# A Wavelet-Based Method for Multiscale Tomographic Reconstruction

M. Bhatia, W. C. Karl,* *Member, IEEE,* and A. S. Willsky, *Fellow, IEEE*

*Abstract*— **We represent the standard ramp filter operator of the filtered-back-projection (FBP) reconstruction in different bases composed of Haar and Daubechies compactly supported wavelets. The resulting multiscale representation of the ramp-filter matrix operator is approximately diagonal. The accuracy of this diagonal approximation becomes better as wavelets with larger numbers of vanishing moments are used. This wavelet-based representation enables us to formulate a multiscale tomographic reconstruction technique in which the object is reconstructed at multiple scales or resolutions. A complete reconstruction is obtained by combining the reconstructions at different scales. Our multiscale reconstruction technique has the same computational complexity as the FBP reconstruction method. It differs from other multiscale reconstruction techniques in that 1) the object is defined through a one—dimensional multiscale transformation of the projection domain, and 2) we explicitly account for noise in the projection data by calculating maximum *a posteriori* probability (MAP) multiscale reconstruction estimates based on a chosen fractal prior on the multiscale object coefficients. The computational complexity of this maximum *a posteriori* probability (MAP) solution is also the same as that of the FBP reconstruction. This result is in contrast to commonly used methods of statistical regularization, which result in computationally intensive optimization algorithms.**

## I. Introduction

In this work we present a multiresolution approach to the problem of reconstructing an image from a complete set[1] of tomographic projections. The conventional, and most commonly used, method for reconstruction from noiseless tomographic projections is the filtered-back-projection (FBP) reconstruction technique [2], wherein the projection data at each angle are first filtered by a high-pass, "ramp" filter and then back-projected. In this paper we work in a multiscale transform space, where the matrix representation of the corresponding multiscale filtering operator is nearly diagonal, leading to an efficient multiscale tomographic reconstruction technique. Perhaps more significantly, however, the different scale components of our proposed multiscale method induce a corresponding multiscale representation of the underlying object and, in particular, provide estimates of (and thus information about) the field or object at a variety of resolutions at *no additional cost.*

Noisy imaging problems arise in a variety of contexts (e.g., low-dose medical imaging, oceanography, and in several applications of nondestructive testing of materials) and in such cases techniques such as FBP often yield unacceptable results. Conventionally, reconstruction from noisy projection data is regularized by one of the following two techniques. First, the FBP ramp filter may be rolled off at high frequencies, thus attenuating high-frequency noise at the expense of not reconstructing the fine-scale features in the object [3], [4]. This roll-off results in a fast, though ad hoc, regularization method. The other common regularization method is to solve for a maximum *a posteriori* probability (MAP) object estimate based on a two—dimensional (spatial) Markov random field (MRF) prior model [5], [6]. This MAP estimate results in a statistically regularized reconstruction that allows the inclusion of prior knowledge in a systematic way, but leads to optimization problems that are extremely computationally intensive. As a result, these methods have not found wide favor in practical applications. In contrast to these methods, we extend our multiscale reconstruction technique to obtain a multiscale MAP object estimate which, while retaining the advantages of statistically-based approaches, is obtained with similar computational complexity to the FBP reconstruction. We accomplish this by constructing prior statistical models *directly in scale space.*

Wavelets have recently been applied to tomography by other researchers as well. Peyrin *et al.* [7] have shown how to relate certain 2-D wavelet transforms of fields to a corresponding set of angularly varying 1-D wavelet transforms of projections of the field. Sahiner and Yagle use the wavelet transform to perform spatially varying filtering by reducing the noise energy in the reconstructed image over regions where high-resolution features are not present [8]. They also apply wavelet based reconstruction to the limited angle tomography problem by assuming approximate a priori knowledge about the edges in the object that lie parallel to the missing views [9]. Wu [10] has recently proposed an image reconstruction technique in which prior information is introduced through the 2-D wavelet transform of the image field. The reconstruction is then obtained by iteratively solving for the corresponding MAP solution and has the advantage that sharp object edges as well as smoothness

Manuscript received December 23, 1993; revised September 25, 1995. This work was supported by the Air Force Office of Scientific Research under Grant F49620-92-J-0002, the Office of Naval Research under Grant N00014-91-J-1004, and the US Army Research Office under Grant DAAL03-92-G-0115. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was M. Viergever. *Asterik indicates corresponding author.*

M. Bhatia is with J. P. Morgan & Co. Inc.

*W. C. Karl is with the Department of Electrical, Computer, and Systems Engineering, Boston University, 44 Cummington Street, Boston, MA 02215 USA (e-mial: karl@mit.edu).

A. S. Willsky is with the Stochastic Systems Group, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Publisher Item Identifier S 0278-0062(96)01217-7.

[1] According to Llacer [1], "a complete data set could be described as sufficient number of line projections at a sufficient number of angular increments such that enough independent measurements are made to allow the image reconstruction of a complete bounded region."

within regions in the object are preserved in the reconstructed image. DeStefano and Olson [11] and Berenstein and Walnut [12] have also used wavelets for tomographic reconstruction problems, in particular to localize the radon transform in even dimensions and reduce the radiation exposure required to image a local region of the object.

## II. Preliminaries

### A. The Tomographic Reconstruction Problem

In tomography, the goal is to reconstruct an object or a field, $f$, from line-integral projection data [2]. For a parallel-beam imaging geometry, the projection data consist of parallel, nonoverlapping strip integrals through the object at various angles. Suppose we have $N_\theta$ uniformly-spaced angular positions between 0° and 180° and $N_s$ parallel strip integrals at each angular position. A discrete version of this situation at angle $k$ can be represented as

$$y_k = T_k f \tag{1}$$

where $T_k$ is an $N_s \times N_s^2$ matrix representing the projection operation at angle $k$, $f$ is an $N_s^2 \times 1$ vector representing $f(u, v)$ on an $N_s \times N_s$ square-pixel lattice, and $y_k$ is the corresponding vector of measurements at that angle. Thus row $\ell$ of $T_k$ is the (discrete) representation of the $\ell$th strip function at angle $k$ and the inner product of $f$ with this strip yields the data contained in the corresponding entry of $y_k$. The tomographic reconstruction problem then reduces to finding an estimate $\widehat{f}$ of the discretized object $f$ given the projection data contained in the $\{y_k;\ k = 1, \cdots, N_\theta\}$.

### B. The Filtered-Back-Projection Reconstruction Technique

The FBP reconstruction technique is based directly on the radon inversion formula, which is valid (i.e., yields exact reconstructions) only when a continuum of noise-free line-integral projections from all angles are used [2]. In practice, we only have access to sampled projection data which are collected using strips of finite width. In this work, we assume that we sample finely enough to produce good reconstructions in the noiseless case [13], [14]. In the FBP reconstruction, the estimated object is represented as a linear combination of the same functions along which the projection data are collected

$$\widehat{f} = \sum_{k=1}^{N_\theta} T_k^T x_k \tag{2}$$

where the $N_s$ vector $x_k$ contains the object coefficient set at angle $k$. Note that (2) can be interpreted as the back-projection operation [2].

To complete the reconstruction the coefficients $x_k$ must now be determined. The standard FBP method calculates them for each angle $k$ by filtering the projection data $y_k$ at that particular angle with a ramp filter $R$ [2]

$$x_k = R\, y_k. \tag{3}$$

Thus (2) and (3) together represent the two operations used in the standard FBP reconstruction.

### C. 1-D Wavelet Transform Based Multiscale Decomposition

Here we present a brief summary of the wavelet-based multiscale decomposition of 1-D functions and intentionally suppress many details. The interested reader is referred to any of the many papers devoted to this topic, e.g., [15]. Given a length $2^N$ 1-D signal represented in the vector $x$, let $x^{(m)}$ be the vector containing the wavelet-based approximation to the signal at scale $m$ (i.e., the scaling coefficients) and let $\xi^{(m)}$ be the corresponding vector of added detail necessary to proceed to the next finer-scale approximation (i.e., the wavelet coefficients at that scale). Further, let $\xi$ be the vector containing the wavelet coefficients at all scales $\{\xi^{(m)}\}$ together with the coarsest-level approximation $x^{(0)}$ of the signal $x$. Then, we will capture the overall operation which takes a discrete signal vector $x$ to its corresponding wavelet transform $\xi$ by the matrix $W$ as follows:

$$Wx = \begin{bmatrix} \xi^{(N-1)} \\ \hline \vdots \\ \hline \xi^{(0)} \\ \hline x^{(0)} \end{bmatrix} \triangleq \xi. \tag{4}$$

Since the 1-D transform is invertible and the wavelet basis functions are orthonormal, it follows that $W^{-1}$ exists and further that $W^{-1} = W^T$. The matrix $W$ depends on the underlying chosen wavelet. In our work in this paper, in addition to the Haar wavelet, we will use wavelets from an especially popular family of these functions due to Daubechies [16], the separate elements of which are denoted $D_n$. To deal with edge effects we cyclically wrap the interval [17], [15]. The methods we describe can be readily adapted to other approaches for dealing with edge effects as in [18] and the references contained therein.

Intermediate approximations $x^{(m)}$ of $x$ and their finest-scale representation may be obtained by using only part of the full wavelet-coefficient set during synthesis, effectively assuming the finer-scale detail components are zero. For convenience, in the discussion to follow we capture this partial zeroing operation in the matrix operator $A(m)$, that nulls the upper $N - m$ subvectors of the overall wavelet vector $\xi$ and thus retains only the information necessary to construct the approximation $x^{(m)}$ at scale $m$

$$A(m) \triangleq \text{block diag}\left[0_{(2^N - 2^m)},\ I_{(2^m)}\right] \tag{5}$$

where $0_p$ is a $p \times p$ matrix of zeros and $I_q$ is a $q \times q$ identity matrix. Also it will prove convenient to define a similar matrix operator $D(m)$, that retains only the information in $\xi$ pertaining to the *detail* component at scale $m$ by zeroing all but the sub-vector corresponding to $\xi^{(m)}$

$$D(m) \triangleq \text{block diag}\left[0_{(2^N - 2^{m+1})},\ I_{(2^m)},\ 0_{(2^m)}\right]. \tag{6}$$

Finally, with these definitions note that we have the following scale-recursive relationship for the partially zeroed vectors:

$$A^{(m+1)}\,\xi = A(m)\,\xi + D(m)\,\xi. \tag{7}$$

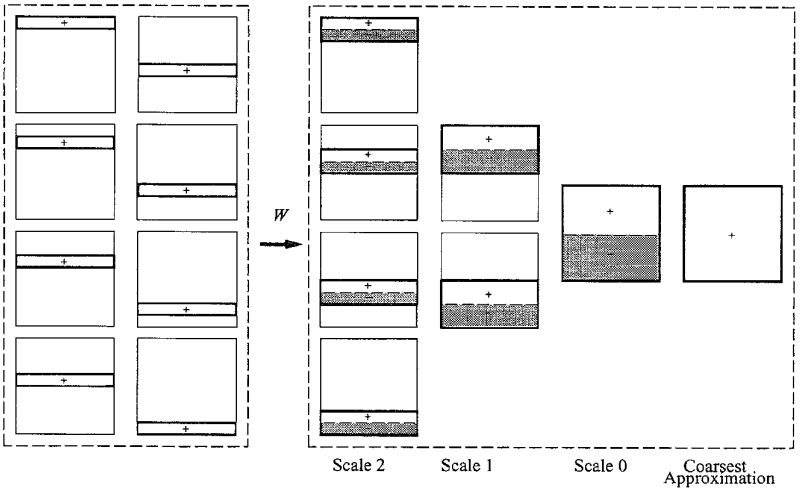



Fig. 1. Example of relationship between (a) original strip basis functions contained in $T_k$ and (b) transformed multiscale basis functions of $\mathcal{T}_k$ for a fixed angle $k$ corresponding to the Haar wavelet.

## III. The Multiscale Reconstruction Technique

### A. Multiscale Object Representation

In this section we derive our 1-D wavelet-based multiscale reconstruction technique. We start by applying a multiscale change of basis, as defined by the matrix $W$ in Section II-C, to the original set of object coefficients $x_k$ at each angle $k$ to obtain an equivalent set of multiscale object coefficients

$$\xi_k = W x_k. \tag{8}$$

Thus, the vector $\xi_k$ forms a multiresolution representation of $x_k$. More importantly, by reflecting this change of basis into the original FBP object representation (2), we naturally induce a corresponding *multiscale representation of the object* through a corresponding set of transformed multiscale basis functions. In particular, substituting (8) into (2) we obtain

$$\widehat{f} = \sum_{k=1}^{N_\theta} \left(T_k^T W^T\right)(W x_k) \triangleq \sum_{k=1}^{N_\theta} \mathcal{T}_k^T \xi_k. \tag{9}$$

The rows of the transformed matrix $\mathcal{T}_k = W T_k$ will now contain the multiscale object basis functions at angle $k$. The wavelet-transform operator matrix $W$, acting identically on each column of $T_k$, will thus form the new multiscale basis functions at that angle from linear combinations of the corresponding original strip functions, where these linear combinations correspond precisely to a *one-dimensional* wavelet transform perpendicular to the projection direction. This transformation of the basis functions is shown schematically in Fig. 1 (which corresponds to the case of the rectangular, Haar wavelet). We may naturally group the multiscale 2-D spatial basis elements into a hierarchy of scale-related components based on their support extent or spatial localization (or, equivalently, their relation to the coefficients in $\xi_k$), as shown in Fig. 1.

A multiresolution object decomposition can now be obtained through (9) by using a series of approximations to $x_k$ at successively finer scales, thereby inducing a series of corresponding approximate representations of the object. In particular, we define the $m$th scale approximation $\widehat{f}^{(m)}$ to $\widehat{f}$ through

$$\widehat{f}^{(m)} \triangleq \sum_{k=1}^{N_\theta} \mathcal{T}_k^T \left(A(m)\,\xi_k\right) \tag{10}$$

where $A(m)$ is defined in Section II-C. The approximation $\widehat{f}^{(m)}$ uses only the $m$ coarsest-scale components of the full vector $\xi_k$. Similarly, by $\Delta\widehat{f}^{(m)}$ we denote the additional detail required to go from the object approximation at scale $m$ to that at scale $(m+1)$, which is given by

$$\Delta\widehat{f}^{(m)} \triangleq \sum_{k=1}^{N_\theta} \mathcal{T}_k^T \left(D(m)\,\xi_k\right) \tag{11}$$

where $D(m)$ is also defined in Section II-C. Combining the object approximation and detail definitions (10) and (11) with the scale recursive relationship (7), we see that the *object itself* satisfies the following scale recursive relationship:

$$\widehat{f}^{(m+1)} = \widehat{f}^{(m)} + \Delta\widehat{f}^{(m)}. \tag{12}$$

Note that our 2-D multiscale object representation given in (10) and corresponding scale recursive construction (12) are induced naturally by the structure of the individual 1-D wavelet-based multiscale decompositions at each angle $k$ and is *not* simply a 2-D wavelet transform of the original object estimate $\widehat{f}$.

### B. Multiscale Coefficient Determination

We now have a natural multiscale object-representation framework through (10), (11), and (12) that is similar in spirit to the FBP case (2). To complete the process and

create multiscale object estimates from data, we must find the multiscale object coefficients $\xi_k$ (which contain all the information we need). To this end, we perform a wavelet-based multiscale change of basis to the data sequences $y_k$ to obtain an equivalent set of multiscale observations

$$\eta_k \triangleq W y_k \tag{13}$$

where, recall, $W$ is a matrix representing the wavelet transform. By combining the two transformations (8) and (13) together with the original FBP relation (3) we obtain

$$\xi_k = \mathcal{R}\,\eta_k \tag{14}$$

where $\mathcal{R} = WRW^T$ is the multiscale data filter, corresponding to the ramp filter $R$ of the usual FBP case. The operator $R$ is compressed by the wavelet operator so that $\mathcal{R}$ is nearly diagonal. Further, higher compression is achieved if Daubechies wavelets $D_n$ with larger $n$ are used, consistent with the observations of Beylkin *et al.* [19].

### C. The Overall Multiscale Algorithm

We are now in a position to present our overall multiscale reconstruction method, which parallels the identical and independent angular processing of the FBP approach, and is thus no more complex than this popular method

*Algorithm 1 (Multiscale Reconstruction):*

1) For a given choice of wavelet, form the multiscale filter matrix $\mathcal{R} = W\,R\,W^T$.
2) For each angle $k$ perform the following:
   a) Find the multiscale observations $\eta_k = W\,y_k$ by taking the 1-D wavelet transform of $y_k$.
   b) Find the multiscale object coefficient set $\xi_k = \mathcal{R}\,\eta_k$ by filtering these observations.
   c) Back-project $\xi_k$ along the corresponding multiscale basis functions $\mathcal{T}_k$, $\mathcal{T}_k^T \xi_k$.
3) Combine the back-projections at each angle to obtain the overall estimate, $\sum_k \mathcal{T}_k^T \xi_k$.

Beyond simply finding a finest-scale object estimate, as described in Algorithm 1, however, we may also reconstruct the underlying object at *multiple resolutions* through (10), (11), and (12), and thus easily obtain information about the object at multiple scales. In particular, if an approximation $\widehat{f}^{(m)}$ at scale $m$ is desired, then in Algorithm 1 we need only replace $\xi_k$ by $(A(m)\,\xi_k)$ in Step 2–c) and 3). Further, if instead we want to reconstruct the detail $\Delta\widehat{f}^{(m)}$ at a particular scale, we need only replace $\xi_k$ by $(D(m)\,\xi_k)$ in Step 2c and 3 of Algorithm 1. Such intermediate scale information about $\widehat{f}$ can even be efficiently found by calculating *only* those elements necessary for reconstructing the scale of interest—i.e., all of $\xi_k$ is not required.

*Examples:* We now show some examples of our multiscale reconstruction framework. Fig. 2 shows the 256 × 256 phantom used in the experiments of this section. Projection data were collected at 256 equally spaced angles ($N_\theta = 256$) with 256 strips per projection ($N_s = 256$) and the Daubechies $D_3$ wavelet was used for the multiscale decomposition $W$. Fig. 3 shows approximate object reconstructions $\widehat{f}^{(m)}$ for the entire range of scales $m = 1, \cdots, 8$. Fig.. 3(i) shows the FBP reconstruction for comparison (which is identical to the finest-scale approximation $\widehat{f}^{(8)}$). The intermediate-scale estimates demonstrate how information is focused at different scales. For example, in the scale 3 reconstruction $\widehat{f}^{(3)}$ [Fig. 3(c)], though only 8 of the full 256 coefficient elements in the vectors $\xi_k$ are being used, we can already distinguish separate objects.

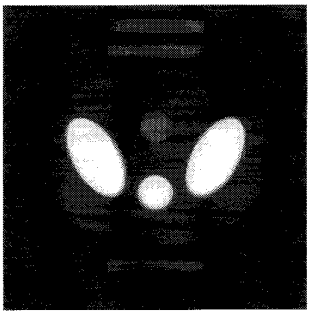


Fig. 2. Phantom used for reconstruction experiments. The phantom is 256 × 256 and projections are gathered at 256 equally spaced angles ($N_\theta = 256$) with 256 strips per angle ($N_s = 256$).

In Fig. 4 we show the corresponding detail components $\Delta\widehat{f}^{(m)}$ for the same phantom. Notice that the fine-scale, edge-based, features of the phantom are clearly visible in the $\Delta\widehat{f}^{(4)}$ and $\Delta\widehat{f}^{(5)}$ reconstructions [Fig. 3(e) and (f)], showing that structural information can be obtained from these detail images alone.

The wavelet-based multiscale transformation also serves to compress the ramp filter matrix $R$ so that the corresponding multiscale filter matrix $\mathcal{R}$ is nearly diagonal. One consequence is that a very good approximation to the exact reconstruction procedure of Algorithm 1 can be achieved by ignoring the off-diagonal terms of $\mathcal{R}$ in (14). Further, this approximation to the exact reconstruction becomes better as Daubechies wavelets $D_n$ with larger $n$ are used. As an illustration, in Fig. 5 we show complete (finest scale) reconstructions $\widehat{f}$ of the same phantom as before, based on the same projection data but using a diagonal approximation to $\mathcal{R}$ in (14) and Algorithm 1 for a variety of choices of the wavelet defining $W$.

## IV. Multiscale Regularized Reconstructions

The presence of noise in projection data often leads to reconstructions by standard methods, such as FBP, that are unacceptable and thus require some form of regularization. In contrast to the standard techniques, we will develop a multiscale MAP object estimate that, while retaining the advantages of statistically based approaches, is obtained with the same computational complexity as the FBP reconstruction. To accomplish this we continue to work in the projection domain and build our statistical models there, rather than in the original object domain.[2] To this end, we start with an observation equation relating the noisy data $y_k$ to the *FBP object coefficients* $x_k$, rather than the corresponding 2-D object

[2] We point out that our approach for obtaining regularized reconstructions by working directly in the projection domain is similar in spirit to the recent work of Fessler [20]. However, an important difference between the two approaches is that the latter is not multiscale based.

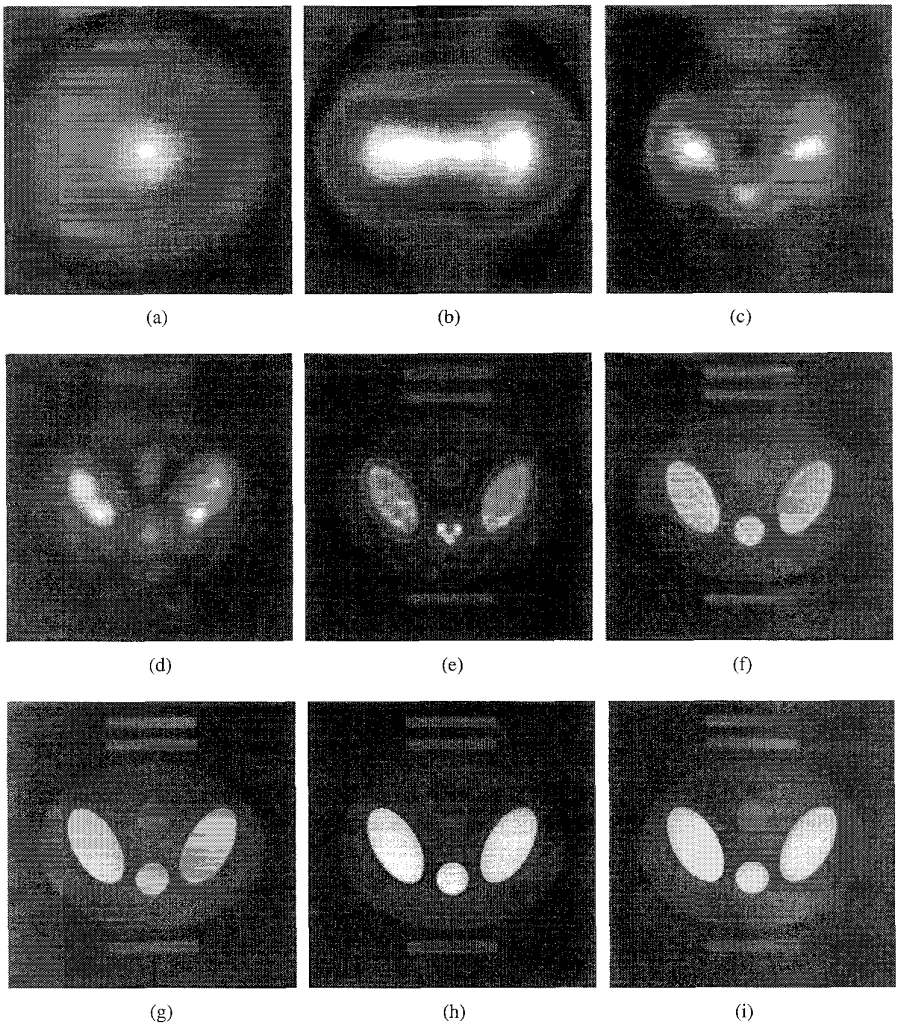



Fig. 3. Approximation reconstructions of phantom of Fig. 2 at various scales, using $D_3$ wavelet. (a) $\widehat{f}^{(1)}$. (b) $\widehat{f}^{(2)}$. (c) $\widehat{f}^{(3)}$. (d) $\widehat{f}^{(4)}$. (e) $\widehat{f}^{(5)}$. (f) $\widehat{f}^{(6)}$. (g) $\widehat{f}^{(7)}$. (h) $\widehat{f}^{(8)}$. (i) Shows the corresponding FBP reconstruction $\widehat{f}$ for comparison. The FBP reconstruction is the same as $\widehat{f}^{(8)}$, since it is the complete reconstruction.

$f$ as is usually done. Such a relationship may be found in the FBP relationship (3), which in the presence of noise in the data becomes

$$y_k = R^{-1}x_k + n_k, n_k \sim \mathcal{N}(0, \Lambda_{n_k}) \tag{15}$$

where $R$ is the FBP ramp-filter operator,[3] and the notation $z \sim \mathcal{N}(m, \Lambda)$ denotes a Gaussian distribution of mean $m$ and covariance $\Lambda$. We assume that $\Lambda_{n_k} = \lambda_k I_{N_s}$ (where $I_n$ denotes an $n \times n$ identity matrix), i.e., that the noise is uncorrelated from strip to strip but may have different noise covariances at different angles. Further, we assume that the noise is uncorrelated from angle to angle, so that $n_k$ is independent of $n_j$, $k \neq j$. This model of independent noise in the projection domain is well justified for most tomographic applications.

[3]Note (15) assumes that $R^{-1}$ exists. For the case where $R$ represents an ideal ramp filter this will indeed not be the case, as this operator nulls the dc component of a signal. For filters used in practice, however, this inverse does exist and the expression given in (15), based on such a filter is well defined [2].

For purposes of estimation we desire a relationship between multiscale representations of the data, object coefficients, and noise. Working in the multiscale-transform domain will again allow us to obtain induced multiresolution estimates of the object, in addition to parsimonious prior models. By combining (15) with the multiresolution orthogonal changes of bases (8) and (13) based on $W$ (defined in Section II-C) we obtain

$$\eta_k = \mathcal{R}^{-1}\xi_k + \nu_k, \nu_k \sim \mathcal{N}(0, \Lambda_{\nu_k}) \tag{16}$$

where $\nu_k = Wn_k$ is the multiscale-transformed noise vector at angle $k$ with $\Lambda_{\nu_k} = W\Lambda_{n_k}W^T = \lambda_k I_{N_s}$ as its corresponding covariance. Note that the assumption of uncorrelated noise from angle to angle and strip to strip in the original projection domain results in uncorrelated noise from angle to angle and multiscale strip to multiscale strip in the multiscale domain as well, since $W$ is an *orthonormal transformation.*

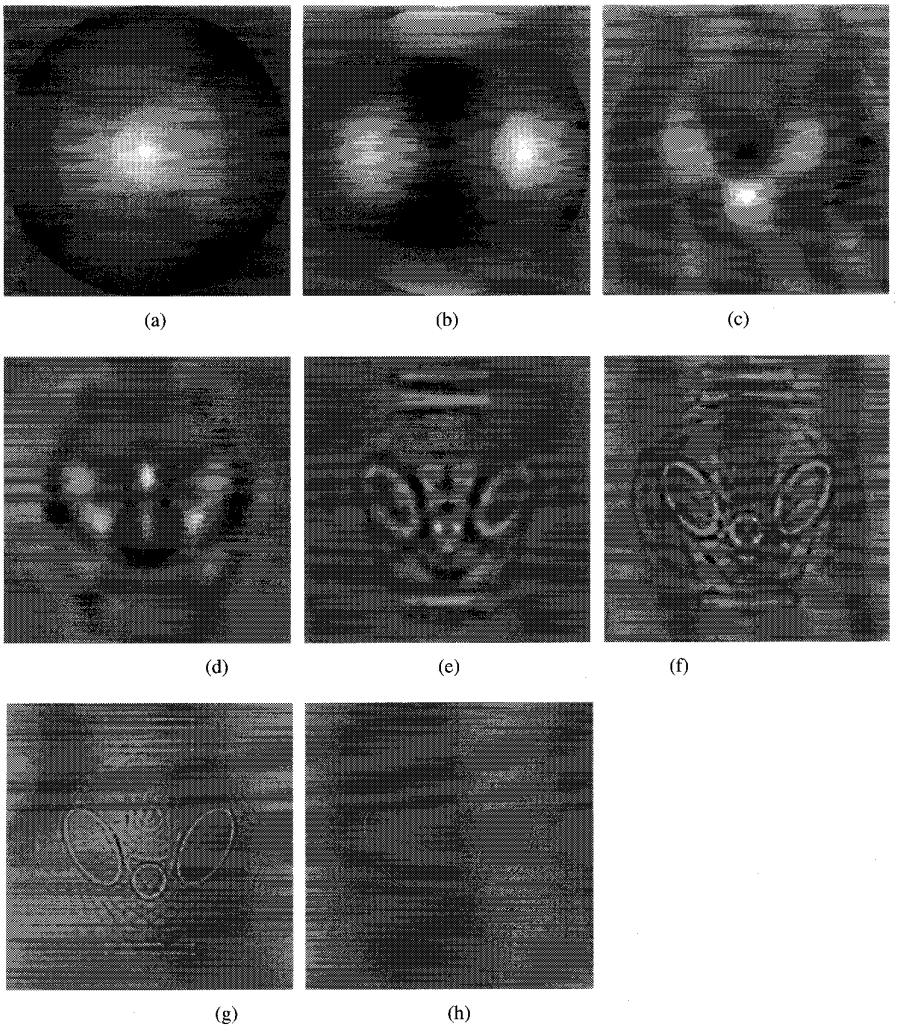


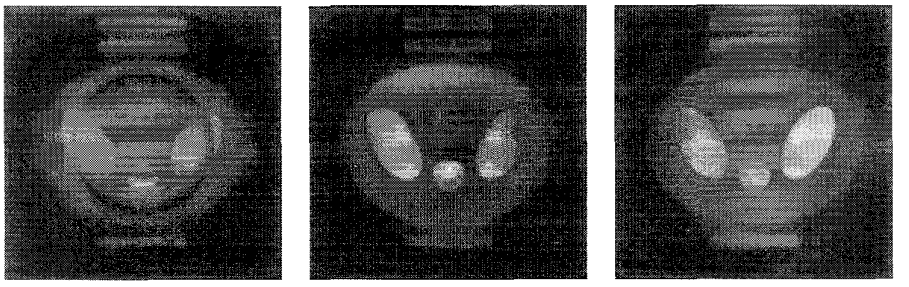

Fig. 4. The detail added between successive scales in the reconstructions of Fig. 3. (a) $\Delta\widehat{f}(0)$. (b) $\Delta\widehat{f}(1)$. (c) $\Delta\widehat{f}(2)$. (d) $\Delta\widehat{f}(3)$. (e) $\Delta\widehat{f}(4)$. (f) $\Delta\widehat{f}(5)$. (g) $\Delta\widehat{f}(6)$. (h) $\Delta\widehat{f}(7)$.

### *A. The Multiscale Prior Model*

To create a MAP estimate of the multiscale object coefficients $\xi_k$, we will combine the observation equation (16) with a prior statistical model for the desired unknown multiscale coefficient vectors $\xi_k$. We base our prior model of the object coefficients *directly* in scale-space. In particular, self-similar, fractal models can simply be obtained by choosing the detail components $\xi_k^{(m)}$ (i.e., the wavelet coefficients at each scale) as independent, $\mathcal{N}(0, \sigma^2 2^{-\rho m})$ random variables [21] whose variance decreases geometrically with scale. The parameter $\rho \geq 0$ determines the nature, i.e., the texture, of the resulting self-similar process while $\sigma^2$ controls the overall magnitude. In addition to defining the probabilistic structure of the detail components of $\xi_k$, we also need a probabilistic model for the DC or coarsest-scale element $x_k^{(0)}$ of $\xi_k$, of which we expect to have little prior knowledge. As a result we choose this element as $\mathcal{N}(0, \overline{\Lambda}_\xi)$, where the (scalar) uncertainty $\overline{\Lambda}_\xi$ is chosen sufficiently large to prevent a bias in our estimate of the average behavior of the field. Our overall prior model is then given by $\xi_k \sim \mathcal{N}(0, \Lambda_\xi)$ with $\xi_k$ independent from angle to angle, and where the *diagonal* matrix $\Lambda_\xi$ is given by

$$\begin{aligned} \Lambda_\xi &= \text{block diag}\left[\Lambda_\xi^{(N-1)}, \cdots, \Lambda_\xi^{(1)}, \Lambda_\xi^{(0)}, \overline{\Lambda}_\xi\right] \\ \Lambda_\xi^{(m)} &= \sigma^2 2^{-\rho m} I_{2^m}. \end{aligned} \tag{17}$$

Such scale-space-based, self-similar models are commonly and effectively used in many application areas such as modeling of natural phenomenon and textures, biological signals, geophysical and economic time series, etc. [21]–[25]. Obviously other choices may be made for the statistics for the multiscale object coefficients. The choice we have made in (17), while simple, is well adapted to many naturally occurring phenomenon. Since the *observation noise* power is uniform across scales or frequencies, the geometrically decreasing variance of this prior model implies that the projection data will most strongly influence the reconstruction of coarse-scale features and the prior model will most strongly influence the reconstruction of fine-scale features.



(a) (b) (c)

Fig. 5. Complete finest-scale multiscale reconstructions for phantom of Fig. 2 for different approximate filtering operators. The three frames show approximate multiscale reconstructions using only the diagonal elements of $\mathcal{R}$corresponding to different choices of the underlying wavelet: (a) Haar, (b) $D_3$, and (c) $D_8$. Note that the approximation to the exact reconstruction becomes better as Daubechies wavelets $D_n$ with larger $n$ are used.

### B. The Multiscale MAP Estimate

We are now in a position to present our overall algorithm for computing a MAP [26] multiscale object estimate $\widehat{\xi}_k$. Since the data at each angle $\eta_k$ and the corresponding prior model for $\xi_k$ are independent from angle to angle, the MAP estimates of the vectors $\xi_k$ decouple. In particular, the estimate of $\xi_k$ at each angle, based on the observations (16) and the prior model (17) is given by

$$\widehat{\xi}_k = \left[\Lambda_\xi^{-1} + \mathcal{R}^{-T}\Lambda_{\nu_k}^{-1}\mathcal{R}^{-1}\right]^{-1}\mathcal{R}^{-T}\Lambda_{\nu_k}^{-1}\,\eta_k \triangleq \overline{\mathcal{R}}\,\eta_k \quad (18)$$

where the regularized multiscale filter operator $\overline{\mathcal{R}}$ is defined in the obvious way. This regularized filtering matrix is exactly analogous to the unregularized filtering operator $\mathcal{R}$ of (14).

Finally, as in the noiseless case, the resulting *object estimate* $\widehat{f}$ is then obtained by back-projecting the MAP estimated multiscale object coefficients $\widehat{\xi}_k$ along the corresponding multiscale basis functions $\mathcal{T}_k$ and combining the result. The overall structure of this regularized reconstruction parallels that of the original FBP method, and therefore is of the *same computational complexity as FBP*. In summary, our overall, efficient regularized multiscale estimation algorithm is given by the following procedure:

*Algorithm 2 (Regularized Multiscale Reconstruction):*

1) For a given choice of wavelet, find the regularized multiscale filter matrix $\overline{\mathcal{R}}$ as follows:
   a) Form the unregularized multiscale filter matrix $\mathcal{R} = W\,R\,W^T$.
   b) Choose the model parameters $\lambda_k$ specifying the observation noise processes $\Lambda_{\nu_k}$ c.f. (16).
   c) Choose the multiscale prior model parameters $\sigma^2$, $\rho$, and $\overline{\Lambda}_\xi$ specifying the magnitude and texture of the model and the uncertainty in its average value, respectively, and generate the prior covariance matrix $\Lambda_\xi$ through (17).
   d) Form $\overline{\mathcal{R}} = \left[\Lambda_\xi^{-1} + \mathcal{R}^{-T}\Lambda_{\nu_k}^{-1}\mathcal{R}^{-1}\right]^{-1}\mathcal{R}^{-T}\Lambda_{\nu_k}^{-1}$.
2) For each angle $k$ perform the following:
   a) Find the multiscale observations $\eta_k = W\,y_k$ by taking the 1-D wavelet transform of $y_k$.
   b) Find the regularized multiscale object coefficient set $\widehat{\xi}_k = \overline{\mathcal{R}}\,\eta_k$ by filtering these observations.
   c) Back-project $\widehat{\xi}_k$ along the corresponding multiscale basis functions $\mathcal{T}_k$, $\mathcal{T}_k^T\widehat{\xi}_k$.
3) Combine the back-projections at each angle to obtain the overall regularized estimate, $\sum_k \mathcal{T}_k^T\widehat{\xi}_k$.

As before, we may also easily obtain regularized reconstructions of the object at multiple resolutions by using (10) and (11) together with the MAP coefficient estimates $\widehat{\xi}_k$. In particular, to obtain the approximation $\widehat{f}^{(m)}$ at scale $m$, we need only replace $\widehat{\xi}_k$ by $(A(m)\,\widehat{\xi}_k)$ in Step 2–c) and 3). Similarly, the corresponding object detail components $\Delta\widehat{f}^{(m)}$ at scale $m$ may be obtained by using $(D(m)\,\widehat{\xi}_k)$ in place of $\widehat{\xi}_k$ in these steps.

While Algorithm 2 is already extremely efficient, additional gains may be obtained by exploiting the ability of the wavelet-transform operator $W$ to compress the FBP filtering operator $R$. In particular, let us assume that the wavelet transform $W$ truly diagonalizes $R$ by effectively ignoring the small, off-diagonal elements in $\mathcal{R}^{-1}$ so that[4]

$$\mathcal{R}^{-1} \approx \mathrm{diag}(r_1, r_2, \cdots, r_{N_s}) \quad (19)$$

where $r_i$ are the diagonal elements of $\mathcal{R}^{-1}$. Now let us represent the diagonal prior model covariance matrix as $\Lambda_\xi = \mathrm{diag}[p_1, p_2, \cdots, p_{N_s}]$, and recall that $\Lambda_{\nu_k} = \lambda_k I_{N_s}$. Using these quantities together with our approximation to $\mathcal{R}^{-1}$ in the specification of the estimate (18) yields an approximate expression for $\widehat{\xi}_k$

$$\widehat{\xi}_k \approx \mathrm{diag}\left(\frac{r_1}{r_1^2 + (\lambda_k/p_1)}, \frac{r_2}{r_2^2 + (\lambda_k/p_2)}, \cdots, \frac{r_{N_s}}{r_{N_s}^2 + (\lambda_k/p_{N_s})}\right)\eta_k \triangleq \widetilde{\mathcal{R}}\,\eta_k \quad (20)$$

where the approximate MAP filtering matrix $\widetilde{\mathcal{R}}$ is defined in the obvious way. Our experience is that when $W$ is defined using Daubechies wavelets of order 3 or higher (i.e., using $D_3$, $D_4$,...), the estimates obtained using $\widetilde{\mathcal{R}}$ in place of the exact regularized filter $\overline{\mathcal{R}}$ in Algorithm 2 are visually indistinguishable from the exact estimates, where $\mathcal{R}^{-1}$ is not assumed to be diagonal. Indeed, it is actually this approximate

[4] One can imagine another level of approximation in which we set the off-diagonal elements of $\mathcal{R}$ itself to zero *prior* to inversion rather than those of $\mathcal{R}^{-1}$. This further approximation results in reconstructions that are visually very similar to what we obtain here.

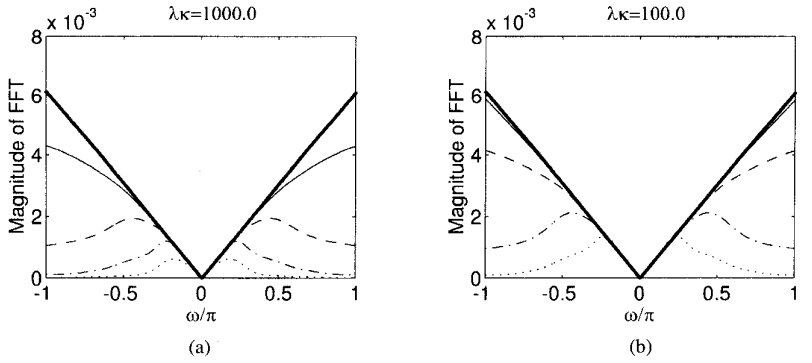



Fig. 6. The Fourier transform of the central row of $R_{\text{eff}}$ for different values of regularization parameters $\rho$ (the decay rate of the added scale detail) and $\lambda_k$ (the noise variance). (a) $\lambda_k = 1000.0$, (b) $\lambda_k = 100.0$. In each of the plots, the V-shaped heavy line corresponds to the standard FBP ramp filter and the four curves from top to bottom correspond to $\rho = 0.5$ (solid line), 1.0 (dashed line), 1.5 (dash-dot line), and 2.0 (dotted line), respectively. In all cases we fixed $\sigma^2 = 1$ (the overall prior model amplitude) and $\overline{\Lambda}_\xi = 1$ (the prior model DC variance).

filtering operator $\widetilde{\mathcal{R}}$ that we use to generate the example reconstructions we show next.

Before proceeding, however, let us examine our MAP regularized filtering operator $\overline{\mathcal{R}}$ in more detail to understand how our multiscale MAP estimation procedure relates both to the standard FBP method and the regularization obtained through apodization of the FBP filter. The multiscale MAP estimation operation specified by (18) imposes a corresponding relationship between the original finest-scale quantities $\widehat{x}_k$ and $y_k$, given by

$$\widehat{x}_k = \left(W^T \overline{\mathcal{R}} W\right) y_k \triangleq R_{\text{eff}}\, y_k \tag{21}$$

where the effective multiscale MAP regularized filtering matrix $R_{\text{eff}}$ is defined in the obvious way. The behavior of the matrix operator $R_{\text{eff}}$ can be most easily understood by examining its corresponding frequency-domain characteristics. To this end, in Fig. 6 we plot the magnitude of the Fourier transform of the central row of the effective regularized matrix $R_{\text{eff}}$ corresponding to a variety of choices of the model or regularization parameters. We also plot, with heavy lines, the magnitude of the Fourier transform of the corresponding central row of the standard, unregularized FBP ramp filter matrix $R$ for comparison. From Fig. 6, we can see that in the multiscale MAP framework regularization is basically achieved by rolling off the ramp filter at high frequencies, the same principle as used in apodization regularized FBP reconstructions. The multiscale-based algorithm, however, provides a rational way of achieving this roll-off as a function of the relative strengths of the observation noise and our belief in the prior model.

*Examples:* Next we show some examples of reconstructions using our multiscale methods in the presence of noise. The same $256 \times 256$ phantom shown in Fig. 2 was used for all experiments. In each case projection data for the phantom were again generated at $N_\theta = 256$ equally spaced angles with $N_s = 256$ strips in each projection. These noise-free values were then corrupted through the addition of independent, zero-mean Gaussian noise to yield our observations. The variance $\lambda_n$ of this additive noise depended on the experiment and was chosen to yield an equivalent signal-to-noise ratio (SNR) of the resulting observations, defined as

$$\text{SNR (dB)} = 10 \log \frac{\sum_{k=1}^{N_\theta} \|T_k f\|^2}{\lambda_n N_\theta N_s} \tag{22}$$

where, recall, $T_k f$ is the noise-free projection data at angle $k$. Finally, in all multiscale reconstructions we show here, the Daubechies $D_3$ wavelet was used in the definition of $W$ for the reconstruction.

Fig. 7 shows the approximate object reconstructions $\widehat{f}^{(m)}$, at various scales $m$, corresponding to our multiscale MAP estimate of $\widehat{\xi}_k$ using noisy data with SNR = 5 dB. The MAP estimate $\widehat{\xi}_k$ was generated using the approximate expression (20), which, for the Daubechies $D_3$ wavelet we are using, was indistinguishable from the corresponding estimate based on the exact expression (18). Again the approximations become finer from left to right and top to bottom in the figure. All images are displayed with the same, common scaling. The effect of the regularization can be readily seen in its ability to suppress noise in the finest-scale reconstruction. For comparison, one can see that the object is completely lost in the unregularized reconstruction for this case [Fig. 7(i)] due to the extreme level of noise. Finally, the multiscale nature of the information focusing can be seen in the scale evolution of the estimates. In particular, there appears to be little difference between scale 5 and finer-scale estimates in the figure, suggesting that little additional information is being obtained in proceeding to such finer scales so that we should stop the reconstruction at this coarser scale.

Finally, in Fig. 8, we show a series of finest-scale multiscale MAP regularized reconstructions, corresponding to different choices of the prior model texture as determined by the parameter $\rho$. The same phantom as before is used, but we use observations with a SNR of $-10$ dB (much worse than used earlier). The unregularized reconstruction [Fig 8(d)] is shown for comparison. All images are displayed with the same common scaling. The object is completely lost in the unregularized reconstruction at this extreme level of noise. The MAP reconstructions are shown in Fig. 8(a)-(c), with a smoother, more correlated prior model being used as we proceed from left to right.

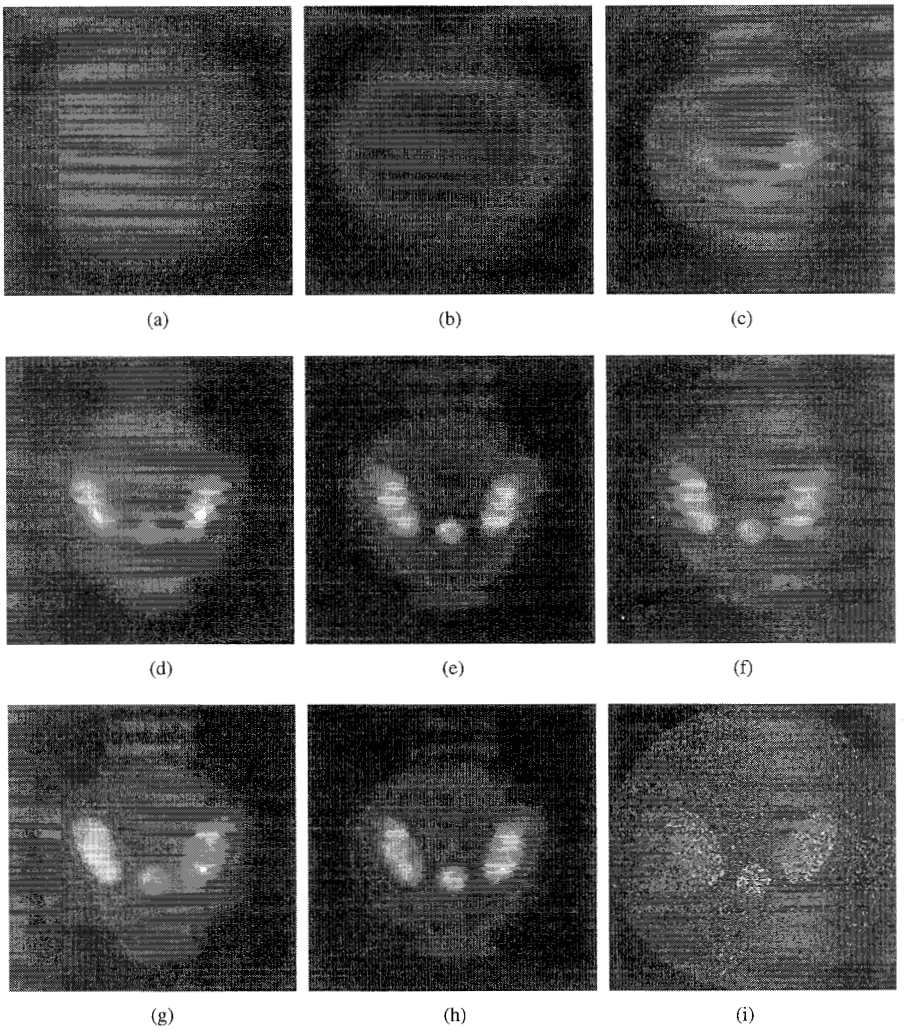



Fig. 7. Multiscale MAP regularized reconstructions at various scales of phantom of Fig. 2 from 5 dB SNR projection data using $D_3$ wavelet. The values of the statistical model parameters used were: observation noise variance $\lambda_k = 5.5 \times 10^5$, added detail decay rate $\rho = 1.5$, overall prior model magnitude $\sigma^2 = 11$, prior model DC variance $\overline{\Lambda}_\xi = 1$. (a) $\hat{f}^{(1)}$. (b) $\hat{f}^{(2)}$. (c) $\hat{f}^{(3)}$. (d) $\hat{f}^{(4)}$. (e) $\hat{f}^{(5)}$. (f) $\hat{f}^{(6)}$. (g) $\hat{f}^{(7)}$. (h) $\hat{f}^{(8)}$. For comparison, the unregularized reconstruction for this case is given in (i).

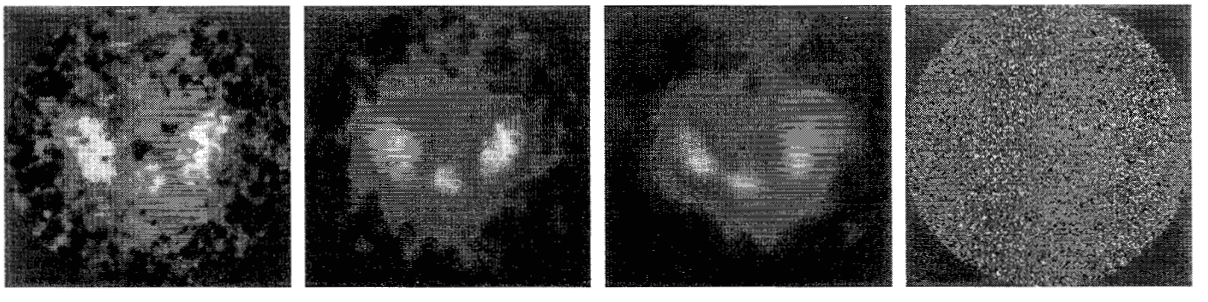



Fig. 8. Multiscale MAP regularized reconstructions of the phantom of Fig. 2 at the finest scale from $-10$ dB SNR observations for different choices of prior model texture, $\rho$, with observation noise variance $\lambda_k = 1.7 \times 10^7$, overall prior model magnitude $\sigma^2 = 17$, and prior model DC variance $\overline{\Lambda}_\xi = 1$, are shown in the first three frames: (a) $\rho = 0.5$. (b) $\rho = 1.0$. (c) $\rho = 1.5$. For comparison the unregularized reconstruction is shown in (d). The RMS variation of the image values is (from left to right): 9.5, 7.7, 7.3, 140.

## V. Conclusion

In this paper we have developed a wavelet-based multiscale tomographic reconstruction technique that is different from other multiscale techniques in the following respects. First, our 2-D multiscale object representation is naturally induced by expanding the FBP coefficients, and hence basis functions (i.e., strips), in a *1-D wavelet basis*. This approach is in contrast to other multiscale reconstruction techniques that *begin* with

a 2-D object representation obtained from a full 2-D wavelet decomposition of the object space. In contrast, the multiscale representation resulting from our approach, arising as it does from the projection strips themselves, is much closer to the measurement domain. This results in a highly efficient method to compute our multiscale object coefficients, in particular, no more complex than the widely used standard FBP operation. Yet, unlike the FBP method, our multiscale reconstructions also provide a framework for the extraction and presentation of information at multiple resolutions from data. Further, our resulting multiscale relationships between data and object allow extremely simple *approximations* to be made to our exact relationships with *virtually no loss in resulting image quality*.

In addition, based on this wavelet-based multiscale framework, we have presented a statistically based multiresolution MAP estimation algorithm. This method provides statistically regularized reconstructions from noisy data and does so at multiple resolutions, with effort similar to that required for the standard FBP method. This approach, based on the construction of prior models directly in scale-space, allows for the inclusion of natural, self-similar prior models into the reconstruction process. In contrast, conventional statistically based regularization methods, utilizing MRF-type prior models constructed directly in (finest scale) object space, lead to extremely complex and taxing optimization problems. The result has typically been that such statistically-motivated methods have been largely shunned in practice in favor of fast, though ad hoc, approaches. Our results provide a bridge between these two extremes. Further, in providing estimates at multiple resolutions, our results provide tools for the assessment of the resolution-versus-accuracy tradeoff, wherein we expect coarser-scale features of data to be more accurately determined than finer-scale ones. Though we did not exploit this ability in the present paper, our formulation also allowed the possibility of combining data from projections of fundamentally different quality, through the specification of different noise variances $\lambda_k$ at different angles. The resulting estimates do not correspond to a simple FBP or even rolled-off FBP reconstruction, yet are easily obtained in our framework. Finally, as before, our multiscale MAP approach leads to algorithms that are amenable to an additional level of approximation, with resulting improved efficiency, again at virtually no loss in corresponding reconstruction quality.

## REFERENCES


[1] J. Llacer, "Theory of imaging with a very limited number of projections," *IEEE Trans. Nucl. Sci.*, vol. NS-26, no. 1, pp. 596–602, Feb. 1979.

[2] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Piscataway, NJ: IEEE Press, 1988.

[3] H. H. Barrett, "Image reconstruction and the solution of inverse problems in medical imaging," in *Medical Images: Formation, Handling and Evaluation, Proceedings of the NATO Advanced Study Institute on the Formation, Handling and Evaluation of Medical Images, Portugal, 1988*. New York: Springer–Verlag, 1992, pp. 21–22.

[4] Donner "Algorithms for Reconstruction Tomography", RECLBL Library Users Manual, Lawrence Berkeley Laboratory, Univ. California, 1977, pp. 35–42.

[5] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified em algorithm," *IEEE Trans. Med. Imag.*, vol. 9, no. 1, pp. 84–93, 1990.

[6] K. Sauer and C. Bouman, "Bayesian estimation of transmission tomograms using segmentation based optimization," submitted to *IEEE Trans. Nucl. Sci.*

[7] F. Peyrin, M. Zaim, and R. Goutte, "Multiscale reconstruction of tomographic images," in *Proc. IEEE-SP Int. Symp. Time-Freq. Time-Scale Anal.*, Oct. 1992, pp. 219–222.

[8] B. Sahiner, and A. E. Yagle, "Image reconstruction from projections under wavelet constraints," *IEEE Trans. Signal Processing*, Dec. 1993.

[9] ——, "Limited angle tomography using the wavelet transform," preprint, Oct. 1993.

[10] Z. Wu, "MAP image reconstruction using wavelet decomposition," *Lec. Notes Comput. Sci.*, vol. 687, pp. 354–371, 1993.

[11] J. DeStefano and T. Olson, "Wavelet localization of the radon transform in even dimensions," in *Proc. IEEE-SP Int. Symp. Time-Frequency and Time-Scale Anal.*, Oct. 1992, pp. 137–140.

[12] C. Berenstein, and D. Walnut, "Local Inversion of the Radon Transform in Even Dimensions Using Wavelets," Center For The Applications Of Mathematics, George Mason Univ., Reference: CAM-21/93, Jan. 1993.

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

[14] A. M. Cormac, "Sampling the radon transform with beams of finite width," *Phys. Med. Biol.*, vol. 23, no. 6, pp. 1141–1148, 1978.

[15] S. G. Mallat, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, 1989.

[16] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992, pp. 167–213.

[17] N. H. Getz, "A perfectly invertible, fast, and complete wavelet transform for finite length sequences: The discrete periodic wavelet transform," in *Proc. SPIE Annu. Conf. Math. Imag.: Wavelet Applicat. Signal and Image Processing*, San Diego, CA, July 1993.

[18] I. Daubechies, "Wavelet on the interval," in *Progress in Wavelet Analysis and Applications*, Y. Meyer and S. Roques, Eds. Santa Rosa, CA: Frontiers, 1992, pp. 95–107.

[19] G. Beylkin, R. Coifman, and V. Rokhlin, "Fast wavelet transforms and numerical algorithms I," *Commun. Pure and Appl. Mathe.*, vol. XLIV, pp. 141–183, 1991.

[20] J. A. Fessler, "Tomographic reconstruction using information-weighted spline smoothing," *Lec. Notes Comput. Sci.*, vol. 687, pp. 372–386, 1993.

[21] G. W. Wornell, and A. V. Oppenheim, "Estimation of fractal signals from noisy measurements using wavelets," *IEEE Trans. Signal Processing*, vol. 40, no. 3, pp. 611–623, Mar. 1992.

[22] E. B. Cargill, H. H. Barrett, R. D. Fiete, M. Ker, D. D. Patton, and G. W. Seeley, "Fractal physiology and nuclear medicine scans," *SPIE Med. Imag. II*, vol. 9, no. 14, pp. 355–361, 1988.

[23] B. J. West and A. L. Goldberger, "Physiology in fractal dimensions," *Amer. Scientist*, vol. 75, pp. 354–365, 1987.

[24] C. -C. Chen, J. S. Daponte, and M. D. Fox, "Fractal texture analysis and classification in medical imaging," *IEEE Trans. Med. Imag.*, vol. 8, no. 2, pp. 133–142, June 1989.

[25] M. Luettgen, W. C. Karl, and A. S. Willsky, "Efficient multiscale regularization with applications to the computation of optical flow," *IEEE Trans. Image Processing*, vol. 3, no. 1, pp. 41–64, Jan. 1994.

[26] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968, pp. 54–63.