavelets for use in this algorithm in section XII.

II. ALGORITHMIC GOALS

The goal of this paper is not to reconstruct the local values of f from local values of the Radon transform. The goal of this paper is to develop a stable algorithm which will reconstruct local values of f from local values of the Radon transform, together with sparsely sampled non-local values of the Radon transform. Thus when we refer to essentially localized data, we are referring to the line integrals which pass through the region of interest, and a small number of line integrals which do not pass through the region of interest.

We emphasize here that inversion of the Radon transform from only the line integrals which pass through a given region is not even well defined, and therefore one must either regularize the reconstruction or gather more data. Our approach is to gather enough data so that a well conditioned true inversion can be made, rather than perform an inversion modulo the nullspace of the operator.

III. BACKGROUND

We will review some of the basic properties of the Radon transform. For details and proofs, we recommend [6]. By S^1 we are simply referring to the one sphere or the unit circle. Throughout we will use the notation F_1 and F_2 to denote the Fourier transform on $L^2(\mathbb{R})$ and $L^2(\mathbb{R}^2)$, respectively, defined by

$$F_1(g)(s) \equiv \hat{g}(s) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(t) e^{-ist} dt$$

and

$$F_2(g)(u,v) \equiv \hat{g}(u,v) = (2\pi)^{-3/2} \int\int_{\mathbb{R}^2} g(x,y) e^{-i(xu+yv)} dx dy.$$

The Radon transform, which is sometimes referred to as the x-ray transform or the projection transform, of a function $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$P_f(\vec{\theta}, \vec{\theta} \cdot \vec{x}) = \int_{\mathbb{R}} f(\vec{x} + t\vec{\theta}^{\perp}) dt. \tag{1}$$

It follows that $P_f(\vec{\theta}, \vec{\theta} \cdot \vec{x})$ is the line integral of f through $\vec{x} \in \mathbb{R}^2$ and perpendicular to $\vec{\theta}$. We will often abuse this notation by associating the vector $\vec{\theta}$ with the angle θ, between $\vec{\theta}$ and the y-axis.

The interior Radon transform is the restriction of the Radon transform given by

$$P_f(\vec{\theta}, t) \to P_f(\vec{\theta}, t) \chi_{[-a,a]}(t), \tag{2}$$

where $|a| < 1$. Thus instead of knowing $P_f(\vec{\theta}, t)$ for $t \in [-1, 1]$ one only knows $P_f(\vec{\theta}, t)$ for $t \in [-a, a]$. One can view this as the restriction of the Radon transform to lines which pass through a circle of radius a about the origin, which we will subsequently refer to as the (local) region of interest.

The operator $P : f \to P_f(\vec{\theta}, \vec{\theta} \cdot \vec{x})$ has been extensively studied in many contexts, and an excellent overview of the literature may be found in [6]. A detailed study of the singular value decomposition of the interior Radon transform (2) can be found in [9, 14]. The results of [9, 14] confirm those in [6], and suggest that the problem of local reconstruction is not severely ill-conditioned, when the inversion is done modulo the nullspace. The singular values $\sigma_{m,l}$ of the operator, which are doubly indexed due to the spherical harmonics used in the construction, behave like $O(m^{-1/2})$ for l fixed, $O(m^{-1/4})$ for $m = l$, and $O(l^{-1/2})$ for fixed m [9].

We assume throughout that we are trying to recover a density function $f : \mathbb{R}^2 \to \mathbb{R}$ which has support in B_1, where B_1 is the unit disc centered at the origin in $\mathbb{R}^2$. Moreover we will view the Radon transform $P_f(\vec{\theta}, t)$ as an element of either the weighted Chebyshev space $L^2(S^1 \times [-1, 1], dx \times (1-t^2)^{-1/2} dt)$ or as elements of L^2 with the usual inner product. We will use the notation $\| \cdot \|_{2,w}$ for the norm in the

weighted space, and $\|\cdot\|_2$ for the norm in L^2 with the usual inner product. We will make use of the following theorem characterizing the range of the Radon transform, which can be found in [6].

Theorem 1 *The Radon transform of a function $f \in L^2(\mathbb{R})$, $P_f(\vec{\theta},t) \in L^2\left[S^1 \times [-1,1]\right]$ can be represented in the form*

$$P_f(\vec{\theta},t) = (1-t^2)^{-1/2}\sum_{l=0}^{\infty} T_l(t)h_l(\theta) \quad (3)$$

where the $T_l(t)$ are Chebyshev polynomials of the first kind,

$$h_l(\vec{\theta}) = \sum_{n\in H_l} a_{l,n}e^{in\theta}$$

and $H_l = \{-l, -l+2 \ldots l-2, l\}$.

One can interpret this theorem by realizing that when viewed as a function of $\vec{\theta}$, for fixed s, $\hat{f}(s\vec{\theta})$ will be essentially bandlimited with bandlimit s. Thus the angular bandlimit in polar coordinates is tied to the radial variable s. Note that in this work, we never need to compute this decomposition. We use this theorem only to characterize this angular bandlimit.

IV. RECONSTRUCTION TECHNIQUES

An essential property of the Radon transform is that the one-dimensional Fourier transform of $P_f(\vec{\theta},\vec{\theta}\cdot\vec{x})$ with respect to the variable $t \equiv \vec{\theta}\cdot\vec{x}$ is a line through the two-dimensional Fourier transform of f, i.e.

$$F_1\left(P_f(\vec{\theta},\vec{\theta}\cdot\vec{x})\right) = F_2(f)(s\vec{\theta}).$$

Using this fact and a polar Fourier inversion formula, one can easily relate the Radon transform $P_f(\vec{\theta},\vec{\theta}\cdot\vec{x})$ to $f(\vec{x})$ through the reconstruction formula

$$\begin{aligned} f(\vec{x}) &= \int_{S^1}\int_{\mathbb{R}} \hat{f}(s\vec{\theta})|s|e^{i(\vec{x}\cdot s\vec{\theta})}dsd\theta \\ &= \int_{S^1}\int_{\mathbb{R}} F_1\left(P_f(\vec{\theta},\vec{\theta}\cdot\vec{x})\right)(s)|s|e^{i(\vec{x}\cdot s\vec{\theta})}d\theta \quad (4) \end{aligned}$$

From (4) we see that the original function may be reconstructed from Fourier transforms of the projections, or line integrals, using a two-dimensional inverse Fourier transform. One problem with this inversion is that the Fourier transform is infamously non-local. Moreover, in this setting the Fourier transform would be sampled on a radial grid and interpolation would be necessary in order to use a conventional Fourier transform based on a rectangular grid.

Filtered backprojection

A commonly used reconstruction method is called filtered backprojection [15]. To obtain the filtered backprojection algorithm we assume that the function f is essentially band-limited on a circle of radius r, and insert a characteristic function, or similar window, into (4) to obtain

$$\begin{aligned} &f_w(\vec{x}) \\ &\quad = \int_{S^1}\int_{\mathbb{R}} F_1\left(P_f(\vec{\theta},\vec{\theta}\cdot\vec{x})\right)(w(s)|s|)e^{i(\vec{x}\cdot\vec{\theta})s}dsd\theta \\ &\quad = \int_{S^1} P_f(\vec{\theta},\vec{\theta}\cdot\vec{x}) * F_1^{-1}(w(s)|s|)d\theta. \quad (5) \end{aligned}$$

The window must be chosen to meet two criteria. First the window should be chosen to agree with the essential band-limit of f, so that f_w will be a good approximation to f. Second, the window also represents a mollification of the inversion of the Radon transform, which is an unbounded operator without the addition of the window [6].

A typical window $w(s)$ might be $w(s) = \mathcal{X}_{[-r,r]}(s)$, where r is chosen to agree with the essential bandlimit of f.

Non-locality

The problem with the reconstruction formula (5) is that the inverse Fourier transform of $|s|w(s)$ will not be locally supported. This stems from the fact that $|s|$ is not differentiable at the origin, and the strict cutoff causes discontinuities at $\pm r$. The non-locality of this inverse Fourier transform implies that local calculation of the convolution in (5) will require global values of the Radon Transform. The discontinuities at the endpoints can be eliminated with with the choice of a suitable window, but the non-differentiability at the origin can not be significantly altered without harming the structure of the image.

We refer to the convolution, or multiplication in the Fourier domain, as a filter $\mathcal{F}$, and we will write it as a two step process. The first is multiplication by (is) in the frequency domain (or differentiation in the time domain). The second is the application of the Hilbert transform $Hf : \hat{f}(s) \rightarrow \frac{1}{i}\text{sign}(s)\hat{f}(s)$. Thus the filtering process is given by

$$\mathcal{F} : \hat{f} \rightarrow \frac{1}{i}\text{sign}(s)(is)\hat{f}(s) = |s|\hat{f}(s).$$

Differentiation is local. The Hilbert transform is not, since it imposes a discontinuity upon the Fourier transform of any function whose average value is not zero, and discontinuities on the higher derivatives which are not zero at the origin.

The imposition of these discontinuities at the origin in the frequency domain will therefore spread the support of functions in the time domain. For this

reason, a local basis will not remain local after filtering. This spreading of local elements is illustrated by a Gaussian in Figure 5.

However, the imposition of discontinuities at the origin in the frequency domain, and the subsequent spreading of the support of the function, will not occur if the function's Fourier transform is zero at the origin, and the function has arbitrarily many zero moments. Further, one can construct functions with only a few zero moments whose support is essentially unchanged by the Hilbert transform. This property is fundamental to the use of our algorithm, as will be described below.

V. WAVELETS

For the reasons outlined above, we would like a basis of functions which are essentially compactly supported, and which possess several zero moments. This second condition will ensure that the basis functions remain essentially compactly supported after the filtering process, and so will allow a reconstruction from localized data. Wavelets are generally constructed with as many zero moments as possible, given other constraints such as locality and smoothness [2]. Thus, the local properties of the high resolution components of a wavelet transform will remain local after the filtering in the reconstruction of the image from the Radon transform.

Wavelet Filtering

We will expand each of the projections in a wavelet transform in order to essentially localize the Hilbert transform, and therefore the filtered backprojection algorithm. If we denote the convolution filter by $\mathcal{F}$ and expand $P_f(\vec{\theta})$ via a wavelet transform, the filtered backprojection routine becomes

$$\begin{aligned} f(\vec{x}) &= \int_{S^1} P_f(\vec{\theta}, \vec{\theta}\cdot\vec{x}) * F_1^{-1}(|s|\chi_{[-r,r]}(s))d\theta \\ &= \int_{S^1} \mathcal{F}\left(P_f(\vec{\theta}, \vec{\theta}\cdot\vec{x})\right) d\theta \\ &= \int_{S^1} \mathcal{F}\left(\sum_{m,n} c_{m,n}(\vec{\theta})\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) d\theta \\ &= \int_{S^1} \sum_{m,n} c_{m,n}(\vec{\theta})\mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) d\theta, \quad (6) \end{aligned}$$

where equality holds in the last line because of the linearity of the filter $\mathcal{F}$.

Localization

There are several advantages to this representation. The first is that as long as ψ has several zero moments, the locality properties of the wavelets will be preserved by the filter $\mathcal{F}$. We can calculate the coefficients $c_{m,n}(\vec{\theta})$ using local information, and since the filtered wavelets have essentially the same support, we need only local information to perform the inversion.

Though certainly encouraging, these statements are only partially true. The derivative of the wavelet will certainly be compactly supported, but the Hilbert transform of this cannot be, although it should decay rapidly. Therefore the filtered wavelets will only have essential compact support. In practice, however, the difference is negligible, as will be shown in section VII.

More importantly, the lower resolution wavelets will have increasingly large support, and therefore so will their filtered versions. We will need non-local data to calculate these coefficients. We would like to use as little non-local data as possible.

Wavelet Sampling

We need to avoid computing the coarse wavelet coefficients for each projection, since the coarse coefficients require the gathering of additional non-local information. We can do this in the following way. We will show that the coarse wavelet coefficients, when viewed as a functions of the angle θ, are essentially bandlimited with the bandlimit dependent upon the essential bandlimit of the coarse wavelet itself. Since the wavelet coefficients are essentially bandlimited, we will be able to compute them from the non-local data for only a few angles and interpolate the values of these coefficients at the other necessary angles.

To see that the above statements are true, we examine the wavelet coefficients and use (1) to expand the projections:

$$\begin{aligned} c_{m,n}(\vec{\theta}) &= \langle \psi_{m,n}(t), P_f(\vec{\theta}, t)\rangle \\ &= \langle \psi_{m,n}(t), (1-t^2)^{-1/2}\sum_{l=0}^{\infty} T_l(t)h_l(\vec{\theta})\rangle. \end{aligned}$$

Applying the Fourier isometry to the above yields

$$\begin{aligned} c_{m,n}(\vec{\theta}) &= \langle \hat{\psi}_{m,n}(s), (\pi/2)^{1/2}\sum_{l=0}^{\infty} i^l J_l(s)h_l(\vec{\theta})\rangle \\ &= (\pi/2)^{1/2}\sum_{l=0}^{\infty}\langle \hat{\psi}_{m,n}(s), i^l J_l(s)\rangle h_l(\vec{\theta}) \quad (7) \end{aligned}$$

where $J_l(s)$ is a Bessel Function of the first kind.

Formula (7) is the key to wavelet sampling of the Radon transform. Notice that as $l \to \infty$ the Bessel functions $J_l(s)$ approach zero on any finite interval about the origin. Coarse wavelets, on the other hand, have Fourier transforms which are essentially supported in a localized region about the origin. Therefore the expansion (7) will have only finitely many

significant coefficients. Thus only low-order exponentials will contribute to the expansion of a coarse wavelet coefficient. From this discussion we see that the representation in Theorem 1 allows us to relate radial band-limiting of $P_f(\theta,t)$ with respect to t to angular band-limiting of the respective coefficients. The following theorem quantifies these statements.

Theorem 2 *Let $P_f(\theta,t) \in L^2\left[S^1 \times [-1,1]\right]$. Then the wavelet coefficients $c_{m,n}(\theta) \equiv \langle \psi_{m,n}, P_f(\theta,t)\rangle$ are essentially band-limited. More precisely we can write $c_{m,n}(\theta) = \sum_n a_n e^{-in\theta}$ where the coefficients a_n satisfy the inequality*

$$|a_n| \leq (\pi/2)^{1/2}\|P_f(\theta,t)\|_{2,w} \times \left(\|\psi_{m,n}\widehat{\chi}_{\mathbb{T}_{n'}}\|_{2,w} + c_1(n')e^{(-\frac{2}{3}n)c(n')^{\frac{3}{2}}}) \right) \quad (8)$$

where $\mathbb{T}_n \equiv \mathbb{R} \setminus [-n+\delta, n-\delta]$ represents the support of the tail of the wavelet, and $c_3(\delta)$ and $c_4(\delta)$ are positive constants. Furthermore, if we assume that $\mathrm{supp}\,[P_f(\theta,t)] \subseteq S^1 \times [-1+\epsilon, 1-\epsilon]$, *then we have the following bounds on the Fourier coefficients of the wavelet coefficients $c_{m,n}(\theta)$:*

$$|a_n| \leq M(\epsilon)\|P_f(\theta,t)\|_2 \times \left(\|\hat{\psi}_{m,n}\chi_{\mathbb{T}_{n'}}\|_2 + c_1(n')e^{(-\frac{2}{3}n)c(n')^{\frac{3}{2}}} \right) \quad (9)$$

where $M(\epsilon) = (\pi/2)^{1/2}(1-(1-\epsilon)^2)^{-1/2} = (\pi/2)^{1/2}(2\epsilon+\epsilon^2)^{-1/2}$.

Thus, if the essential support of $\hat{\psi}_{m,n}$ is outside the region $[-n+\delta, n-\delta] \sim [-n, n]$, then the Fourier coefficients for $c_{m,n}(\theta)$ will decay exponentially after the first n terms. Therefore the Fourier transform of $c_{m,n}(\theta)$ will be essentially band-limited with band-limit n. Furthermore the dilation of the wavelet by a factor of two will double or halve the band-limit of the corresponding wavelet coefficient. This discussion can be formalized in the following sampling corollary.

Corollary 1 (Wavelet sampling)
Suppose that the optimal Nyquist sampling rate for a wavelet coefficient $c_{m,n}(\theta)$ is δ. Then the optimal Nyquist sampling rate for the coefficient $c_{m+1,n}(\theta)$ is 2δ and the corresponding rate for $c_{m-1,n}(\theta)$ is $\frac{1}{2}\delta$.

The proofs of the above theorems can be found in the Appendix.

VI. THE ALGORITHM

The simple algorithm

The band-limits on the coarse wavelet coefficients allow us to interpolate all their values from samples taken at a reduced number of angles $\vec{\theta}$. Only for these angles must we expose outside the local area. For the remaining angles, the localized support of the filtered high-resolution wavelets allows us to expose only the local region of interest (plus a small extra margin). By interpolating the coarse coefficients, we can fill in the data missing from these "under-exposed" angles. Of course, we won't reconstruct the whole image exactly, but the reconstruction of the local region will be nearly exact.

In practice one assumes that the projections are also band-limited in the s variable, and samples them at the Nyquist rate. This imposes a "natural" highest resolution on the wavelet transform. It is at the next lower resolution that the first possibility of undersampling arises. Figure 6 shows a plot of the inner products $|\langle\hat{\psi}_{m,n}(s), J_l(s)\rangle|$ for this second-highest-resolution wavelet, using the Daubechies D_6 wavelet. As one might expect, the effective cutoff is around half the Nyquist frequency. This behavior continues in the lower scales, each having an effective cutoff half that of the scale before it.

The simplest version of our algorithm, appropriate for a 128-by-128 pixel image, is as follows. Take full-exposure projections at every other angle, just as in current tomography systems. For the remaining angles, however, expose only the region of interest, plus a margin on either side corresponding to the support of the analysing wavelet. Then, for each angle, or "line", in the projection data, perform a wavelet transform (the standard pyramid algorithm will do). From the low-order coefficients on the full-exposure lines, interpolate the low-order coefficients for the localized lines, retaining the highest-order coefficients which were calculated locally.

Next, invert the approximate wavelet transforms so constructed, line-by-line, and use a standard backprojection algorithm to produce the final image. This image is nearly identical in the region of interest to the image produced by filtered backprojection. The differences between this image in the region of interest and that produced by filtered backprojection using full data will come from two sources. First, the filtered high-resolution wavelet is not compactly supported, it merely has rapid decay. However, for many wavelets, the difference is negligible. Second, the low-order wavelets may not be truly band-limited, resulting in slight aliasing in the interpolation. Again, in practice, the difference is small for smooth wavelets.

The full algorithm

The full algorithm is the natural extension of the simple algorithm, and obtains greater savings by taking full advantage of locality. Although the highest-order coefficient is the most localized, it is not the only level of coefficient on which we can save. Coefficients at several more scales can be computed for the region

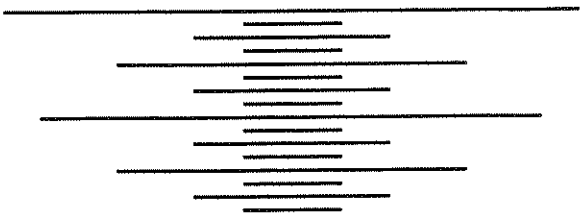


Figure 1: Schematic of exposure pattern

of interest from less-than-complete projection data. Figure 1 is a schematic view of the exposure pattern needed to image a small region in the center of a larger image. Each line represents a different angle of projection, with the length and position of the line showing where projections are required for that angle. Figure 3 shows the Radon transform of the Shepp-Logan phantom, and the data gathered in this sampling scheme.

Note that as the wavelet gets wider at lower scales, we need to widen the exposure to reconstruct the region of interest, but we need data from fewer angles to get coefficients for all angles, thanks to the band-limit.

The number of scales at which less-than-complete exposures can be used depends not only on the size of the region of interest, but also on the resolution of the full image. A 512-by-512 pixel image provides opportunity for a greater percentage decrease in radiation than a 128-by-128 pixel image, since more of the high-order wavelets are "local", relative to the size of the image, and so fewer full-exposure angles are needed. While the savings for the simple version of the algorithm is clearly bounded by 50%, savings on higher resolution images are much less limited. For example, imaging a 32-by-32 pixel region in a 256-by-256 image, using the D_6 wavelet, would result in a reduction in radiation exposure of 58%. At twice the resolution, imaging a 64-by-64 pixel region in a 512-by-512 image, the savings increases to 69%.

Note that these savings are in addition to the savings that can be realized using the well-known technique of interlaced sampling [21, 20, 19], which is discussed in section VII. When interlaced sampling is used, the savings for these two examples increase to 79% and 84%, respectively.

A table outlining the savings at a variety of scales and with a variety of sizes of regions of interest is shown in Figure 4. Note that as the size of the array increases the amount of data used by our algorithm approaches the amount of "local data". For instance, a reconstruction of an inner 1/8 radius region of interest in a 512x512 array uses 20% of the data, 12.5% of which passes through the region of interest. If the array size is increased to 2048 the percentage of data needed is only 17%.

Off center regions

Although the algorithm as presented assumes that the region of interest is centered in the image, off-center regions can also be imaged. The collection of data remains the same (although, of course, the data is collected around the off-center region of interest), but a simple transformation on the projection data is required to center the region before interpolation is performed. This transformation may require one to pad the data with zero values, and possibly double the size of the array. This results in a doubling of the computation time for the wavelet transforms, but does not significantly affect the accuracy of the reconstruction and does not increase the amount of x-ray exposure required.

VII. ERROR ANALYSIS

The performance of the filtered backprojection algorithm has been studied extensively [6, 16, 17, 18]. We will show that the image produced by our algorithm is only additively different from that produced by the filtered backprojection of the complete sinogram, and that in the region of interest, this additive difference can be chosen to be lower than the resolution of the imaging machinery.

We will denote the region of interest by $\mathcal{RI}$, and the wavelet coefficients which have support in the projections through the local region by $\mathcal{L}$. The wavelet coefficients which do not have support in the local region will then be denoted by $\mathcal{NL}$. Thus, our algorithm consists of identifying the coefficients $\mathcal{L}$ which are of interest, and determining them with a minimal amount of data. The coefficients in $\mathcal{NL}$ are not used, and are never computed.

Truncation Error

We begin with the filtered backprojection formula (6),

$$f(\vec{x}) = \int_{S^1} \sum_{n,m} c_{n,m}(\vec{\theta}) \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) d\theta. \qquad (10)$$

This can be separated into two parts, corresponding to the portion which we will compute, and the portion which we will ignore. This decomposition is then given by

$$\begin{aligned} f(\vec{x}) &= \int_{S^1} \sum_{\mathcal{L}} c_{n,m}(\vec{\theta}) \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) d\theta \\ &+ \int_{S^1} \sum_{\mathcal{NL}} c_{n,m}(\vec{\theta}) \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) d\theta \quad (11) \end{aligned}$$

The result obtained by our algorithm is therefore

$$f_a(\vec{x}) = \int_{S^1} \sum_{\mathcal{L}} c_{n,m}(\vec{\theta}) \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) d\theta \quad (12)$$

and the difference between our algorithm and the result given by filtered backprojection of the complete sinogram is given by

$$Err = \left| \int_{S^1} \sum_{\mathcal{NL}} c_{n,m}(\vec{\theta}) \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) d\theta \right|. \quad (13)$$

We will show that this term can be made smaller than the resolution of standard imaging software. Thus the error caused by this term will not be detectible when viewing the images.

To obtain bounds on (13), we apply the Cauchy Schartz Lemma in the following way,

$$\begin{aligned} Err &= \left| \int_{S^1} \sum_{\mathcal{NL}} c_{n,m}(\vec{\theta}) \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) d\theta \right| \\ &\leq \int_{S^1} \sum_{\mathcal{NL}} \left| c_{n,m}(\vec{\theta}) \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) \right| d\theta \\ &\leq \int_{S^1} \left(\sum_{\mathcal{NL}} |c_{n,m}(\vec{\theta})|^2 \right)^{1/2} \times \\ &\quad \left(\sum_{\mathcal{NL}} \left| \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) \right|^2 \right)^{1/2} d\theta. \quad (14) \end{aligned}$$

Now since the wavelets are an orthonormal basis, and the $c_{n,m}$ are merely the coefficients of an expansion in that basis, the first factor is less than $\|P_f(\theta)\|_2$ for each fixed θ. We can then bound this factor by the maximum value of f, since the projections are merely integrals of f in the unit circle. We take the maximum of the second term and we get that

$$\begin{aligned} Err &\leq \int_0^{\pi} \|P_f(\theta)\|_2 \left(\sum_{\mathcal{NL}} \left| \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) \right|^2 \right)^{1/2} d\theta \\ &\leq \pi \max_{\vec{x}} |f(\vec{x})| \times \\ &\max_{\theta} \max_{x\in\mathcal{RI}} \left(\sum_{\mathcal{NL}} \left| \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) \right|^2 \right)^{1/2} \quad (15) \end{aligned}$$

Thus the error will be negligible if

$$\max_{\theta} \max_{x\in\mathcal{RI}} \left(\sum_{\mathcal{NL}} \left| \mathcal{F}\left(\psi_{n,m}(\vec{\theta}\cdot\vec{x})\right) \right|^2 \right)^{1/2} \quad (16)$$

is sufficiently small.

Let us define $S = \{\vec{x}\cdot\vec{\theta} : x \in \mathcal{RI}, \vec{\theta} \in S^1\}$. This is the portion of the sinogram which intersects the region of interest. If we center our region of interest in the image space, then S will be a strip down the center of the Radon transform which is invariant with respect to θ. This simplifies the error term (16) to the form

$$\max_{t\in S} \left(\sum_{\mathcal{NL}} |\mathcal{F}\left(\psi_{n,m}(t)\right)|^2 \right)^{1/2}. \quad (17)$$

This term measures the amount that the non-local wavelets will "bleed" into the region of interest after they are filtered. There are two ways to reduce this term. The first is to insist on more zero moments in the wavelets, resulting in less spreading of the non-local wavelets. The second would be to keep more of the non-local coefficients. Numerical tests show that the use of wavelets with several zero moments is sufficient to make this term negligible.

Wavelet	Savings	Error	Savings 2
D2	72	.0001	86
D4	65	.00003	82
D6	58	.00001	79

Figure 2: A comparison of the error bounds and savings for the Daubechies wavelets D2, D4, and D6. The second savings number is the number for the algorithm coupled with interlaced sampling.

In Figure 8 we consider the error terms for the Daubechies D2, D4, and D6 wavelets. These errors were lower than we had originally hoped, and are certainly beyond the resolution of standard imaging software (i.e. 1/256). Even if one has two byte capability, or $(1/256^2)$ resolution, the D6 error term is still below this. As expected the error term is reduced for the wavelets with higher moments, but was surprisingly low for even the D2 wavelet. In the same figure we consider the corresponding savings of the algorithm for a 256x256 pixel image. The increased error reductions are paid for by an increase in the amount of data needed, due to the increased support of the wavelets.

The above analysis might be slightly misleading in that one typically zooms in on the most significant variances in an image, rather than displaying the complete range of the image unscaled. Thus the additional orders of magnitude achieved by the D4 or D6 wavelets might be necessary even for a one byte imaging system.

Interpolation Error

The above results show that the D2 wavelet produces an error bound of only .0001, which is below the resolution of nearly any imaging software, not to mention the human eye. Since this wavelet produces the greatest savings, it would seem that this would be the optimal choice. However, many of the coefficients in

the set $\mathcal{L}$ will have to be approximated by our algorithm via interpolation. Thus, there will be an additional error in the region of interest from the first term in (11) which depends on the degree to which the wavelet used is essentially band-limited. Aliasing will occur, which is significant with the D2 wavelet, but the effects of the aliasing are greatly reduced with the D4 and D6 wavelets for a number of reasons.

The first is that the filter in filtered backprojection must closely resemble the ramp $|s|$. Thus low frequencies will be diminished and high frequencies will be emphasized. Since we will be sampling only the coarse resolution wavelet coefficients, which correspond to low frequencies, the aliasing will be deemphasized by the filtering process.

Although this helps to reduce the interpolation errors, it will not reduce them sufficiently if the wavelet in question is not sufficiently localized in frequency. This can be seen in Figure 9, where the D6 wavelet generates an excellent image, with the reconstruction errors being negligible to the resolution of the imaging device, but the reconstruction using the D2 wavelet shows considerable artifacts. Given the extremely low bound on the truncation error for the D2 wavelet, this must be attributed to the aliasing due to the fact that the D2 wavelet is not sufficiently band-limited.

To illustrate this, the absolute value of the Fourier transforms of the D2, D4, and D6 wavelets are shown in Figure 17. The D2 wavelet is not very localized in frequency, and therefore it is not surprising that the aliasing error is very noticeable. The D4 and D6 wavelets, however, are much more localized in frequency, which explains their superior performance in our algorithm.

We would like to be certain that the errors induced while using the D2 wavelet with our algorithm are due to aliasing. To check this, in Figure 10 we used the local coefficients of the D2 wavelet, but have calculated them exactly, from non-local data if necessary. Thus there is no interpolation error, and the approximation is actually that depicted in (12). The reconstruction is essentially exact, illustrating that the errors are due to the aliasing, and our truncation error bounds are correct.

Tradeoffs

As outlined above, the tradeoff in the algorithm is that better truncation error bounds, as well as lower aliasing errors are achieved when wavelets with larger support are used. Using wavelets with larger support requires the gathering of slightly more data, which is the tradeoff between locality and accuracy. Our numerical experiments to date have shown that in general, the D6 wavelet minimizes the aliasing errors to where they are completely acceptable, and the truncation errors for the D6 wavelets are far below the tolerance thresholds of our imaging system. The D4 wavelet also has a negligible truncation threshold, and the aliasing errors for the D4 have not been significant, but have been noticeable. The D2 wavelet, on the other hand, produces unacceptable aliasing errors, and therefore is unacceptable.

On the basis of this numerical work, it seems that one could use either the D6 or D4 wavelets, with the D4 wavelet using less data, and the D6 wavelet providing more accuracy.

As opposed to the simple analysis for the truncation error, a precise analysis of the interpolation error is beyond the bounds of this paper. A simple analysis could be done rather easily, but it is the belief of the authors that this might be misleading. The reason is that the filtered backprojection algorithm will effectively integrate through the aliasing, and as a result, may very well minimize some of the effects of the aliasing as in [18]. Emperical evidence from the Daubechies wavelets suggests that this error will be minimal in practice when a smooth enough wavelet is employed.

VIII. WAVELETS VS. LAMBDA TOMOGRAPHY

Lambda-tomography refers to an algorithm introduced in [7, 8] to recover local features of an object, from local values of the Radon transform. The algorithm of [7, 8] does not recover the original density function f, but rather the function $\Lambda f - \mu\Lambda^{-1} f$, where

$$\widehat{\Lambda f}(\xi) \equiv |\xi| \hat{f}(\xi).$$

The beauty of the Λ transform is that it preserves smooth regions of an image, and emphasizes edges or discontinuities. Moreover, the Λ transform changes the filter in filtered backprojection from the inverse Fourier transform of $|s|w(s)$ to the inverse Fourier transform of $s^2 w(s)$, where $w(s)$ is a window chosen to agree with the essential band-limit of the object. This is significant because the inverse Fourier transform of $s^2 w(s)$ can be made to have a very small essential support, since $s^2 w(s)$ can be made infinitely differentiable. As has been noted above, the inverse transform of $|s|w(s)$ will not be essentially compactly supported because of the discontinuity at the origin. Thus the Lambda transform allows local convolutions, and therefore local application of filtered backprojection.

The drawback of the Λ transform is that it produces a cupping of regions which are essentially constant. To counteract this, it was shown in [7] that $\Lambda^{-1} f$ contains a countercup, which will essentially neutralize the cupping in Λf if an appropriate constant μ is chosen. Thus, one reconstructs $\Lambda f + \mu\Lambda^{-1} f$. Since $\Lambda^{-1} f$ is simply backprojection without filtering, it follows that this algorithm can be

employed from localized values of the Radon transform.

Although $\Lambda f + \mu\Lambda^{-1} f$ is not an approximation of f, it was shown in [7] that $\Lambda f + \mu\Lambda^{-1} f$ is very similar to f in many ways. Given this, one would like to find a way to reconstruct a function g, in a local manner, so that $\Lambda f + g$ would be an approximation to f.

Unfortunately, this is not possible, since this would imply that the interior Radon transform is invertible. Our algorithm, however, uses a small amount of exterior data in order to stabilize the inversion. The natural question then becomes, can a small amount of exterior data be used to recover $f - \Lambda f$?

To examine this question we note that the convolution kernel for a filtered backprojection reconstruction of $f - \Lambda f$ would be given by the inverse transform of $(|s| - s^2)w(s)$. This kernel is just as non-local as the original convolution kernel, since it is the difference of the original non-local kernel, and the local kernel $s^2 w(s)$. Thus one can certainly not use local samples to recover this information.

Since our algorithm uses non-local information, sampled on a sparse grid, the question is whether or not sparse non-local information can be used to recover $f - \Lambda f$. The answer to this question is no, for the following reasons.

Sampling of $f - \Lambda f$

In order utilize sparse sampling to recover a function one must know the essential band-limit of the function. The necessary tool to understanding sampling of the Radon transform in two dimensions is to understand the essential support of its two dimensional Fourier transform. The range characterization (1) of the Radon transform shows us that the Fourier transform of the Radon transform is supported essentially on a bowtie in the transform domain [6, p.74].

Now suppose that $P_{\Lambda f}(\theta, r)$ is the Radon transform of Λf. Then the projection slice theorem states that the Fourier transform of this with respect to r will be $|s|\hat{f}(s\vec{\theta})$. Thus the Fourier transform of $P_{f-\Lambda f}(\theta, r)$ with respect to r will be given by $(1 - |s|)\hat{f}(s\vec{\theta})$. The Fourier transform of this with respect to θ, which is necessarily the two dimensional Fourier transform of $P_{f-\Lambda f}(\theta, r)$, is $(1 - |s|)$ times the Fourier transform of $f(s\vec{\theta})$. The Fourier transform of $f(s\vec{\theta})$ with respect to θ is the two dimensional Fourier transform of $P_f(\theta, r)$. Thus the essential support of the Fourier transforms of $P_f(\theta, r)$ and $P_{f-\Lambda f}(\theta, r)$ are identical except at $|s| = 1$. Therefore one cannot sample the Radon transform of $f - \Lambda f$ more sparsely than the Radon transform of f.

Necessary conditions for essential localization

The essential difference between our method and Lambda tomography is the following. Lambda-tomography is dependent upon the multiplication by $|s|$ to change the filter from $|s|w(s)$ to $s^2 w(s)$. If one were to obtain an approximation to f from Lambda-tomography, one would need to recover $(1-|s|)\hat{f}(s\vec{\theta})$. Since the essential support of the Fourier transform of $(1-|s|)\hat{f}(s\vec{\theta})$ with respect to θ is the same as that of $\hat{f}(s\vec{\theta})$ the sampling conditions cannot be loosened in order to recover the "coarse" features of $(1-|s|)\hat{f}(s\vec{\theta})$.

Local calculation of high frequency data

The algorithm we have proposed, however, uses wavelets to recover the local high frequency components on the region of interest. As a result, one needs only to recover the low-pass version of f, $\hat{f}_{low} = w(s)\hat{f}(s\vec{\theta})$. Since the essential support of $w(s)$ is concentrated around the origin, the Fourier transform of the Radon transform of f_{low}, which is the Fourier transform of $\hat{f}_{low} = w(s)\hat{f}(s\vec{\theta})$ with respect to θ, will be $w(s)$ times the two-dimensional Fourier transform of $P_f(\theta, r)$. Since $w(s)$ has essential support concentrated about the origin, and since the Fourier transform of $P_f(\theta, r)$ is a bowtie centered about the origin, the Fourier transform of the Radon transform of f_{low} can be made to have support contained in a small rectangle about the origin in the frequency domain.

This is illustrated in Figure 18, where the essential support of the Radon transform of a function which is supported inside the unit circle is shown in (a), and the essential support of the Radon transform of a function, times a window $w(s)$ in frequency is shown in (b). If the essential support of $w(s)$ is between the dashed lines, then the essential support of the Radon transform the windowed function will be contained within the small square in the center.

Sparse sampling of low frequency data

This allows us to subsample the non-local information in a very sparse manner, and therefore recover an approximation to the actual density function f, on a local region, with less information than is needed to recover f on the whole region. Thus the necessary condition for an algorithm which obtains high frequency information locally, and then uses sparse samples to obtain the low frequency information is the following. If one recovers f_{high} locally, the remaining remaining low passed version of f, $f_{low} = f - f_{high}$, must have a Radon transform whose Fourier transform has much smaller essential support than the Fourier transform of the Radon transform of f. In practice this is also accomplished by ignoring the high frequency data away from the region of interest, namely $f_{low} = f_{loc} - f_{high}$, where f_{high} is the high passed version of the image on the region of interest, and f_{loc} is the low passed image plus high passed image on the region of interest.

IX. NUMERICAL RESULTS

We have tested the full algorithm on several images and it performs as predicted. Figure 11 shows an example of the algorithm using the D6 wavelet. We reconstructed the Shepp-Logan phantom from 256 projections, with 256 samples per projection, at a 256x256 pixel resolution. We used both full data and 42% of the data, while localizing on the inner 1/8 of the image. The relative difference between the images within the region of interest is less than 1/256 and therefore the reconstructions are pixel for pixel identical within the region of interest.

In Figure 12 we reconstructed using the full algorithm with the D4 wavelet. The local reconstruction required 35% of the data. Because the D4 wavelet is not as localized in frequency there is more error due to aliasing, but once again the images are almost indistinguishable in the region of interest.

In Figure 13 we reconstructed an off-centered region from local data. This required that the projection data be "shifted," so that the region of interest is centered in the data. As a result, the size of the projection array had to be doubled. This does not alter the amount of data that needs to be collected, however; it only requires zero padding. There was a more pronounced aliasing in the off-centered region, which is not completely unexpected. It is primarily due to the discrete nature of the data. We are forced to "shift" the data by full pixels, or to attempt interpolation to move by partial pixels. Thus, the result of the shifting operation is not exactly a true sinogram, and the results of Theorem 2 do not apply completely. The difference, however, is quite minimal, and can be mollified by interpolating at fewer scales of wavelet coefficients. The results of localizing on the first two scales is shown in (b) and (c) shows the result of localizing the first three scales. Note that these few scales account for most of the total savings in the full algorithm, and so the loss in savings is not significant.

Interlaced Wavelet Sampling

Interlaced sampling of the Radon transform for computerized tomography was first suggested by Cormack in [19], and has been thoroughly analyzed in [21, 6, 18, 20]. It was proven in [21, 20] that one can effectively take half the number of samples, at each fixed angle, and recover an image with the same resolution as the full sampling scheme. This is accomplished via a standard theorem from multidimensional sampling and wise use of the range theorem (1).

The reconstruction, which is described in great detail in [18, 20], involves collecting every other sample of $P_f(\theta_j, t_i)$, but interlacing the gathered samples. Thus at each angle θ_j, one collects $P_f(\theta_j, t_i)$ whenever $i \equiv j \bmod 2$. At each angle, the projection $P_f(\theta, t)$ will be undersampled by a factor of two. One ignores this fact, and does a sinc interpolation accross t for each angle, with the corresponding aliasing errors. The miracle then occurs when the aliasing errors cancel in the integration of the filtered backprojection [20, 21].

To make this work most efficiently, it was shown in [18] that one should use a window other than the standard sinc or square window, to do the interpolation. The addition of this window to the process allows the interpolation to be done locally, and therefore it adapts perfectly to our algorithm. The price for this local interpolation is a slight decrease in the resolution of the object.

Examples of using our algorithm with the interlaced sampling scheme are shown in Figure 14. In Figures and 4 we have also included the savings of the interlaced sampling scheme.

A comparison to SVD methods

Our method relies on a small amount of additional data from outside the region of interest to do the reconstruction, as opposed to the singular value decomposition methods for the interior Radon transform [9, 14]. However, the price which is paid for full locality in the singular value decomposition methods is that there is a nullspace of the operator, which cannot be recovered, as well as singular vectors which correspond to singular values which are too small to recover.

It has been noted in a number of places [9, 6, 14], that the elements of the nullspace do not vary much well within the region of interest. This can be proven in some form, but we feel that there are serious questions to be answered about how much variance there is, and how far within the region of interest one must go before the variance is small.

For comparison, we assume that we are doing reconstructions with 256 samples per projection, and that we are interested in the interior 32 pixels of the image. We construct an element of the nullspace, as in [6], by defining its projections, at each angle, to be zero in the interior 32 pixels, and 1 for 16 pixels on either side of this region. It is easily checked that this is an element of the nullspace, and that it is consistent with the Helgason-Ludwig conditions [6]. The projection of this circularly symmetric nullspace element is shown on the right of Figure 15. A reconstruction of the nullspace element is made using our method is shown in the same figure, along with cross sections of the nullspace element and our reconstruction of this function. Note that our method, as predicted, reconstructs the nullspace element exactly on the region of interest.

The striking thing about this nullspace element is that although it is smooth inside of the region, it is varies quite a bit, and therefore would significantly

change the quality of any reconstruction which ignored it. This does not violate the theorems proven in [9, 6, 14], but rather points to their asymptotic nature, which does not take effect when the region of interest is truly local.

The need for non-local information when using SVD methods

The variance of the nullspace element is greatest at the edges of the region of interest, so one approach would be to use SVD methods, but on a region larger than the region of interest. This would reduce the amount of variance within the true region of interest. Thus one must gather more than localized data to do a localized reconstruction, which eliminates the supposed advantage of the SVD over our algorithm.

The beauty of our algorithm is that is uses essentially the same amount of data as reconstruction from completely local values, but recovers the image exactly. For instance, at a resolution of 512x512, which is fairly common in medical imaging, we would use 32% of the data to image an local region of radius 1/4 with a D4 wavelet. The local data comprises 25% of this data, so we are only adding 7% of additional data in order to guarantee that the image is identical to that produced by filtered backprojection with full data within the region of interest.

If one were to try to use local information exclusively with the SVD for the interior Radon transform, one would certainly have to increase the size of the region to be imaged since the nullspace elements ring significantly on the border of the region of interest. Thus one will need to gather data beyond the 25% level, and still won't be guaranteed that the values in the final image are at all exact.

Reconstruction methods other than filtered backprojection

Although we used the filtered backprojection algorithm to design our sampling scheme, and the resulting algorithm, our algorithm is not inherently dependent upon the filtered backprojection method.

The filtered backprojection method in its mollified form is merely a reconstruction of a bandlimited version f_B of the object function f [6]. Our algorithm specifies the gathering of data which corresponds to f_B, in the region of interest, and an even more bandlimited version of f, f_{MB}, outside of the region of interest. Thus the algorithm specifies the gathering and processing of data, to produce a data set corresponding to another image. This second image, however, is equal to the original image in the region of interest.

There is no reason why reconstruction methods other than filtered backprojection won't work equally well. Examples would include the circular harmonic Radon transform [23] and the Hankel transform reconstructions [24]. Our method merely specifies projection data corresponding to an image which is equal to the bandlimited version f_B of the object function within the region of interest. Thus any algorithm which inverts the Radon transform in a stable manner should work in conjunction with our method.

Local tomography via the Lambda transform

Another way of doing local tomography has been outlined and studied extensively in [8, 7]. It must be emphasized that the two methods are different, and comparison of our method to their method is not really appropriate. First of all, the two methods are designed to do two different things. Our method uses slightly more than local data to reconstruct an approximation to the actual density function, or the actual image. Lambda localized tomography is primarily intended for edge detection and does not employ much data from beyond the local region of interest. Lambda tomography does not recover an approximation to the original density function, but rather an image which is related to the original density function in several ways, particularly in the locations of edges. As a result, these methods are complementary rather than competing algorithms.

For the sake of completeness we have included two comparisons of Lamda-tomography and our algorithm. The first, in Figure 19, shows a local reconstruction of Λf, and a local reconstruction of the center region using our algorithm. Note that both algorithms reconstruct the image well in the center of the region. Our algorithm is more accurate in the center of the image, which is due to the fact that we use more data.

The Lambda-tomography reconstruction of the Shepp-Logan phantom does exhibit a rather large artifact at the boundary of the phantom. This is due to the fact that Λ is the square root of the Laplacian and therefore will accentuate all discontinuities, especially large ones.

To emphasize the difference between our algorithm and Lambda-tomography we illustrate in Figure 20 that our algorithm does not exhibit the artifact at the boundary of the image, when we localize on that portion of the image. Once again, this is due to the fact that we can sample non-locally to obtain the course structure of the image, and then obtain the fine structure locally.

In this example μ was chosen using a least squares criterion, in order to minimize the variation on the interior region, which was known to be essentially constant. One of the difficulties of using Lambda-tomography is that at present, there is no theoretical basis for determining μ [27].

X. A WAVELET HIGH-PASS FILTER

Local tomography, or Lambda tomography, as developed in [7, 8], is of interest not only because of its locality, but also because of the enhanced edge detection properties which it exhibits. It was proven in [7] that Λf has the same edges as f, and that these edges are enhanced. This is principally due to to the fact that Λ is the square root of the positive Laplacian $-\Delta$. In [13] an algorithm is given for recovering $\Delta^{1/2} f$ using wavelet and Gabor transforms on the Radon transform.

Moreover, it has been shown that edge detection is equivalent to detection of maxima in the wavelet transform [5]. In [10] an algorithm to detect edges within an image directly from the projections $P_f(\vec{\theta}, t)$ was developed. It therefore seems that a variety of enhancement and edge detection algorithms could be developed from the wavelet transform applied to the projection data.

We will present a basic technique which illustrates the potential of the wavelet transform for edge detection. For edge detection we use a simple high-pass wavelet filter, keeping only the finest scale wavelets from the projection data. This is equivalent to using a donut shaped high-pass filter in the Fourier transform, with a cross section of the donut being the Fourier transform of the wavelet. We then use filtered backprojection to reconstruct the image.

Figure 16 shows the results of the application of this high-pass filter to the Shepp-Logan phantom, using the D_6 wavelet. The first image is a high-pass filter reconstruction of the image using full projection data. The second image is a local high-pass filter reconstruction of the center portion of the image, using only those projections which pass through a small neighborhood of the center. These are then compared against the corresponding images using Lambda-tomography.

The important thing to note concerning the local application of this high-pass filter is that it is truly local. One needs only the high-resolution wavelet coefficients, which can be calculated from values in a neighborhood of a local region. There is no need for the subsampling and interpolation of low-order wavelet information which is necessary for reconstruction of f.

XI. WHY WAVELETS?

We have alluded throughout to the fact that the properties which allow a basis to essentially localize the Radon transform are not only conducive to the use of wavelets, they actually specify that a wavelet type basis be used. At this point it is beneficial to outline those properties which are necessary for our algorithm to be successful. There are three properties that a basis must possess in order to be suitable.

1. **Essential compact support in time:** This is necessary so that the coefficients of the expansions of the projections $P_f(\theta, t)$ can be computed locally.
2. **Zero moments:** The basis must have zero moments so that the Hilbert transform does not cause low-order discontinuities in its Fourier transform, and correspondingly cause the basis elements to become non-local after the filtering in the filtered backprojection routine.
3. **Essential compact support in frequency:** The basis must be localized in frequency, so that the first bound in Theorem 2

$$\|\hat{\psi}_{m,n} \chi_{\mathbb{R}\setminus[-n+\delta, n-\delta]}\|_2$$

will vanish appropriately, allowing subsampling of the non-local, low order wavelet coefficients. The Theorem could be rewritten for any basis which is localized in frequency, but rests heavily on the localization.

Note that smoothness in time is not an explicit criterion, but is implicitly required by the third property.

By definition, any basis which satisfies these three criterion is a time-frequency localization basis. These properties are not shared by polynomials, spherical harmonics, or many of the other traditional bases which have been used in the tomography literature.

There has been a recent paper, [22], which proposed a set of rotationally symmetric basis functions which are essentially time and band-limited. They do not have zero moments, however, and therefore they will not remain local under the influence of the Radon transform.

It should be noted that there is no particular reason why the local cosine bases of Coifman and Meyer [26], or wavelet packets could not be used.

We will now examine four standard wavelets, those of Meyer, Chui-Wang, Lemarié-Battle, and Daubechies in terms of these three criterion. We will begin with the Meyer wavelets.

The Meyer Wavelets

The Meyer wavelet [2, p.117], has its Fourier transform defined as

$$\hat{\psi}(\zeta) = \begin{cases} (2\pi)^{-1/2} e^{i\zeta/2} \sin[\frac{\pi}{2}\nu(\frac{3}{2\pi}|\zeta| - 1)], & \frac{2\pi}{3} \le |\zeta| \le \frac{4\pi}{3}, \\ (2\pi)^{-1/2} e^{i\zeta/2} \cos[\frac{\pi}{2}\nu(\frac{3}{2\pi}|\zeta| - 1)], & \frac{4\pi}{3} \le |\zeta| \le \frac{8\pi}{3}, \\ 0 & \text{otherwise} \end{cases}$$

where ν is a C^k or C^∞ function satisfying

$$\nu(x) = \begin{cases} 0 & \text{if} x \le 0, \\ 1 & \text{if} x \ge 1, \end{cases} \tag{18}$$

with the additional property

$$\nu(x) + \nu(1-x) = 1.$$

The key consideration when constructing ν is that $\hat{\psi}$ will inherit the smoothness properties of ν. We now analyze the optimality of the Meyer wavelet in terms of the three criterion developed above.

1. **Essential compact support in time:** As is pointed out in [2, p.119] even if $\nu \in C^\infty$ which would imply that for any N there exists a C_N such that

$$|\psi(x)| \leq C_N(1+|x|^2)^{-N},$$

the numerical decay of ψ may not be as fast as desired. In other words we don't really care about the decay at infinity, since we will not go out that far. What we really want is rapid decay in the near vicinity of the center of the wavelet. Although we do have the bounds listed above, the constants C_N could tend to ∞ as $N \to \infty$. Thus if we want $\inf\{n : |\psi(x)| \leq \epsilon \text{ for all } |x| \leq n\}$ to be minimized, the Meyer wavelet will probably fail. One may, however, be able to construct a ψ which is close to achieving this desired goal.
2. **Zero moments:** The Meyer wavelet is optimal in terms of moments, since its Fourier transfrom is zero in a region about the origin and therefore posesses an infinite number of zero moments. This implies that the image of the wavelet after filtering will have the same theoretical decay properties as the original wavelet. Since we must jointly localize these two functions this may prove to be a very important property. However, one should check to see that this is the numerical truth, after one chooses an appropriate function ν.
3. **Essential compact support in frequency:** The Meyer wavelet is optimal according to this criterion also. The Fourier transform is compactly supported and therefore the first error term in Theorem 2,

$$\|\hat{\psi}_{m,n}\chi_{\mathbb{R}\setminus[-n+\delta,n-\delta]}\|_2,$$

will be zero for n picked such that $\text{supp}[\hat{\psi}] \cap \mathbb{R} \setminus [-n+\delta, n-\delta] = \emptyset$. Thus the Meyer wavelet leads to strictly band-limited wavelet coefficients $c_{m,n}(\theta)$.

B-Spline Wavelets

B-Spline wavelets were independently constructed by Chui and Wang [1], and Unser et al [11]. The B-Spline wavelets of order n correspond to the multiresolution subspaces $\{S^n_{(i)}, i \in \mathbb{Z}\} = \{f \in L^2[\mathbb{R}] : f \in C^{n-1}$ and f is equal to a polynomial of degree n on each interval $[k2^i, (k+1)2^i]$ for all $k \in \mathbb{Z}\}$.

1. **Essential compact support in time:** The B-spline wavelets have compact support, so they would appear to be optimally suited to this criterion. The problem is that they form a biorthogonal system, and their biorthogonal pairs are not as compactly supported as the original wavelets. Thus one must gather more data to compute the coefficients than the support of the wavelet.
2. **Zero moments:** From the multiresolution subspaces of the B-Spline wavelets it follows that any function in a difference subspace $S^n_{(i+1)} \setminus S^n_{(i)}$ must be orthogonal to all polynomials of degree n and therefore the B-spline wavelets have n zero moments. This is good, but not optimal.
3. **Essential compact support in frequency:** It has been shown in [12] that the B-Spline wavelets are very close to being optimal in terms of time-frequency localization. Therefore we cannot ask much more of the combined conditions 1 and 3, although the broadened support of the biorthogonal pairs seems to cause problems.

The Lemarié-Battle Wavelets

The Lemarié-Battle wavelets are orthonormalized versions of the B-spline wavelets.

1. **Essential compact support in time:** The price of orthonormality is the loss of compact support in time. Their support is all of $\mathbb{R}$. They die out exponentially, however, so although they are not optimal in terms of this criterion, they score higher than the Meyer wavelet.
2. **Zero moments:** Since the Battle-Lemarié wavelets span the vsame subspaces as the B-spline wavelets, they too must have n zero moments. This is good, but not optimal.
3. **Essential compact support in frequency:** It has been shown [12] that the B-Spline wavelets are very close to being optimal in terms of time-frequency localization. Since the orthogonalization process that creates the Battle-Lemarié wavelets from the B-Splines does not increase the frequency support, the two wavelets are equivalent in the criterion.

The Daubechies Wavelets

No analysis would be complete without considering the optimality of the Daubechies compactly supported wavelets.

1. **Essential compact support in time:** The Daubechies wavelets D_N are compactly supported in time, with support length $(2N-1)$. This is optimal, although we will have to widen the support for the second criterion.
2. **Zero moments:** The Daubechies wavelets D_N have N vanishing moments. Thus we have a

classical tradeoff between criteria 1 and 2. The more vanishing moments we require, the wider the support of the wavelet. If there are not very many vanishing moments, the support of the wavelet will be small, but the filtered version of the wavelet will not die out rapidly and therefore will not be essentially compactly supported.

3. **Essential compact support in frequency:** The Daubechies wavelets are essentially localized in frequency, but do have small non-local nodes. These nodes adversely affect the band-limited nature of the wavelet coefficients. This can be seen through Theorem 2 because the first term in the estimate,

$$\|\hat{\psi}_{m,n}\mathcal{X}_{\mathbb{R}\setminus[-n+\delta,n-\delta]}\|_2$$

will remain small but non-zero long after the majority of the support of the wavelet is within $[-n+\delta, n-\delta]$. This can also be seen in the plot of inner products of the wavelets versus the Bessel functions in Figure 6. This fact could potentially cause undesireable aliasing, although we have not noticed significant errors in practice.

Numerical Comparisons

Numerical comparisons of the Daubechies and Battle-Lemarié wavelets are shown in Figure 3. Although the Battle-Lemarié wavelet is definitely more compactly supported in frequency, we have been using the Daubechies wavelets in our examples because of their compact support, and the essential compact support of their filtered versions.

XII. CONCLUSION

Through the properties wavelets, and through the connection between radial and angular band-limits presented in Theorem 2, our algorithm greatly reduces the radiation exposure required for local tomographic imaging. In addition, the image created is that of the true density function, not its lambda transform, nor that of the density function modulo the nullspace of the localization operator, as with SVD methods. It has been shown that the requirements of joint time-frequency localization and high-order zero moments specify that a wavelet type basis be used for the decomposition.

The results for synthetic data clearly show the advantages of the algorithm, and we predict similar experimental success with the genuine medical data we are currently acquiring. The reduction in x-ray exposure resulting from our algorithm are significant, and we have been able to achieve even further savings by combining our algorithm with an interlaced parallel sampling geometry, similar to that presented in [6]. Furthermore, it has been shown that there is a potential for constructing a variety algorithms similar to the lambda transform for selective edge detection and enhancement. We feel that the techniques developed here provide an important tool for use in medical imaging.

APPENDIX

To prove Theorem 2 we will begin with two lemmas.

Lemma 1 *If $\|\hat{\psi}\|_2 = 1$ and J_n is a Bessel function of the first kind, then for any $n' > n$ the following inequality holds:*

$$\langle\hat{\psi}, J_n\rangle \leq |\langle\hat{\psi}\mathcal{X}_{\mathbb{T}_{n'}}, J_n\rangle| + \pi^{-1/2}c(n')^{-\frac{1}{4}}e^{(-\frac{1}{3}n)c(n')^{\frac{3}{2}}} \quad (19)$$

where $\mathbb{T}_{n'} = \mathbb{R} \setminus [-n'+\sqrt{n'}, n'-\sqrt{n'}]$, and $c(n') = \frac{1}{1+\sqrt{n'}}$.

Proof: We begin with the simple inequality

$$\begin{aligned}\langle\hat{\psi}, J_n\rangle &= \int \hat{\psi}(t)J_n(t)dt \\ &\leq \int_{[-n'+\sqrt{n'},n'-\sqrt{n'}]} |\hat{\psi}(t)J_n(t)|dt + \\ &\quad \left|\int_{\mathbb{R}\setminus[-n'+\sqrt{n'},n'-\sqrt{n'}]} \hat{\psi}(t)J_n(t)dt\right| \quad (20)\end{aligned}$$

In order to investigate the first term,

$$\int_{[-n'+\sqrt{n'},n'-\sqrt{n'}]} |\psi(t)J_n(t)|dt$$

we need the Bessel function inequality

$$\begin{aligned}0 &\leq |J_n(\theta n)| \\ &\leq (2\pi n)^{-\frac{1}{2}}(1-\theta^2)^{-\frac{1}{4}}e^{(-\frac{n}{3})(1-\theta^2)^{\frac{3}{2}}} \quad (21)\end{aligned}$$

where $|\theta| < 1$ [6].

Applying the Cauchy–Schwartz inequality to this term yields

$$\begin{aligned}&\left(\int_{-n'+\sqrt{n'}}^{n'-\sqrt{n'}} |\hat{\psi}(t)J_n(t)|dt\right)^2 \\ &\quad \leq \int_{-n'+\sqrt{n'}}^{n'-\sqrt{n'}} |J_n(t)|^2dt \int_{-n'+\sqrt{n'}}^{n'-\sqrt{n'}} |\hat{\psi}(t)|^2dt \\ &\quad \leq \int_{\frac{-n'+\sqrt{n'}}{n}}^{\frac{n'-\sqrt{n'}}{n}} |J_n(n\theta)|^2 n d\theta \ \|\hat{\psi}\|_2^2.\end{aligned}$$

Applying (21) and using the fact that $\|\hat{\psi}\|_2 = 1$ we have

$$\left(\int_{-n'+\sqrt{n'}}^{n'-\sqrt{n'}} |\hat{\psi}(t)J_n(t)|dt\right)^2 \leq \int_{-n'+\sqrt{n'}}^{n'-\sqrt{n'}} |J_n(t)|^2dt$$

$$\leq \int_{\frac{-n'+\sqrt{n'}}{n}}^{\frac{n'-\sqrt{n'}}{n}} (2\pi n)^{-1}(1-\theta^2)^{-\frac{1}{2}} e^{(-\frac{2n}{3})(1-\theta^2)^{\frac{3}{2}}} n d\theta$$

$$\leq \int_{\frac{-n'+\sqrt{n'}}{n}}^{\frac{n'-\sqrt{n'}}{n}} (2\pi)^{-1}(1-\theta^2)^{-\frac{1}{2}} e^{(-\frac{2n}{3})(1-\theta^2)^{\frac{3}{2}}} d\theta.$$

Now we note that $1-\theta^2$ is bounded below by

$$c(n,n') \equiv 1 - \left(\frac{n'-\sqrt{n'}}{n}\right)^2.$$

For each fixed value of n' this is an increasing function of n and since by assumption $n > n'$ it will attain its minumum value for each n' at $n = n'+1$. We denote this value by $c(n')$ and from this it follows that the above quantity is dominated by

$$\frac{1}{\pi} c(n')^{-\frac{1}{2}} e^{(-\frac{2n}{3})c(n')^{\frac{3}{2}}} d\theta.$$

A quick check of $c(n')$ shows that it is a decreasing function of n', but that $(\frac{2n'}{3})c(n')^{\frac{3}{2}}$ is asymptotically an increasing function of n', which places a finite bound on these bounds. □

Lemma 2 *If J_n is a Bessel function of the first kind then*

$$\sum_{n>n'} \left|\langle \hat{\psi} \mathcal{X}_{\mathbb{T}_{n'}}, J_n \rangle\right|^2 \leq \frac{2}{\pi} \|\psi \hat{\mathcal{X}}_{\mathbb{T}_{n'}}\|_{2,w}^2.$$

Proof: We use the Fourier isometry to get

$$\sum_{n>n'} \left|\langle \hat{\psi} \mathcal{X}_{\mathbb{T}_{n'}}, J_n \rangle\right|^2 = \sum_{n>n'} \left|\langle \psi * \hat{\mathcal{X}}_{\mathbb{T}_{n'}}, T_n(x)(1-x^2)^{-1/2} \rangle\right|^2 \quad (22)$$

where $T_n(x)$ is a Chebyshev polynomial of the first kind. The Chebyshev polynomials form an orthogonal system in the weighted space $L^2([-1,1],(1-x^2)^{-1/2})$, which we can normalize with multiplication by $\frac{\pi}{2}$. Bessel's inequality applied to the above then yields

$$\sum_{n>n'} \left|\langle \hat{\psi} \mathcal{X}_{\mathbb{T}_n}, J_n \rangle\right|^2 \leq \frac{2}{\pi} \|\psi * \hat{\mathcal{X}}_{\mathbb{T}_{n'}}\|_{2,w}^2. \qquad \square$$

Proof of Theorem 2: We recall that any set of projections has a special expansion

$$P_f(\theta,t) = (1-t^2)^{-1/2} \sum_n h_n(\theta) T_n(t)$$

where $h_n(\theta) = \sum_{n' \in H_n} a_{n,n'} e^{in'\theta}$ with $H_n = -n, -n+2, -n+4 \ldots n-2, n$. Now the wavelet coefficients are given by

$$c_{m,n}(\theta) = \langle \psi_{m,n}(t), (1-t^2)^{-1/2} \sum_n h_n(\theta) T_n(t) \rangle.$$

If we apply the Fourier isometry in the variable t, and use the fact that the weighted Chebyshevs will transform to Bessel functions of the first kind we have

$$\begin{aligned} c_{m,n}(\theta) &= \langle \hat{\psi}_{m,n}(t), (\pi/2)^{1/2} \sum_n h_n(\theta) i^n J_n(t) \rangle \\ &= (\pi/2)^{1/2} \sum_n h_n(\theta) \langle \hat{\psi}_{m,n}(t), i^n J_n(t) \rangle \\ &= (\pi/2)^{1/2} \sum_n \sum_{n' \in H_n} a_{n,n'} e^{in'\theta} \langle \hat{\psi}_{m,n}(t), i^n J_n(t) \rangle. \end{aligned}$$

We would like to rearrange the terms in the last sum to examine the trigonometric coefficients of $c_{m,n}(\theta)$. We have

$$\begin{aligned} c_{m,n}(\theta) &= \\ &(\pi/2)^{1/2} \sum_n \sum_{n' \in H_n} a_{n,n'} e^{in'\theta} \langle \hat{\psi}_{m,n}(t), i^n J_n(t) \rangle \\ &= (\pi/2)^{1/2} \sum_{n'} b_{n'} e^{in'\theta} \end{aligned}$$

where

$$b_{n'} = \left(\sum_{n > |n'|} a_{n,n'} \langle \hat{\psi}_{m,n}(t), i^n J_n(t) \rangle \right).$$

We need to show that the terms $b_{n'}$ decay quickly as $|n'| \to \infty$.

Applying Lemma 1 we get

$$\begin{aligned} |b'_n| &= \left| \sum_{n > |n'|} a_{n,n'} \langle \hat{\psi}_{m,n}(t), i^n J_n(t) \rangle \right| \\ &\leq \sum_{|n| > |n'|} |a_{n,n'}| \left(|\langle \hat{\psi} \mathcal{X}_{\mathbb{T}_{n'}}, J_n \rangle| \right. \\ &+ \left. \pi^{-1/2} c(n')^{-\frac{1}{4}} e^{(-\frac{1}{3}n)c(n')^{\frac{3}{2}}} \right) \\ &\leq \sum_{n > |n'|} \left(|a_{n,n'}| \left| \langle \hat{\psi}_{m,n} \mathcal{X}_{\mathbb{T}_{n'}}, J_n \rangle \right| \right. \\ &\left. + |a_{n,n'}| \pi^{-1/2} c(n')^{-\frac{1}{4}} e^{(-\frac{1}{3}n)c(n')^{\frac{3}{2}}} \right) \quad (23) \end{aligned}$$

Now the terms $a_{n,n'}$ are the coefficients of an orthogonal expansion of $P_f(\theta,s)$, in the weighted

space $L^2([-1,1],(1-s^2)^{-1/2})$. Applying the Cauchy Schwartz inequality and Lemma 2 to (23) shows that

$$\begin{aligned}
&\left| \sum_{n>|n'|} a_{n,n'} \langle \hat{\psi}_{m,n}(t), i^n J_n(t) \rangle \right| \\
\leq\ & \frac{2}{\pi} \|P_f(\theta,t)\|_{2,w} \|\psi_{m,n} \widehat{\mathcal{X}}_{\mathbb{T}_{n'}}\|_{2,w} \\
& + \frac{2}{\pi} \|P_f(\theta,t)\|_{2,w} \pi^{-1} \sum_{n>|n'|} c(n')^{-\frac{1}{2}} e^{(-\frac{2}{3}n)c(n')^{\frac{3}{2}}} \\
=\ & \frac{2}{\pi} \|P_f(\theta,t)\|_{2,w} (\|\psi_{m,n} \widehat{\mathcal{X}}_{\mathbb{T}_{n'}}\|_{2,w} \\
& + c_1(n') e^{(-\frac{2}{3}n)c(n')^{\frac{3}{2}}}). \qquad (24)
\end{aligned}$$

The final term in the above is due to the fact that exponential bounds remain exponential bounds upon summation.

We would like an ordinary L^2 bound for the terms on the right. We can do this simply if we assume that supp $[P_f(\theta,t)] \subseteq S^1 \times [-1+\epsilon, 1-\epsilon]$. Then the bound M is merely the maximum of the weight on the support of the function, times the constant from the previous inequality, and we are done. □

REFERENCES

[1] C.K.Chui, and J.Z.Wang, On compactly supported spline wavevlets and a duality principle. CAT rep. 213, Texas A&M Univ., 1990. (To appear *Trans. Am. Math Soc.*).

[2] I.Daubechies, *Ten Lectures on Wavelets*, SIAM-CBMS Series, 1992.

[3] I.Daubechies, Orthonormal Bases of Compactly Supported Wavelets, Comm. Pure and Applied Math, XL1, 1988.

[4] J.Destefano and T.Olson, Wavelet Localization of the Radon Transform in Even Dimensions, Proc. IEEE-SP Int. Symp. of Time-Frequency and Time-Scale Analysis, (in press).

[5] S.Mallat, Wavelet transform maxima and multiscale edges, in *Wavelets and Their Applications*, Mary Beth Ruskai et al, eds., Jones and Bartlett, Cambridge, MA 1991.

[6] F.Natterer, *The Mathematics of Computerized Tomography*, John Wiley, 1986.

[7] A. Faridani, E. Ritman, and K.T.Smith, Local Tomogaphy, Vol.52, No.2, pp.459-484, April 1992, SIAM J Appl Math.

[8] A. Faridani, F. Keinert, F. Natterer, E.L. Ritman, and K.T. Smith, Local and global tomography, *Signal Processing*, IMA Vols in Math and Appl, vol 23, Springer Verlag, New York, 1990, pp 241-255.

[9] P.Maass, The Interior Radon transform, SIAM J. APPL. MATH, Vol.52, No.3, pp. 710-724, June 1992.

[10] N. Srinivasa, K.R. Ramakrishanan, and K. Rajgopal, Detection of Edges from Projections, IEEE Trans. on Med. Imaging, Vol.11, No.1, March 1992.

[11] M.Unser, A. Aldroubi, and M.Eden, "A family of polynomial spline wavelet transforms." NCRR rep. 153/90, Nat. Inst. Health.

[12] M.Unser, A.Aldroubi, and M.Eden, "On the Asympototic Convergence of B-Spline Wavelets to Gabor Functions", IEEE Trans on Inf. Theory, Vol.38, No.2, March 1992.

[13] D.Walnut, Applications of Gabor and Wavelet Expansions to the Radon Transform, preprint.

[14] A.K. Louis and A. Reider, "Incomplete Data Problems in X-Ray Computerized Tomography", Numeriche Mathemtik, **56**, 1989.

[15] G.N. Ramachandran and A.V. Lakshminarayanan, "Three dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms", Proc. Nat. Acad. Sci., 68, 2236-2240, 1971.

[16] R.M. Lewitt, R.H.T. Bates, and T.M. Peters, "Image reconstruction from projections II: Modified backprojection methods", Optik, 50, 180-205.

[17] O.J. Tretiak, "The point spread function for the convolution algorithm in R.Gordon (ed.), Image processing for 2-D and 3-D reconstructions from projections: Theory and practice in medicine and the physical sciences." Digest of technical papers, Stanford, California,

August 4-7, 1975, Sponsored by the Opt. Soc. of Am.

[18] A. Faridani, "Reconstruction from efficiently sampled data in parallel-beam computed tomography", G.F. Roach (ed.): Inverse problems and imaging, Pitman Research Notes in Mathematics, Vol. 245, pp. 68-102, 1991.

[19] A.M. Cormack, "Sampling the Radon transform with beams of finite width", Phys. Med. Biol., 23, 1141-1148, 1978.

[20] A. Faridani, "Sampling and resolution in diffraction tomography II: An error analysis of the filtered backpropogation algorithm", To appear, SIAM J. Appl. Math.

[21] A.G. Lindgren and P.A. Rattey, "The inverse discrete Radon transform with applications to tomographic imaging using projection data", Adv. in Elec. and Electron Phys, **56**, 359-410, 1981.

[22] R.M. Lewitt, "Multidimensional digital image representations using generalized Kaiser-Bessel window functions", J. Opt. Soc. Am. A, **7**, No. 10, Oct. 1990.

[23] C.H. Chapman and P.W. Cary, "The circular harmonic Radon transform", Inverse Problems **2**, 1986, 23-49.

[24] W.E. Higgins and D.C. Munson, "A Hankel tranform approach to tomographic image reconstruction", IEEE Trans. on Med. Im., March 1988.

[25] C. Hamaker, K.T. Smith, D.C. Solmon, S.L. Wagner: "The divergent beam X-ray transform." Rocky Mountaion J. Math, **10**, 253-283, 1980.

[26] P.Auscher, G.Weiss, and M.V.Wickerhauser, "Local Sine and Cosine Bases of Coifman and Meyer and the Construction of Smooth Wavelets", in Wavelets, A Tutorial in Theory and Applications, C.Chui, Ed., AP, 1990.

[27] A. Faridani, E. Ritman, and K.T.Smith, Examples of Local Tomogaphy, Vol.52, No.4, pp.1193-1198, August 1992, SIAM J Appl Math.

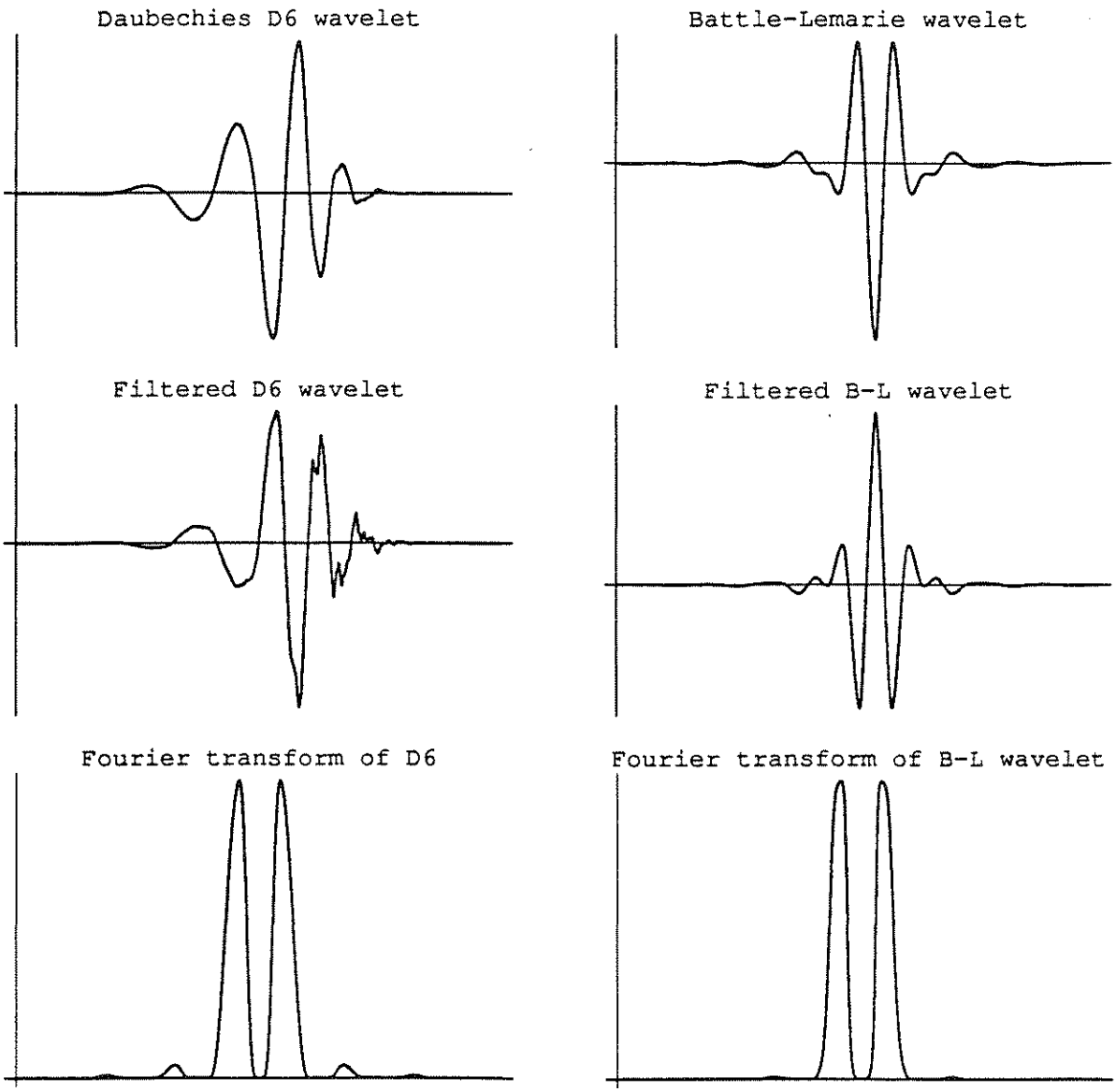


Figure 3: Two common wavelets, their filtered versions, and their Fourier transforms. The left column shows the Daubechies D_6 wavelet, and the right column the Battle-Lemarie wavelet. Note that in both cases the effective time-domain support is unchanged by the filter $\mathcal{F}$. The last row shows the magnitude of the Fourier transforms. The two main lobes of both wavelets fall completely inside the 2π band limit, while the smaller side lobes lie outside it.

Target size ->	0.0625	0.125	0.25	0.375	0.5
Image size					
128	72:86	66:83	54:77	43:72	32:66
	57:79	52:76	41:71	32:66	23:61
	48:78	43:71	34:67	24:62	16:58
256	81:90	75:87	62:81	51:75	39:69
	70:85	65:82	53:77	42:71	31:66
	63:82	58:79	47:73	36:68	26:63
512	86:93	80:90	68:84	55:78	43:72
	79:90	73:87	62:81	50:75	38:69
	74:87	69:84	57:78	45:73	34:67
1024	89:95	83:92	71:85	58:79	46:73
	85:93	79:90	67:83	55:77	43:71
	82:91	76:88	64:82	52:76	40:70
2048	91:96	85:93	73:86	60:80	48:74
	89:94	83:91	70:85	58:79	46:73
	87:93	81:90	66:84	56:78	44:72

Figure 4: Table of savings realized by wavelet localization. For each image size and target size there are three entries, one for D2, one for D4, and one for D6. The number before the colon is the savings without interlaced sampling; the number after the colon is the savings with interlaced sampling.

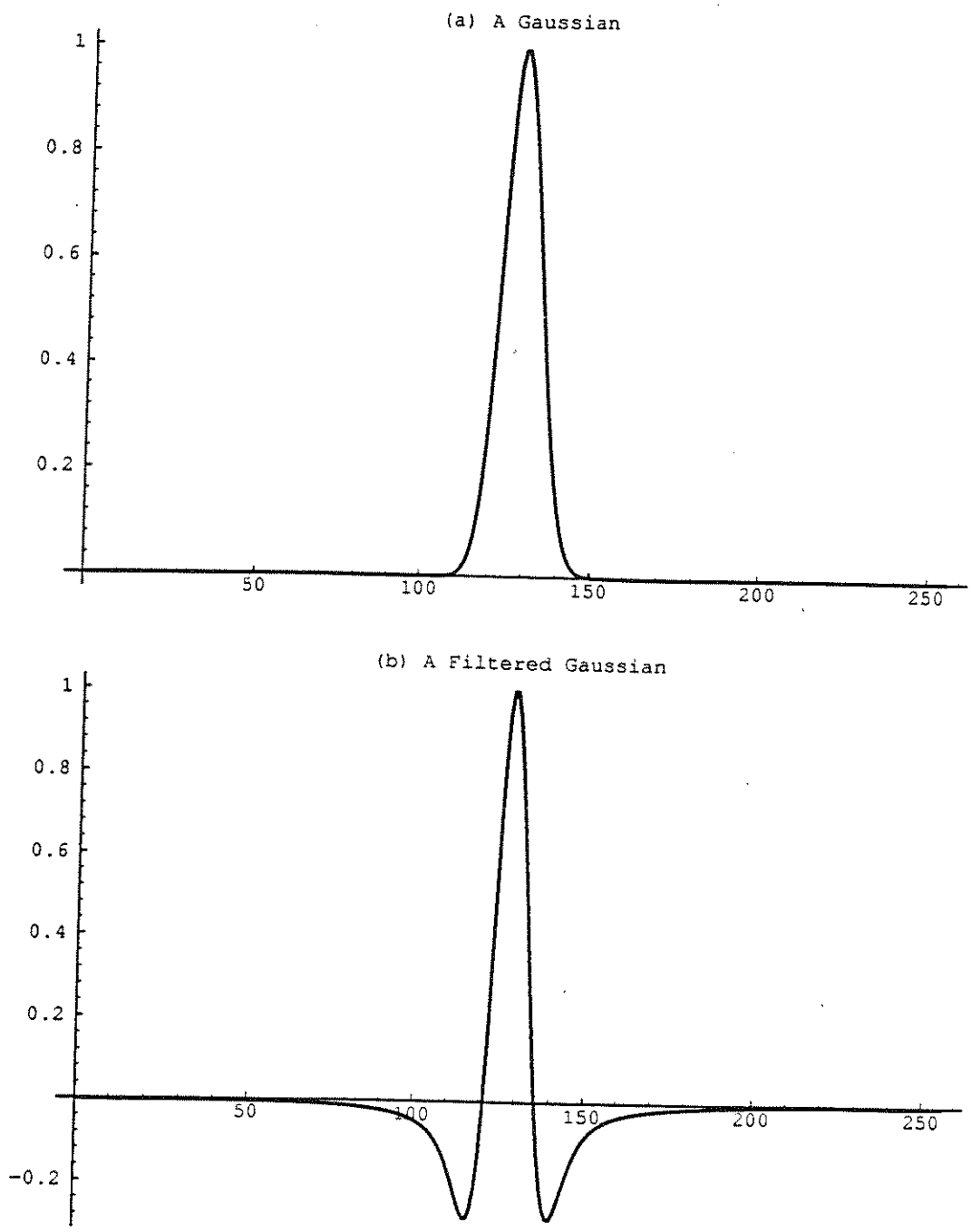


Figure 5: The non-locality of the filtering is illustrated with the original function, a Gaussian, and its filtered version, which is spread noticeably by the filtering process. To be more precise, there are 42 values in the original Gaussian which are above a threshold of .001, and all 256 values in the filtered version are above .001.

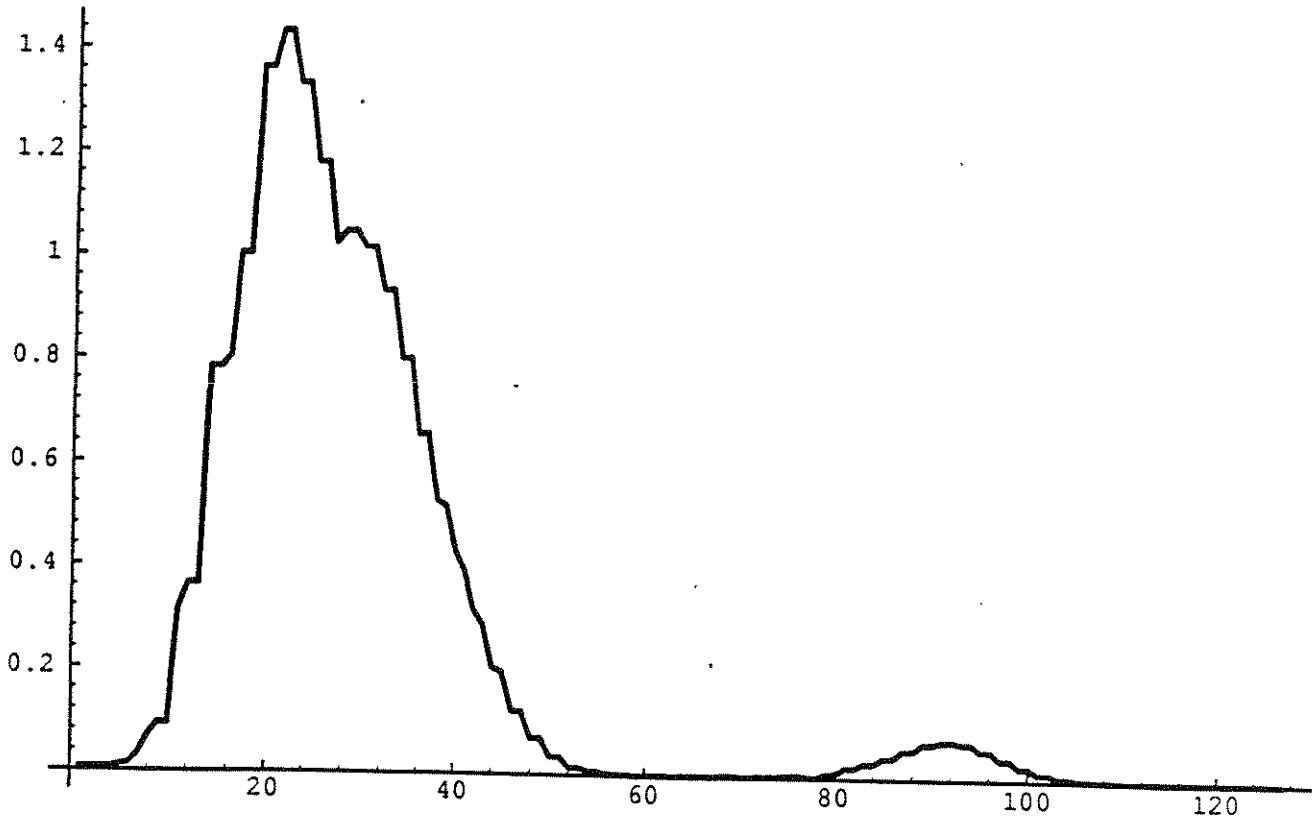


Figure 6: A plot of the inner products $|\langle \psi_{m,n}, J_l \rangle|$ as l increases. The figure is actually a "silhouette" of the true function, which fluctuates too much to produce a legible plot. Each value of the original was replaced by the maximum of it and its two neighbors.

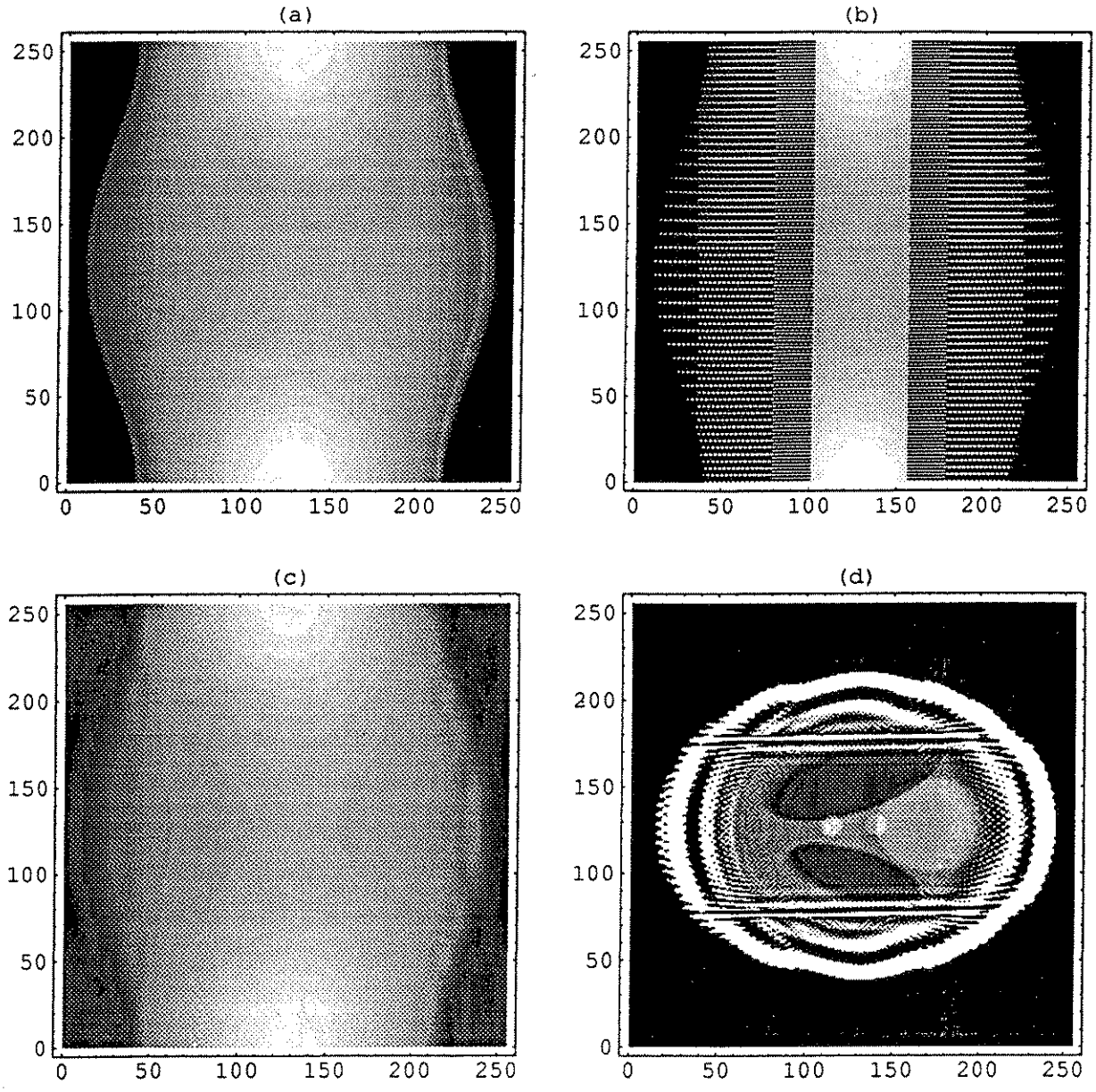


Figure 7: The wavelet sampling scheme and algorithm is illustrated by the the examples above. The standard sampling scheme is shown in (a), and the wavelet sampling scheme is shown in (b). In (c) the interpolated sinogram is shown and in (d) is a reconstruction using the interpolated sinogram.

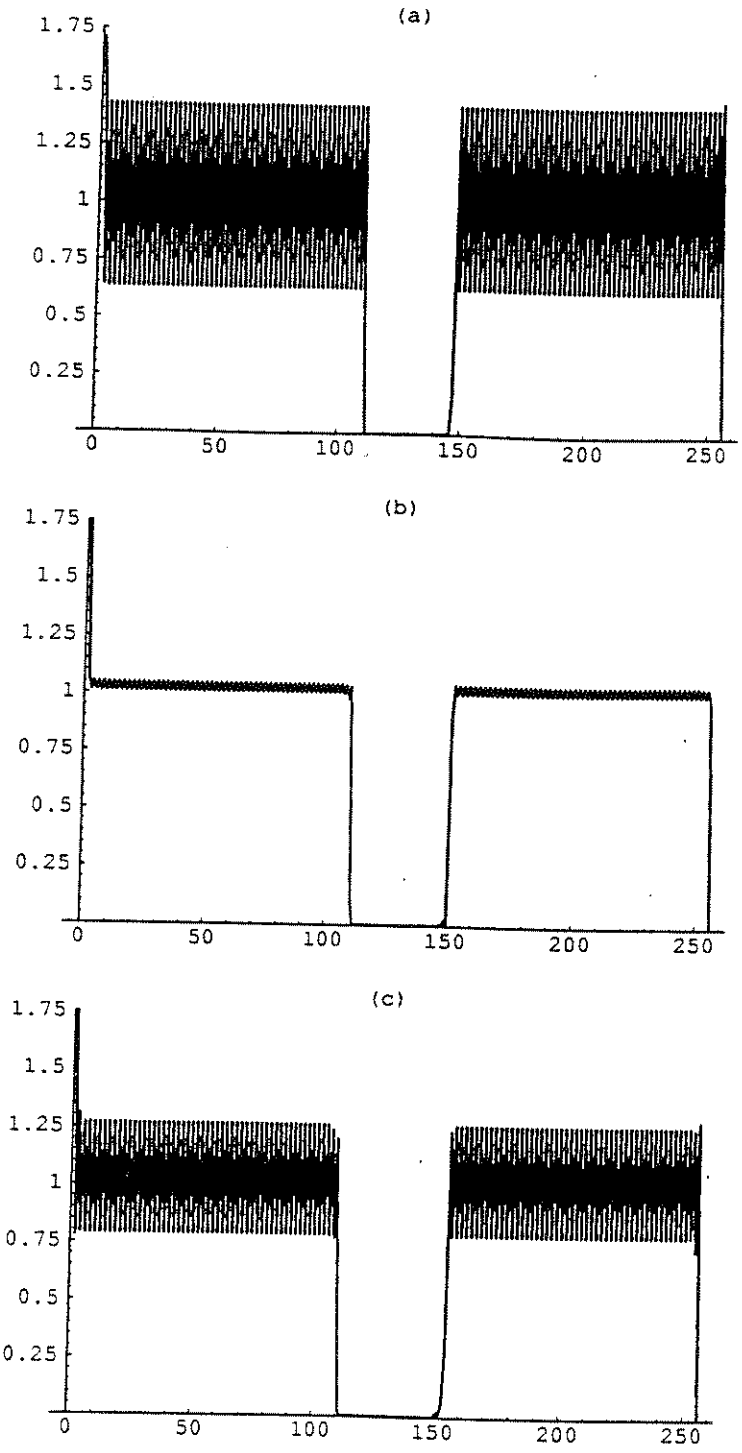


Figure 8: The error bound from (15),

$$\max_{t \in S} \left(\sum_{\mathcal{NL}} |\mathcal{F}(\psi_{n,m}(t))|^2 \right)^{1/2}$$

is plotted below. The fractal behavior of the D2 and D6 errors, compared to the D4 errors, away from the region of interest is of no practical interest, yet is quite puzzling.

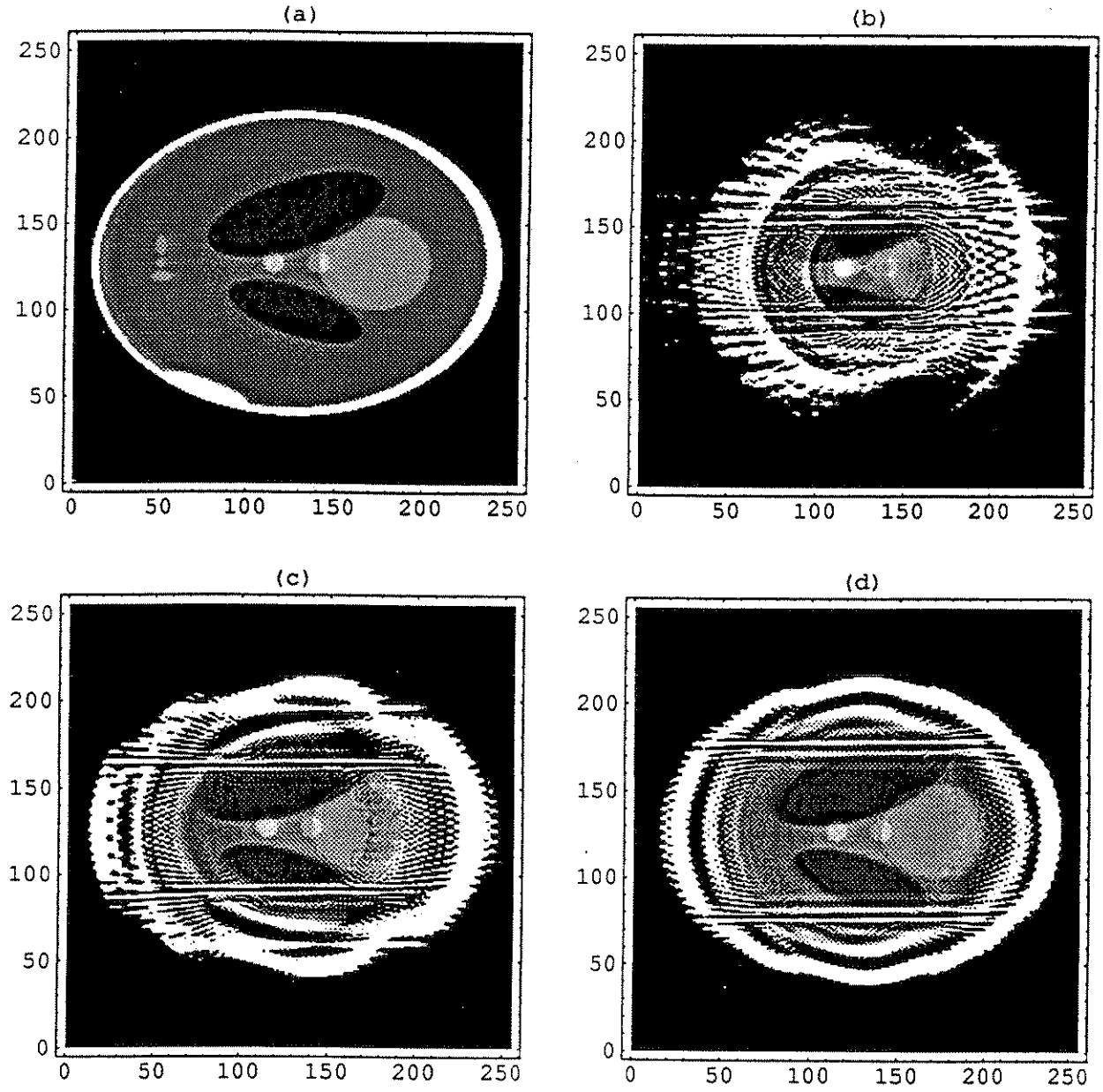


Figure 9: We illustrate the aliasing caused in the interpolation portion of the algorithm above. The reconstruction of the Shepp-Logan phantom with full data (a) is contrasted against the reconstruction of a circle with radius 1/8, using the D2 (b), D4 (c), and (d) D6 wavelets. Note the errors caused by interpolation in the D2 case. The D4 reconstruction is very good in the interior 1/8 region of interest, and the D6 reconstruction is pixel for pixel exact in the region of interest.

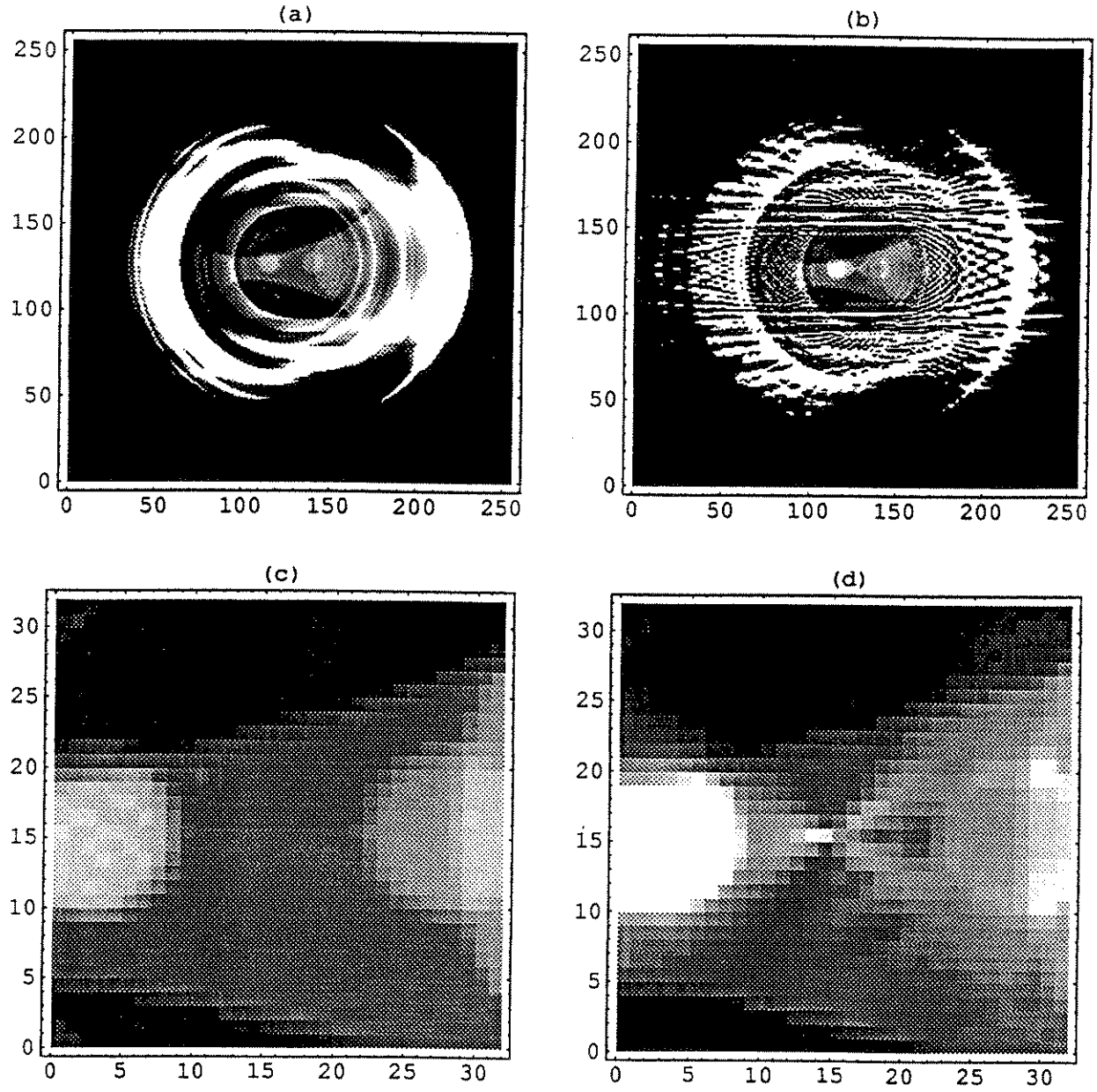


Figure 10: To illustrate that the errors in the reconstruction using the D2 wavelet in Figure 9 were caused by aliasing, we calculated the local approximation using the exact local coefficients, with no interpolation. In (a) this exact local approximation is used and in (b) the reconstruction is done from local data, with interpolation. We show the region of interest in (c) and (d).

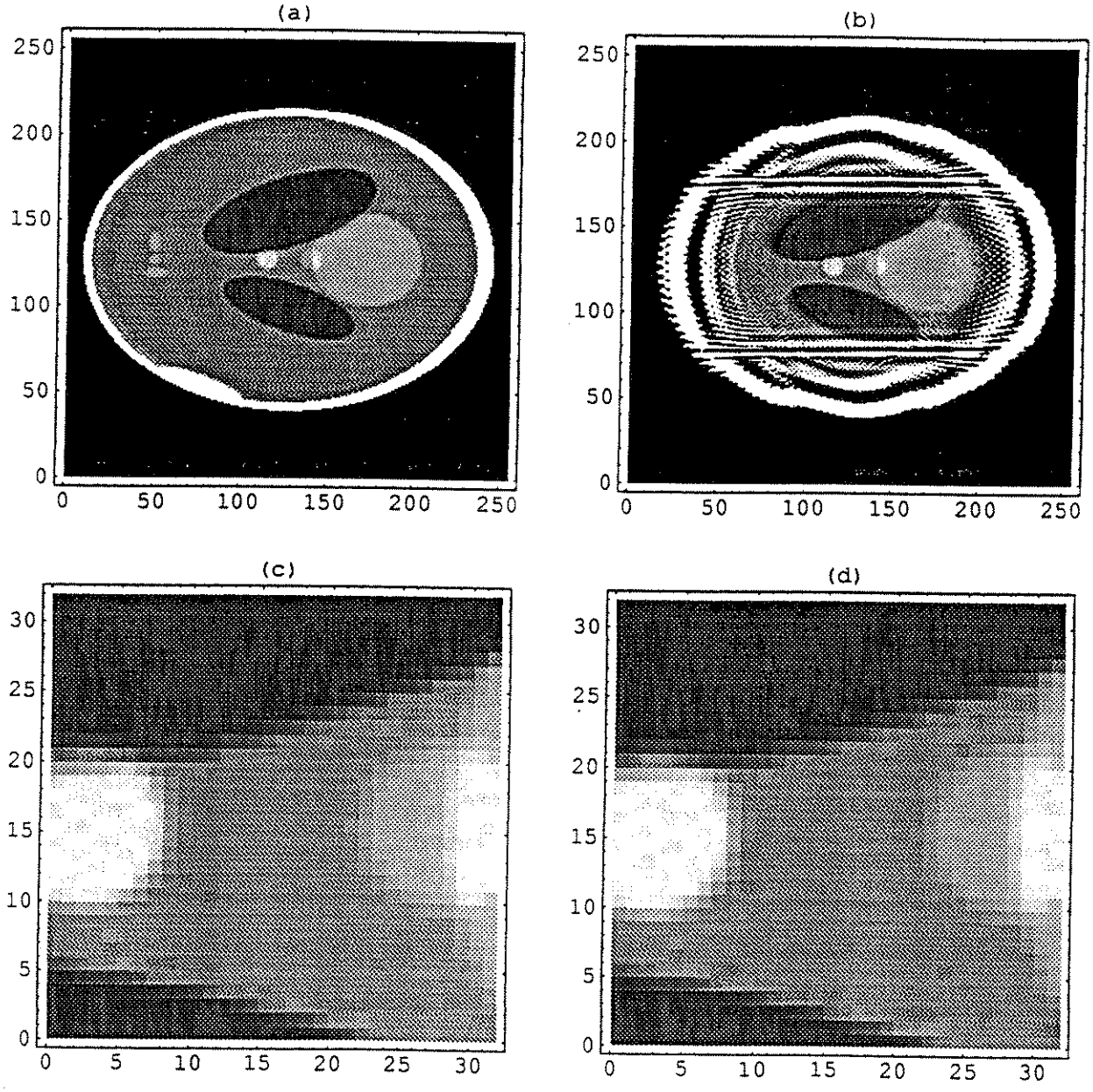


Figure 11: Above are reconstructions of the Shepp-Logan phantom: (a) From a full set of data, and (b) From 42% of the data with our algorithm, imaging a central circle of radius 1/8. A blow up of the corresponding central sections are shown in (c) With full data, and in (d) With 42% of the data and our algorithm. The circles contained inside these central sections are pixel for pixel identical. In these examples we were using Daubechies' D6 wavelet.

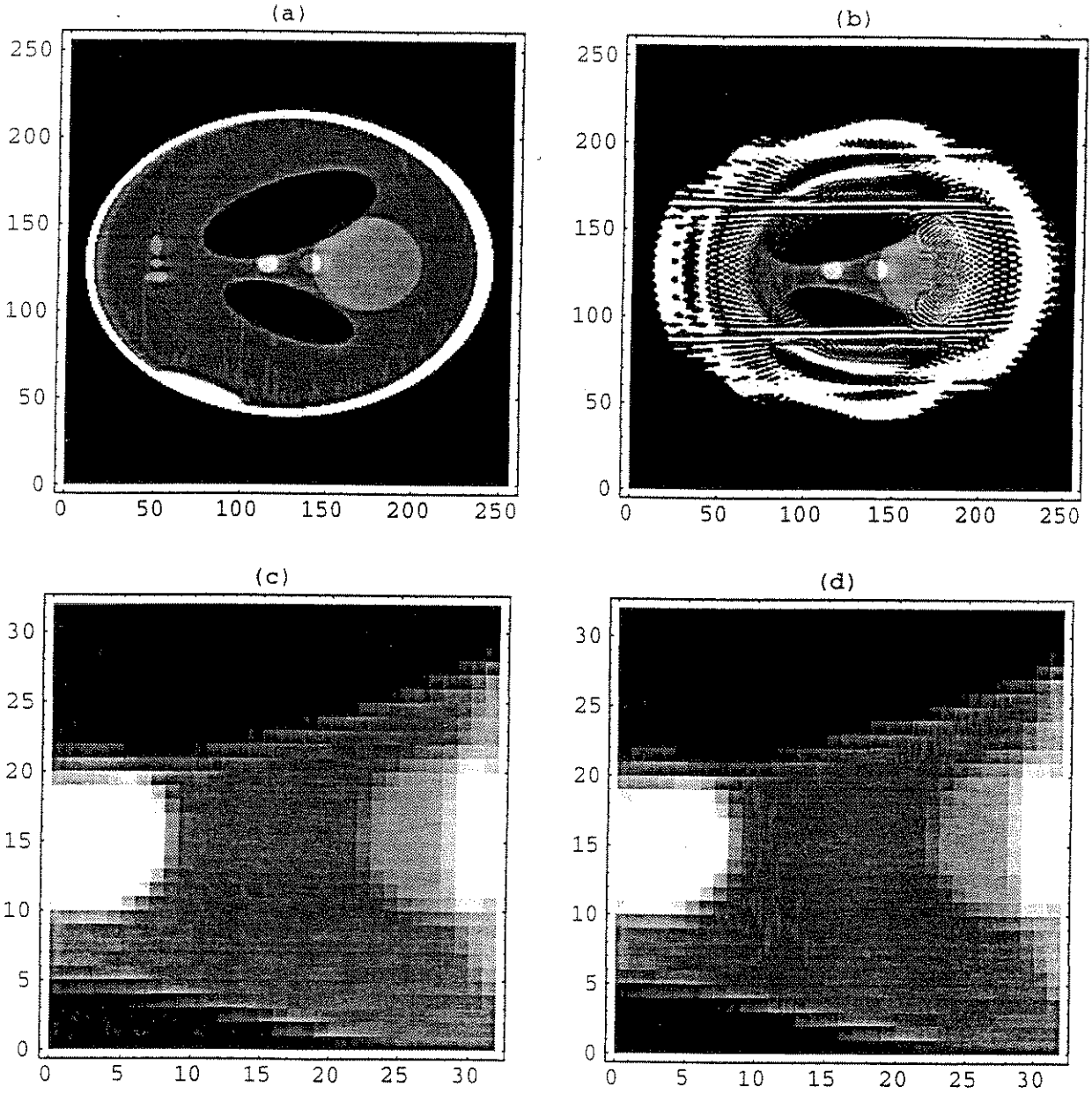


Figure 12: Above are reconstructions as in Figure 11, but this time using the Daubechies D4 wavelet. (a) From a full set of data, and (b) Using 35% of the data and our algorithm, to image a central circle of radius 1/8. A blow up of the corresponding central sections are shown in (c) With full data, and in (d) With 35% of the data and our algorithm. The additional errors can be attributed to a lower number of zero moments, and the aliasing induced by the loss in frequency localization.

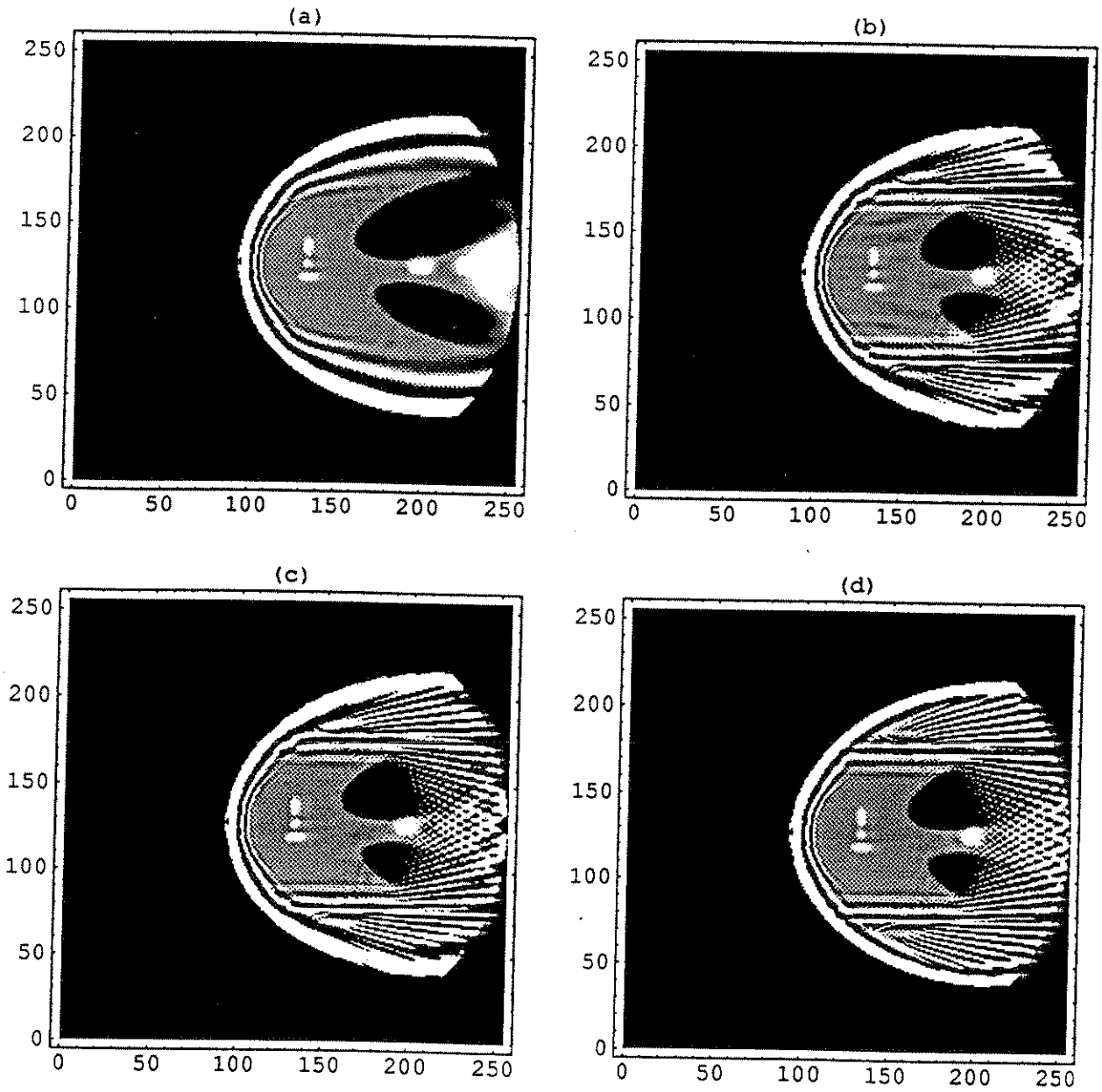


Figure 13: We reconstruct an off center region of our original image. The aliasing errors are increased in the off center regions, but can be combated by interpolating at fewer scales. The figure shows (a) the algorithm with perfect interpolation for reference, (b) with as many scales as possible interpolated, for maximum savings of 50%, (c) with fewer scales interpolated and savings of around 48%, and (d) with fewer scales yet interpolated and savings of around 45%. Thus the interpolation aliasing can be controlled by increasing the amount of data slightly.

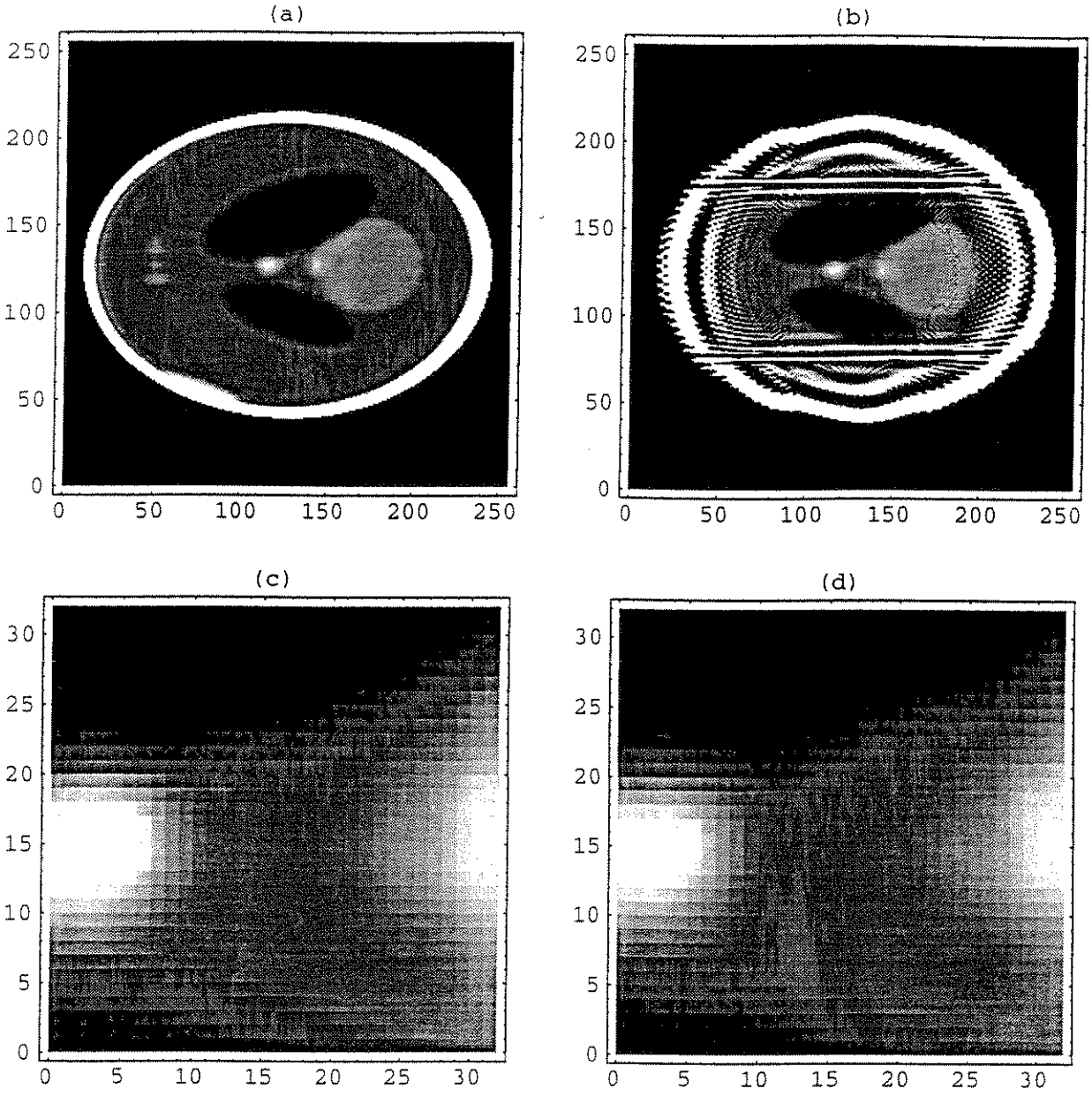


Figure 14: We illustrate the use of interlaced sampling with filtered backprojection (a) With full interlaced data, corresponding to 50% of the standard data, and (b) With local interlaced data, which is only 22% of the standard data. The corresponding close ups of these are shown in (c) and (d) respectively. Notice that there is a slight loss of resolution even with full interlaced data. This is primarily due to approximating the two dimensional sinc interpolations from the original sampling theorem as in [6].

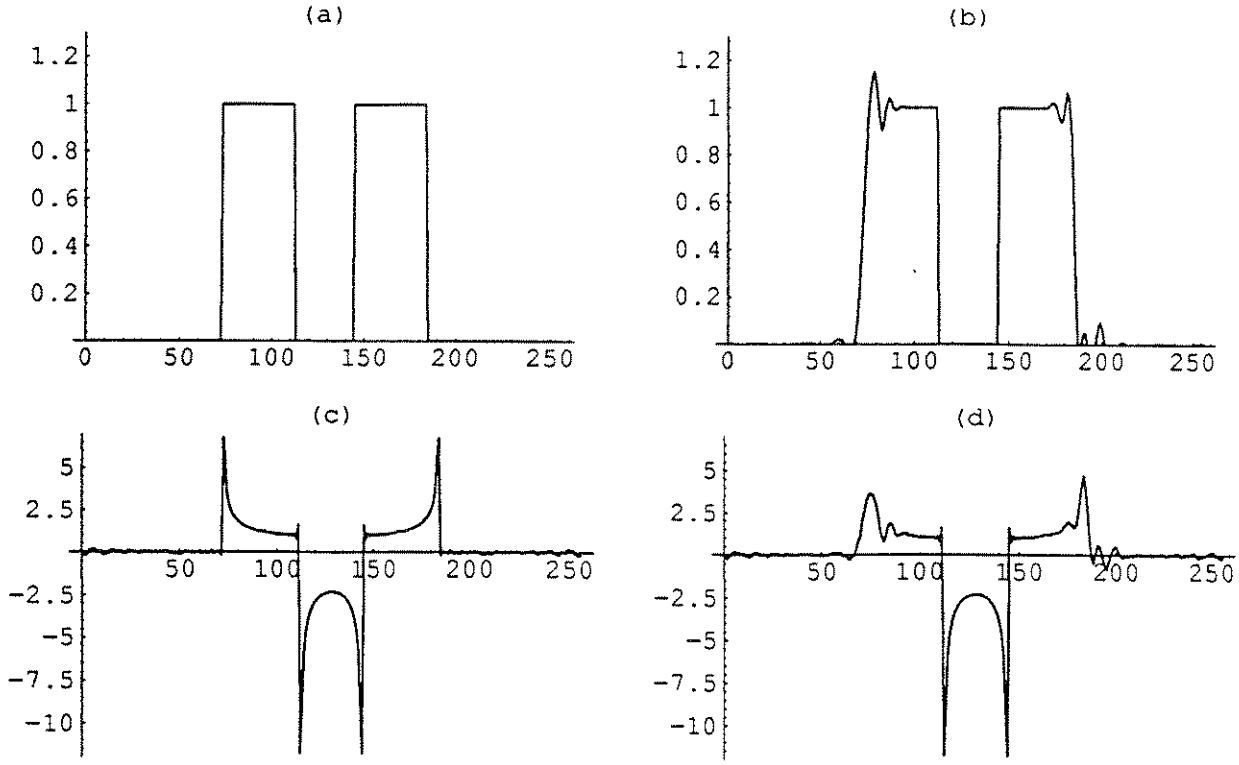
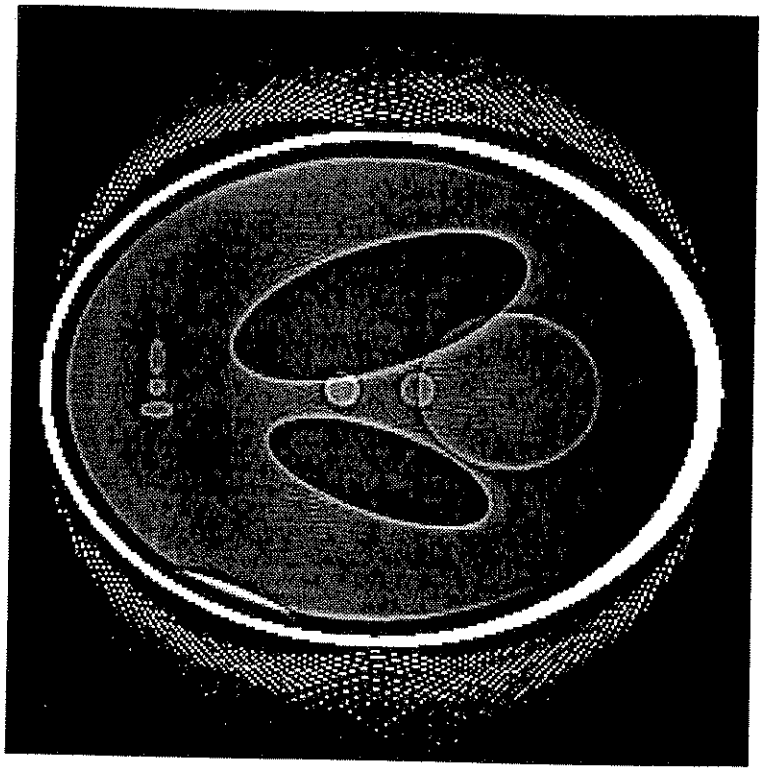


Figure 15: The projection of a circularly symmetric element of the nullspace is illustrated in (a), with the "sparse data" approximation to it from our algorithm in in (b). In (c), a cross section of a reconstruction of the nullspace element is shown, and in (d) a cross section of the local reconstruction of the nullspace element, using our algorithm is shown. Notice that (c) and (d) are essentially identical within the region of interest. The maximum error in this region is .0003.

(a)



(b)

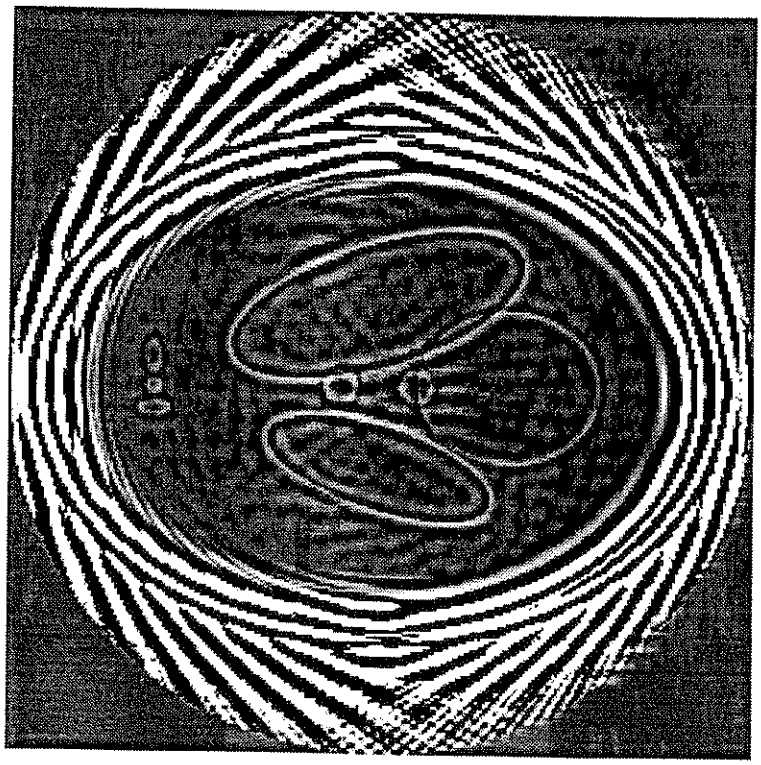


Figure 16: We illustrate the high pass wavelet filter (a) against Lambda-tomography (b). The wavelet high pass filter does induce local maxima and minima which could possibly be considered undesireable. On the other hand, it has been proven that lambda-tomography does not induce local maxima and minima.

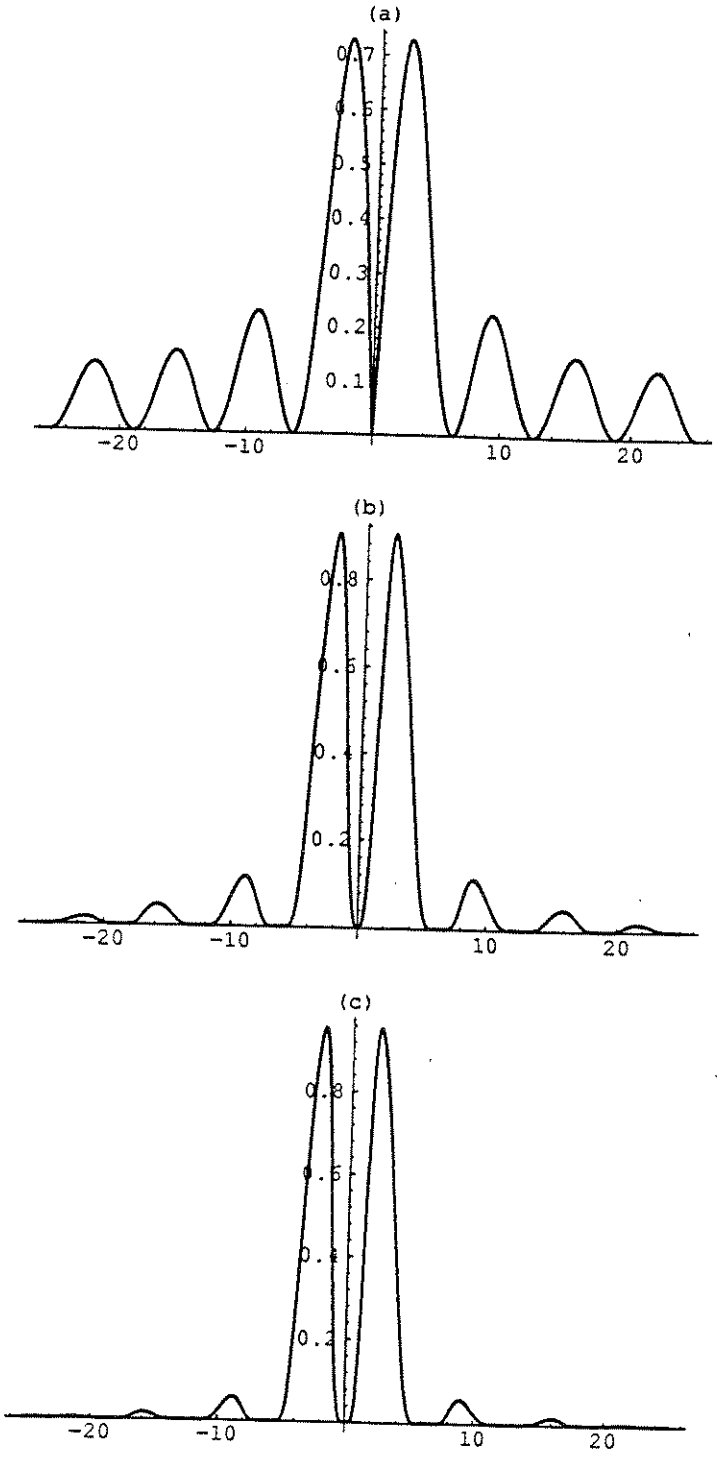


Figure 17: To illustrate the aliasing errors which occur with the D2, D4, and D6 wavelets we plot the absolute values of the Fourier transforms of (a) the D2, (b) the D4, and (c) the D6 wavelets. Since the D6 wavelet shows the most frequency localization it is no surprise that it performed best in our numerical examples. The D2 is not all that localized in frequency, so the aliasing from the interpolation of its coefficients is very noticeable.

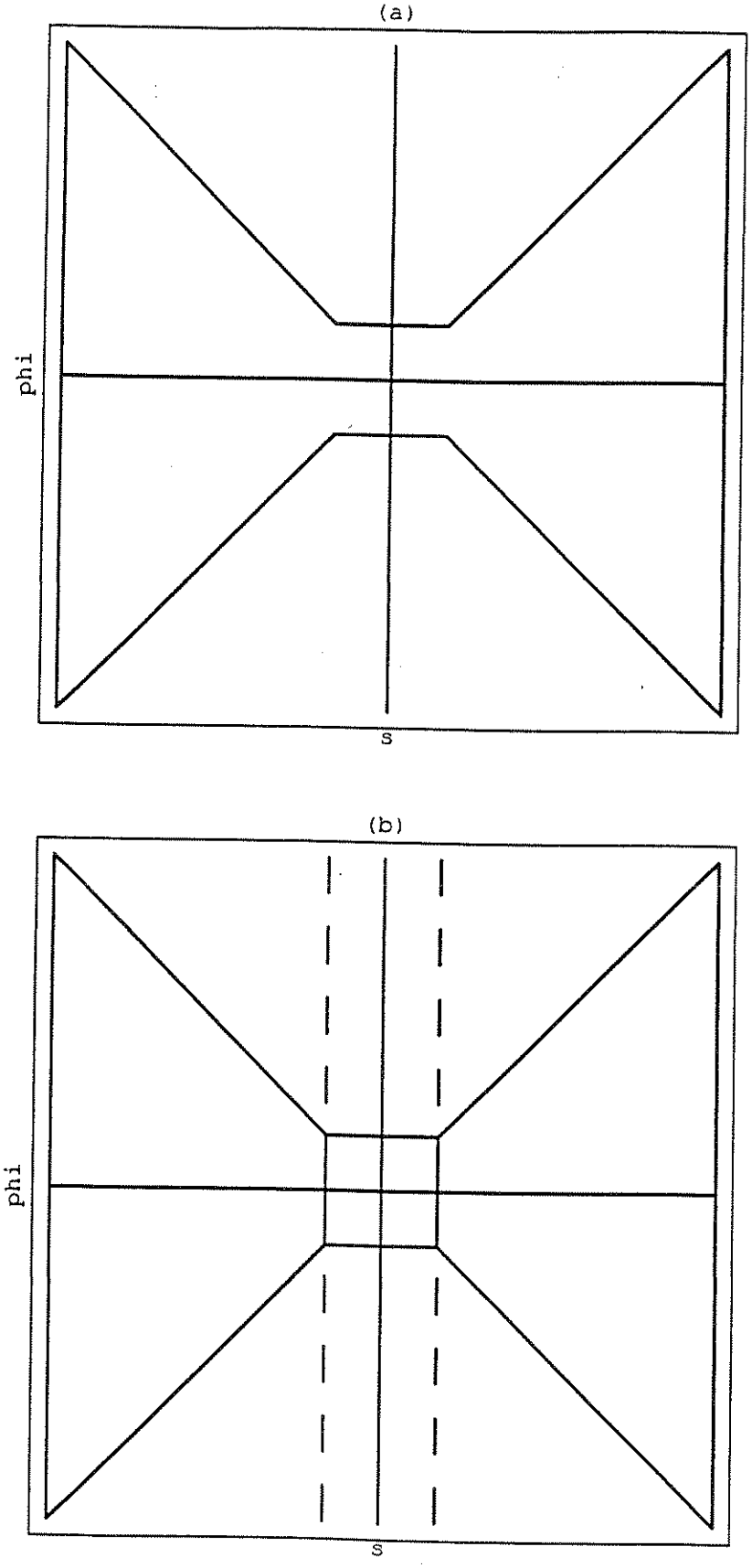


Figure 18: The support of the Fourier transform the Radon transform of a function f which is supported in the unit circle is illustrated in (a). In (b) we illustrate the result of imposing a radial window $w(s)$ on the Fourier transform $\hat{f}(s\vec{\theta})$.

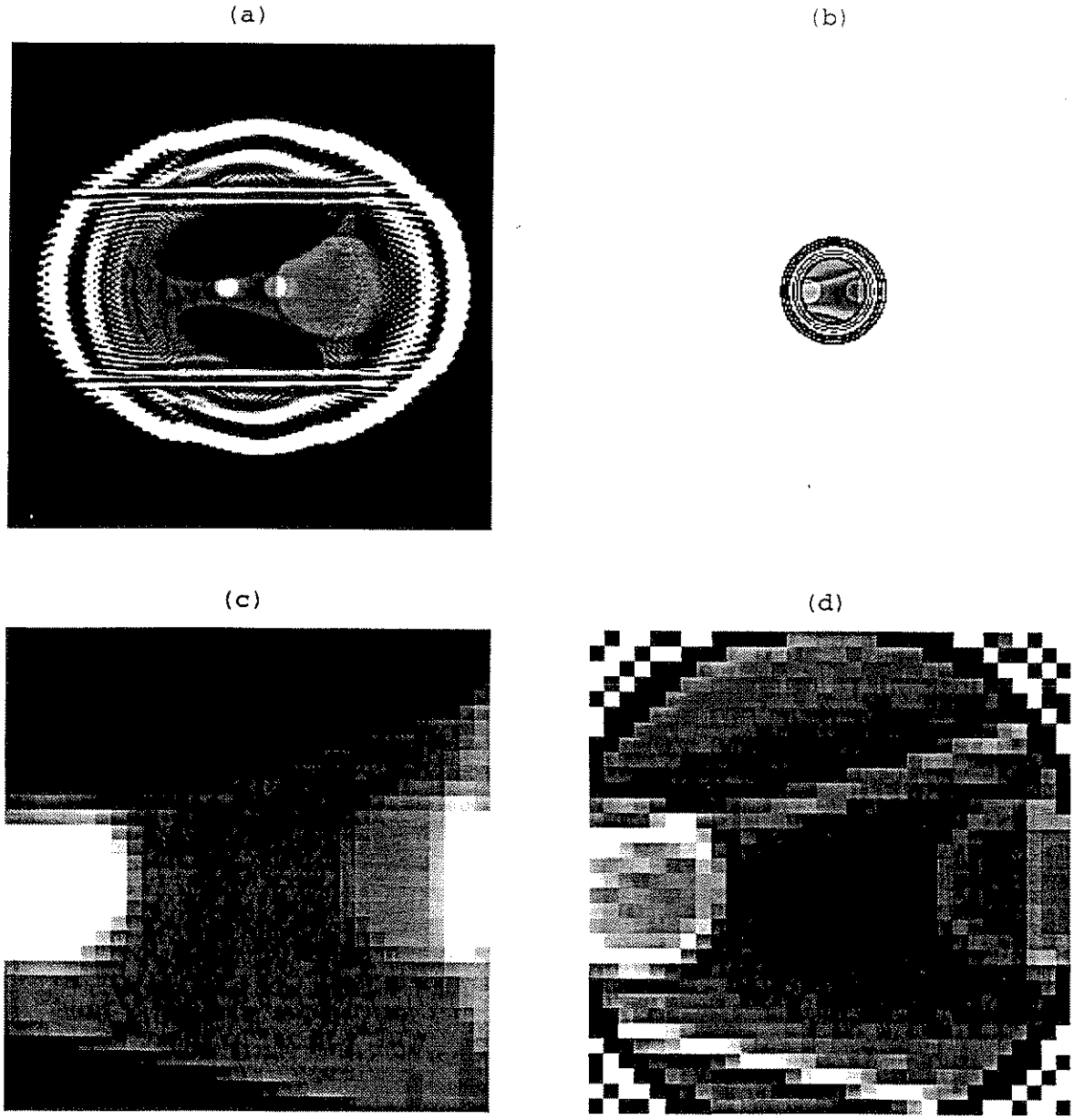


Figure 19: Local reconstructions of (a) f on a local region about the origin with our algorithm, and (b) Λf on a local region about the origin. The region of interest, which is a circle imbedded in the square emphasized in (c), with our algorithm, and (d) with Lambda-tomography. The cup correction of $\Lambda^{-1} f$ did not seem necessary for this local reconstruction. One must note that although our algorithm recovers the original image, we need more data to accomplish this than is required by Lambda-tomography.

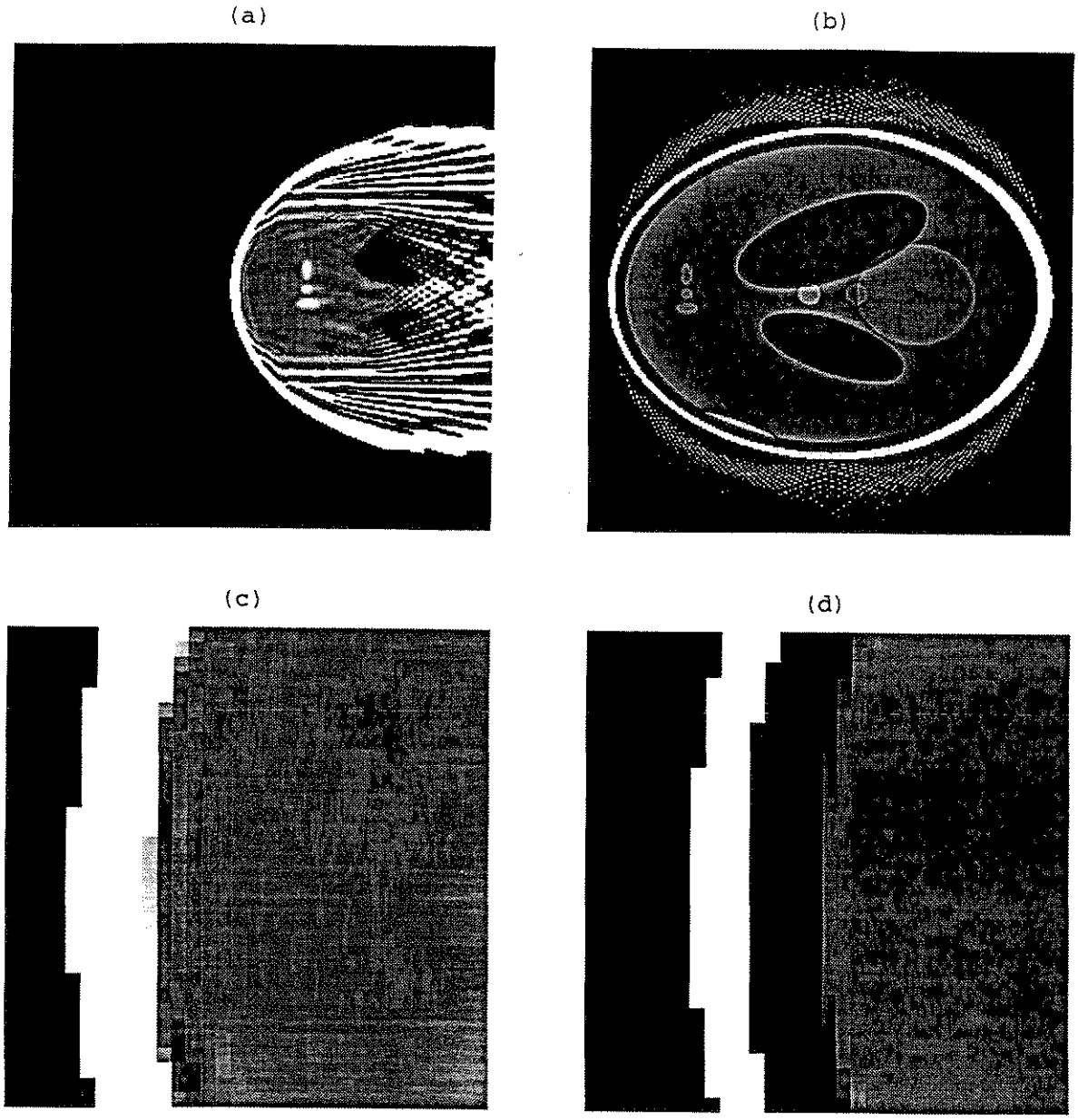


Figure 20: We compare a local reconstruction of the edge of the phantom, (a) using our algorithm, and (b) using Λ-tomography with cup correction. In (b) we have reconstructed $\Lambda f - \mu \Lambda^{-1} f$, where μ was chosen using a least squares criterion to make the interior of the region essentially constant. The large ring artifact at the edge is highlighted in (c) where our algorithm does not contain this artifact, and in (d) where the artifact remains, despite the cup correction.