Lawrence Berkeley National Laboratory

Recent Work

Title
Generalized Approach to Inverse Problems in Tomography: Image Reconstruction for Spatially Variant Systems Using Natural Pixels

Permalink
https://escholarship.org/uc/item/5132d67h

Authors
Baker, J.R.
Budinger, T.F.
Huesman, R.H.

Publication Date
1992-10-14
Presented at the High Performance Computing and Biomedical Research Conference, Durham, NC, October 14–15, 1992, and to be published in the Proceedings

Generalized Approach to Inverse Problems in Tomography: Image Reconstruction for Spatially Variant Systems Using Natural Pixels

J.R. Baker, T.F. Budinger, and R.H. Huesman

October 1992

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. Neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California and shall not be used for advertising or product endorsement purposes.

Lawrence Berkeley Laboratory is an equal opportunity employer.
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.
Generalized Approach to Inverse Problems in Tomography: Image Reconstruction for Spatially Variant Systems Using Natural Pixels

John R. Baker†+, Thomas F. Budinger†+, and Ronald H. Huesman‡
† Department of Electrical Engineering and Computer Sciences
University of California, Berkeley

and

‡ Research Medicine and Radiation Biophysics Division
Lawrence Berkeley Laboratory
University of California
1 Cyclotron Road
Berkeley, CA 94720

This work was supported in part by the Director, Office of Energy Research, Office of Health and Environmental Research, Medical Applications and Biophysical Research Division of the U.S. Department of Energy under contract No DE-AC03-SF00098 and in part by NIH grant HL07367.
Abstract

A major limitation in tomographic inverse problems is inadequate computation speed, which frequently impedes the application of engineering ideas and principles in medical science more than in the physical and engineering sciences. Medical problems are computationally taxing because a minimum description of the system often involves 5 dimensions (3 space, 1 energy, 1 time), with the range of each space coordinate requiring up to 512 samples. The computational tasks for this problem can be simply expressed by posing the problem as one in which the tomograph system response function is spatially invariant, and the noise is additive and Gaussian. Under these assumptions, a number of reconstruction methods have been implemented with generally satisfactory results for general medical imaging purposes. However, if the system response function of the tomograph is assumed more realistically to be spatially variant and the noise to be Poisson, the computational problem becomes much more difficult. Some of the algorithms being studied to compensate for position dependent resolution and statistical fluctuations in the data acquisition process, when expressed in canonical form, are not practical for clinical applications because the number of computations necessary exceeds the capabilities of high performance computer systems currently available. Reconstruction methods based on natural pixels, specifically orthonormal natural pixels, preserve symmetries in the data acquisition process. Fast implementations of orthonormal natural pixel algorithms can achieve orders of magnitude speedup relative to general implementations. Thus, specialized thought in algorithm development can lead to more significant increases in performance than can be achieved through hardware improvements alone.
1 Background

The tomographic projection formation can be described by the discrete-continuous model [1], [2], [3], [4]

\[ p_{\theta k} = F_{\theta k} \cdot b \]  
\[ = \int_{\mathbb{R}} dy \int_{\mathbb{R}} dx f_{\theta k}(x, y) b(x, y) \]  

where \( p_{\theta k} \) is the measured projection at angle index \( \theta \) and bin position \( k \). \( F_{\theta k} \) is a second order tensor functional operating on the two dimensional object distribution \( b \). The operation of equation 1 represents the integration of the product of the impulse response \( f_{\theta k}(x, y) \) and the object distribution \( b(x, y) \) over the imaging field as depicted in figure 1. The symbol \( \cdot \) indicates integration over the imaging field. The impulse response function is the spatial response of a projection at angle \( \theta \) and bin \( k \) to a point source moved to every position within the sampling domain. There are \( \Theta \) different angles and \( K \) projection bins at each angle.

![Figure 1: Schematic of projection formation.](image)

To simplify notation\(^1\), the projection formation equation is written in vector form

\[ \mathbf{p} = F \cdot \mathbf{b} \]  

\(^1\)Lower case bold symbols denote vectors, lower case script symbols denote functions or scalars, upper case bold symbols denote matrices, and upper case script symbols denote operators.
by combining the $\theta$ and $k$ indices into one index. Specifically[5],

\begin{align}
    b & \in L_2[\mathbb{R}^2], \\
    p & \in \mathbb{R}^{\Theta K}, \text{ and} \\
    F & : L_2[\mathbb{R}^2] \rightarrow \mathbb{R}^{\Theta K}.
\end{align}

Because the model is based on the fact that the detection process is defined on a discrete domain and the original distribution is defined on a continuous domain, the model is easily adapted to include a variety of physical effects found in many imaging modalities. For positron emission tomography (PET), $F_{\theta k}$ can include radioactive decay, positron range, non-collinearity of photons, sampling geometry, attenuation, inter-detector-crystal scatter, detector crystal penetration, and detection efficiency [6] [7] [8].

The singular value decomposition of the tomograph system response function is

\[ F = US \cdot V^T. \] (7)

$U$ is an orthogonal matrix containing the left singular vectors of $F$ and is defined by the eigenvalue decomposition of the projection normal matrix,

\[ A \equiv F \cdot F^T = US \cdot S^TU^T \] (8) (9)

where $F^T$ is the adjoint operator of $F$ [5]. An element $A_{\theta k'\theta k}$ is the projection at angle $\theta'$ and bin $k'$ of the backprojection at angle $\theta$ and bin $k$ of a unit projection value, $p_{\theta k} = 1$. A schematic of this operation is shown in figure 2. The projection normal matrix is symmetric and positive semidefinite. $V$ are the right singular functions of $F$ and are defined by the relationship

\[ F^T F = V \cdot S^T S \cdot V^T. \] (10)

The functional $S$ operates similarly to $F$ in that it maps continuous domain functions to discrete domain samples. Thus, $V$ performs an infinite dimensional rotation on the continuous domain object space, $S$ selects and scales a finite number of the rotated functions, and $U$ performs a finite dimensional rotation into the discrete domain projection measurement space as shown in equations 11-13.

\begin{align}
    U & : \mathbb{R}^{\Theta K} \rightarrow \mathbb{R}^{\Theta K}, \\
    S & : L_2[\mathbb{R}^2] \rightarrow \mathbb{R}^{\Theta K}, \text{ and} \\
    V & : L_2[\mathbb{R}^2] \rightarrow L_2[\mathbb{R}^2].
\end{align}

A geometric interpretation is given in figure 3.
Figure 2: Schematic of the computation of one element of the projection normal matrix. A unit projection is backprojected and reprojected to a new projection to form a projection matrix element.
The goal of tomography is to reconstruct the unknown distribution, $b$, from one realization, $p$, of the projection measurement process $p$ [9] [10] [11] [12] [13] [14] [15] [16]. Reconstructing the true continuous space distribution, $b$, from sampled projections is probably impossible without prior information about the distribution [3] [17] [18] [19]. Instead, a discretized representation, $c_{mn}$, is estimated from the measurements where $B_{mn}(x, y)$ defines a generalized pixel.

$$b(x, y) \approx \sum_{mn} B_{mn}^T(x, y)c_{mn}$$  \hspace{1cm} (14)

In vector form the pixelization is

$$b \approx B^Tc.$$  \hspace{1cm} (15)

Using this representation, a least squares estimator (LSE) [5] [20] [21] [22] [23] is formulated to estimate the mean intensity of the generalized pixels, $\hat{c}$, used to describe the unknown spatial distribution from one measured projection dataset, $p$. While the resulting LSE formula is quite general, specific application to image reconstruction using square pixels, Buonocore’s natural pixels [24] [25], and the new orthonormal natural pixels is shown. The least squares estimator for the mean intensity of the generalized pixel image is found by minimizing the square of the $L_2$ norm of the difference between the projection vector, $p$, and the estimated projection, $F \cdot B^Tc$, over all possible image vectors, $c$; i.e.,
Generalized Approach to Inverse Problems in Tomography

\[ \hat{c} \equiv \arg \min_{\hat{c}} \left\{ \| \mathbf{p} - F \cdot B^T \hat{c} \|_2^2 \right\} \]  \hspace{1cm} (16)

\[ = \left( S \cdot V^T \cdot B^T \right)^+ U^T \mathbf{p}. \]  \hspace{1cm} (17)

The operator \(^+\) is the Moore-Penrose pseudo-inverse [26] [27] [28]. Using the measured projection vector, \( \mathbf{p} \), as a single sample estimate of the mean projection vector gives

\[ \hat{c} = \left( S \cdot V^T \cdot B^T \right)^+ U^T \mathbf{p}. \]  \hspace{1cm} (18)

The fluctuations of the generalized pixel least squares estimator due to random variations in the measurements can typically be characterized in terms of the covariance between pixel estimates. The covariance matrix for the generalized pixel estimator is

\[ \Sigma_{\hat{c}} \equiv E \left[ (\hat{c} - E\hat{c})(\hat{c} - E\hat{c})^T \right] \]  \hspace{1cm} (19)

\[ = \left( S \cdot V^T \cdot B^T \right)^+ U^T \Sigma_{\mathbf{p}} U \left( B \cdot V \cdot S^T \right)^+. \]  \hspace{1cm} (20)

where \( \Sigma_{\mathbf{p}} \) is the covariance of the projections.

The estimator for the mean of the intensity of the object in continuous space is found by applying the adjoint of the basis operator, \( B \), to the generalized pixel estimator of the intensity mean. For the continuous space object, the least squares estimator for the mean of the intensity is

\[ \hat{b} = B^T \hat{c} \]  \hspace{1cm} (21)

\[ = B^T \left( S \cdot V^T \cdot B^T \right)^+ U^T \mathbf{p}. \]  \hspace{1cm} (22)

The basis set used to describe the pixels influences the types of artifacts that appear in the reconstructed image [25] [29] [30]. In this work, three bases are evaluated. The first, \( B_1 \), is the traditional square pixel or Heaviside basis. The second basis set, \( B_2 \), consists of the set of functions that comprise the tomograph system response functional and was proposed by Buonocore [24]. The third, \( B_3 \), has been proposed by us and is composed of the right singular functions \( V \), defined by equation 7, that have been selected by \( S \) and normalized by the \( L_2 \) norm.

\[ B_1 \equiv \text{Heaviside} \]  \hspace{1cm} (23)

\[ B_2 \equiv F \]  \hspace{1cm} (24)

\[ B_3 \equiv \left[ (S \cdot S^T)^+ \right]^\frac{1}{2} S \cdot V^T \]  \hspace{1cm} (25)

Small singular values in the singular value filter, \( \left[ (S \cdot S^T)^+ \right]\frac{1}{2} \), of equation 18 can lead to large statistical errors in the reconstructed image, \( \hat{b} \). By applying a diagonal weighting
matrix, $D$, to the filter, the mean square error of the object estimates may be decreased [3]. The resulting estimates and covariance are

$$
\hat{c} = D \left[ (S \cdot S^T)^+ \right]^{\frac{1}{2}} U^T \bar{P} \\
\Sigma_{\hat{c}} = D \left[ (S \cdot S^T)^+ \right]^{\frac{1}{2}} U^T \Sigma_p U \left[ (S \cdot S^T)^+ \right]^{\frac{1}{2}} D \\
b = V \cdot S^T D (S \cdot S^T)^+ U^T \bar{P}.
$$

Determining the weighting values is the subject of ongoing research. Since basis vectors are assumed to be arranged so the singular values are in the accepted non-increasing order, one possibility is to truncate the number of singular values used in the singular value filter so only the $J$ largest singular values will be included. A weighting matrix with elements

$$
D_{j',j} = \begin{cases} 
1 & \text{if } j' = j \text{ and } j < J, \\
0 & \text{otherwise}
\end{cases}
$$

will select only the $J$ largest singular values. Since the basis is orthonormal, the resulting object estimate is the sum of the estimates of each pixel that was multiplied by one; i.e.,

$$
\hat{b} = \sum_{j=0}^{J-1} B^T j \hat{c}_j \\
= \sum_{j=0}^{J-1} V \cdot S^T j (S \cdot S^T)^+ U^T \bar{P}.
$$

## 2 Example

As a simple example, consider the spatial sampling system of figure 4 that has three projection angles and two projection bins at each angle. The angles are equally spaced between zero and $\pi$ radians. The impulse response functions are defined by

$$
f_{\theta k}(x,y) = \begin{cases} 
1 & \text{if } k - 1 \leq -x \sin \left( \frac{\pi}{3} \right) + y \cos \left( \frac{\pi}{3} \right) < k \text{ and } x^2 + y^2 \leq 1, \\
0 & \text{otherwise}
\end{cases}
$$

and are shown in figure 4. The projection normal matrix and the singular value decomposition of the projection normal matrix for this example are
where the left singular vectors, $U$, are given as the product of a matrix and the inverse of the norm of that matrix. There are four non-zero singular values for this example parallel beam sampling system meaning that of the six measurements only four are linearly independent. The linear dependence between the measurements is simply demonstrated by subtracting one of the impulse response functions at one angle from all of the impulse response functions at the other angle, the resulting difference is equal to the impulse response function that was not included in the difference; e.g., for the system of figure 4,

$$f_{00}(\cdot, \cdot) + f_{01}(\cdot, \cdot) - f_{10}(\cdot, \cdot) = f_{11}(\cdot, \cdot).$$

Figures 5-7 show, respectively, a possible set of basis functions for square pixels, Buonocore's natural pixels, and orthonormal natural pixels using the sampling defined by the impulse response functions of figure 4.

Consider the wedge shaped object of figure 8 which is defined as

$$b(x, y) = \begin{cases} 
1 & \text{if } |\frac{y}{x}| < \frac{1}{2} \text{ and } x \geq 0 \text{ and } x^2 + y^2 \leq 1, \\
0 & \text{otherwise.}
\end{cases}$$

(36)
Figure 4: Schematic of spatial sampling functions for a simple parallel beam tomographic system with three equally spaced projection angles and two projection bins at each angle.
Figure 5: Square pixel or Heaviside basis for a simple parallel beam tomographic system with three equally spaced projection angles and two projection bins at each angle.
Figure 6: Buonocore's natural pixel basis for a simple parallel beam tomographic system with three equally spaced projection angles and two projection bins at each angle.
Figure 7: New orthonormal pixel basis for a simple parallel beam tomographic system with three equally spaced projection angles and two projection bins at each angle.
The pixel coefficients using the square pixel basis of figure 5 are

\[ c = \left[ \frac{1}{3} \ 0 \ 0 \ rac{1}{3} \right]^T. \]  \hfill (37)

In Buonocore's natural pixel basis of figure 6, the pixel coefficients are

\[ c = \frac{1}{72} \left[ 4 \ 4 \ 13 \ -5 \ 13 \ -5 \right]^T \]  \hfill (38)

and for the orthonormal natural pixel basis shown in figure 7

\[ c = \frac{\sqrt{\pi}}{4} \left[ 1 \ \frac{1}{\sqrt{6}} \ -\frac{1}{\sqrt{2}} \ 0 \ 0 \ 0 \right]^T. \]  \hfill (39)

Figure 8: Wedge shaped phantom object.

Figure 9 shows, respectively, the continuous space representation of the wedge shaped object using square pixels, Buonocore's natural pixels, and orthonormal natural pixels with the sampling defined by the impulse response functions of figure 4.

All systematic reprojection errors due to pixelization are eliminated using a basis like \( B_2 \) or \( B_3 \) as shown in table 1 for the wedge phantom. While this is true in particular for \( B_2 \) and \( B_3 \), any basis that spans the subspace of functions defined by \( S \cdot V^T \) will also have this property. A geometric interpretation of the error associated with computing projections of pixelized object is shown in figure 10. Many of these bases may represent the original object distribution, \( b \), better than \( B_2 \) or \( B_3 \). However, no information about the coefficients for the functions that are outside the space \( S \cdot V^T \) is available from the projection measurements. By using \textit{a priori} information about the continuous space distribution of \( b \), the formulation of Bayesian estimators that use basis functions not in \( S \cdot V^T \) is an exciting area for future
Figure 9: Representation of a wedge shaped object using top) square pixels, middle) Buonocore’s natural pixels, and bottom) orthonormal natural pixels.
Table 1: Errors due to pixelization for a wedge shaped phantom sampled with a parallel beam tomographic system with three equally spaced projection angles and two projection bins at each angle.

<table>
<thead>
<tr>
<th>Basis</th>
<th>Object Estimate Squared Error $|b - B^T c|_2^2$</th>
<th>Projection Estimate Squared Error $|p - F \cdot B^T c|_2^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B1$</td>
<td>$\frac{4}{3}$</td>
<td>$\frac{\pi^2}{324}$</td>
</tr>
<tr>
<td>$B2$</td>
<td>$\frac{7}{6}$</td>
<td>$0$</td>
</tr>
<tr>
<td>$B3$</td>
<td>$\frac{7}{6}$</td>
<td>$0$</td>
</tr>
</tbody>
</table>

Figure 10: Simplified geometric representation of the systematic error that results from computing model projections from pixelized versions of an object. Pixel bases that include the subspace $S \cdot V^T$ eliminate systematic pixelization error.
research. Thus, the basis subset contained in $S \cdot V^T$ is from a channel model for $F$ and the subset contained in $V^T - S \cdot V^T$ is from a process model for $b$.

To test the LSE in the presence of noise, a noisy projection dataset,

$$p = \begin{bmatrix} 0.8579 & 0.2920 & 1.4252 & 0.0000 & 0.4829 & 0.0000 \end{bmatrix}^T$$ (40)

will be used. This projection dataset was created by sampling an independent multivariate normal (Gaussian) distribution with mean and variance equal to the noiseless projections of the wedge shaped object. The estimated pixel coefficients and the covariance of those estimates using the square pixel basis of figure 5 are

$$\hat{c} = \begin{bmatrix} 0.5508 & -0.3602 & 0.0981 & 1.0092 \end{bmatrix}^T$$ (41)

$$\Sigma_{\hat{c}} = \begin{bmatrix} 0.1738 & 0.0014 & -0.0854 & 0.0870 \\ 0.0014 & 0.0677 & -0.0191 & -0.0854 \\ -0.0854 & -0.0191 & 0.0677 & 0.0014 \\ 0.0870 & -0.0854 & 0.0014 & 0.1738 \end{bmatrix}$$ (42)

In Buonocore's natural pixel basis of figure 6, the pixel coefficient estimates and covariance are

$$\hat{c} = \begin{bmatrix} 0.1534 & 0.0629 & 0.5383 & -0.3220 & 0.1336 & 0.0827 \end{bmatrix}^T$$ (43)

$$\Sigma_{\hat{c}} = \begin{bmatrix} 0.1810 & -0.1771 & -0.1428 & 0.1468 & 0.1556 & -0.1517 \\ -0.1771 & 0.1810 & 0.1556 & -0.1517 & -0.1428 & 0.1468 \\ -0.1428 & 0.1556 & 0.1899 & -0.1771 & -0.1384 & 0.1512 \\ 0.1468 & -0.1517 & -0.1771 & 0.1722 & 0.1512 & -0.1561 \\ 0.1556 & -0.1428 & -0.1384 & 0.1512 & 0.1899 & -0.1771 \\ -0.1517 & 0.1468 & 0.1512 & -0.1561 & -0.1771 & 0.1722 \end{bmatrix}$$ (44)

and for the orthonormal natural pixel basis shown in figure 7

$$\hat{c} = \begin{bmatrix} 0.5751 & 0.6879 & -0.3641 & -0.2124 & 0.0000 & 0.0000 \end{bmatrix}^T$$ (45)

$$\Sigma_{\hat{c}} = \begin{bmatrix} 0.0556 & 0.0340 & -0.0589 & 0.0000 & 0.0000 & 0.0000 \\ 0.0340 & 0.1250 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ -0.0589 & 0.0000 & 0.1250 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.5000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}$$ (46)

Figure 11 shows, respectively, the continuous space reconstruction of the wedge shaped object using square pixels, Buonocore's natural pixels, and orthonormal natural pixels with the sampling defined by the impulse response functions of figure 4. Table 2 shows the mean and observed object squared error and the the observed projection squared error for these
reconstructions. The natural pixel and orthonormal natural pixel estimates have better observed squared error than the square pixel estimates. The square pixel estimator has better mean squared error characteristics than the unfiltered estimators based on natural pixels or orthonormal natural pixels for the wedge shaped object.

Figure 11: Least squares reconstruction of a wedge shaped object using top) square pixels, middle) Buonocore’s natural pixels, and bottom) orthonormal natural pixels.

Table 3 shows the effects of using the diagonal weighting matrix defined in equation 29 with the value $J$ varied from one to four for the example of figure 8. The projection estimate error decreases with the inclusion of each orthonormal natural pixel basis function; however, adding the image corresponding to the third basis vector, $J = 4$, increases the mean object squared error. The increase in mean object squared error is due to noise being added to the reconstructed image while no new information about the object is being added since the wedge phantom only has non-zero projections onto the first three orthonormal natural basis
Table 2: Least squares reconstruction errors for a wedge shaped phantom sampled with noise by a parallel beam tomographic system with three equally spaced projection angles and two projection bins at each angle.

<table>
<thead>
<tr>
<th>Basis</th>
<th>Mean Object Squared Error $E | b - B^T \hat{c} |_2^2$</th>
<th>Observed Object Squared Error $E | b - B^T \hat{c} |_2^2$</th>
<th>Observed Projection Squared Error $E | p - F \cdot B^T \hat{c} |_2^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B1$</td>
<td>1.1775</td>
<td>3.2636</td>
<td>0.3001</td>
</tr>
<tr>
<td>$B2$</td>
<td>1.3571</td>
<td>2.6294</td>
<td>0.2348</td>
</tr>
<tr>
<td>$B3$</td>
<td>1.3571</td>
<td>2.6294</td>
<td>0.2348</td>
</tr>
</tbody>
</table>

Table 3: Effects of truncating the number of singular values included in the singular value filter on orthonormal natural pixel least squares reconstruction errors for a wedge shaped phantom sampled with noise by a parallel beam tomographic system with three equally spaced projection angles and two projection bins at each angle.

<table>
<thead>
<tr>
<th>Truncation Index $J$</th>
<th>Mean Object Squared Error $E | b - B^T \hat{c} |_2^2$</th>
<th>Observed Object Squared Error $E | b - B^T \hat{c} |_2^2$</th>
<th>Observed Projection Squared Error $E | p - F \cdot B^T \hat{c} |_2^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.4762</td>
<td>1.9655</td>
<td>1.5270</td>
</tr>
<tr>
<td>2</td>
<td>1.3839</td>
<td>2.8223</td>
<td>0.5360</td>
</tr>
<tr>
<td>3</td>
<td>0.8571</td>
<td>2.4571</td>
<td>0.2584</td>
</tr>
<tr>
<td>4</td>
<td>1.3571</td>
<td>2.6294</td>
<td>0.2348</td>
</tr>
</tbody>
</table>

3 Implementation

The computational complexity of reconstruction algorithms can be reduced by using the orthonormal natural pixel basis because it preserves symmetries in the data acquisition process. In this section, an implementation of the orthonormal natural pixel least squares estimator for rotationally invariant systems is described and analyzed.
3.1 Orthonormal Least Squares

The continuous space representation of the truncated minimum $L_2$ norm least squares estimator for the mean value of the orthonormal natural pixel image is from equation 31

$$\hat{b} = \sum_{j=0}^{J-1} V \cdot S_j^T (S \cdot S^T)^+ U^T \mathcal{P}. \quad (47)$$

This estimate of the object is converted to a square pixel representation using the Heaviside basis operator, $B_1$, of equation 23. The resulting square pixel representation is

$$\hat{c} = \sum_{j=0}^{J-1} B_1 \cdot V \cdot S_j^T (S \cdot S^T)^+ U^T \mathcal{P} \quad (48)$$

and after rearrangement using equation 7,

$$\hat{c} = \sum_{j=0}^{J-1} B_1 \cdot F^T U (S \cdot S^T)^+ U^T \mathcal{P}. \quad (49)$$

A block diagram of the orthonormal natural pixel least squares estimation algorithm is shown in figure 12. In general, the computation of the projection normal matrix requires $O(2^2 K^2)$ integral evaluations. An $O(3^3 K^3)$ singular value or eigenvalue decomposition routine is used to compute $U$ and $S \cdot S^T$. After the singular value decomposition of the projection normal matrix is known, evaluation of equation 49 uses $O(2^2 K^2)$ operations.

For a rotationally invariant system, the evaluation of the projection normal matrix requires only $O(K^2)$ integral evaluations and its singular value decomposition requires $O(K^3)$ operations using the block circulant singular value decomposition algorithm described in section 3.2. The rotationally invariant orthonormal least squares estimator is computationally tractable on current computer systems.

---

2Let $n, n_0 \in \mathbb{N}$ and $\epsilon \in \mathbb{R}, \epsilon > 0$. Also, $f, g : \mathbb{N} \to \mathbb{R}$. Then, define

3Using the symbol $\Theta$ for the number of projection angles measured and the combined bound function $\Theta(\cdot)$ is somewhat confusing but parenthesis distinguish between the two uses.
Figure 12: Block diagram of orthonormal natural pixel least squares estimation algorithm.
3.2 Block Circulant Singular Value Decomposition (BCSVD)

When the elements of the projection normal matrix are a function of only the difference \( \Delta \theta = \theta' - \theta \) modulo \( \Theta \) as shown in equation 50, the system is rotationally invariant. If it is not a function of the difference between \( k \) and \( k' \) modulo \( K \), then the system is radially variant. The system is spatially variant if it is rotationally or radially variant.

\[
A_{\theta'k'\theta k} = A_{[(\theta - \theta') \mod \Theta] k' \theta k} = A_{[\Delta \theta \mod \Theta] k' \theta k} = \int_{\mathbb{R}} dy \int_{\mathbb{R}} dx f_{[\Delta \theta \mod \Theta] k} (x, y) f_{\theta k} (x, y)
\]

When the system is rotationally invariant, the projection normal matrix can be written in block circulant form [32]. The block circulant structure is

\[
A = \begin{bmatrix}
A_0 & A_1 & A_2 & \cdots & A_{\Theta-2} & A_{\Theta-1} \\
A_{\Theta-1} & A_0 & A_1 & \cdots & A_{\Theta-3} & A_{\Theta-2} \\
A_{\Theta-2} & A_{\Theta-1} & A_0 & \cdots & A_{\Theta-4} & A_{\Theta-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
A_2 & A_3 & A_4 & \cdots & A_0 & A_1 \\
A_1 & A_2 & A_3 & \cdots & A_{\Theta-1} & A_0 
\end{bmatrix}
\]

There are \( \Theta \times \Theta \) blocks each of size \( K \times K \).

For the simple system of figure 4 the projection normal matrix can be written in block circulant form by reversing the direction of the projection axis for projection angle \( \theta = 1 \). With this change,

\[
A = \frac{\pi}{6} \begin{bmatrix}
3 & 0 & 1 & 2 & 1 & 2 \\
0 & 3 & 2 & 1 & 2 & 1 \\
1 & 2 & 3 & 0 & 1 & 2 \\
2 & 1 & 0 & 3 & 2 & 1 \\
1 & 2 & 1 & 2 & 3 & 0 \\
2 & 1 & 2 & 1 & 0 & 3 
\end{bmatrix}
\]

There are three by three blocks each of size two by two.

By using the rotational invariance properties of the projection normal matrix, computationally fast and efficient algorithms have been implemented for the reconstruction procedures described in section 1. As an example for the rest of this section, the case where \( \Theta = K = 64 \) shall be used because it is representative of some typical tomographic inverse problems.

A \( \Theta (K^2 \Theta \log \Theta) \) fast Fourier transform (FFT) technique [33] [34] and an \( \Theta (\Theta K^3) \) singular value decomposition (SVD) algorithm are used to compute the factorization [35].
Generalized Approach to Inverse Problems in Tomography

\[ A = (F_\Theta \otimes I_K)^\dagger D (F_\Theta \otimes I_K) \]
\[ = (F_\Theta \otimes I_K)^\dagger U_D \Lambda_D U_D^T (F_\Theta \otimes I_K) \]
\[ = U\Lambda_D U^T \]
\[ = U S \cdot S^T U^T \]

where \( F_\Theta \) is a normalized \( \Theta \times \Theta \) discrete Fourier operator matrix and \( I_K \) is \( K \times K \) identity matrix. The operator \( ^\dagger \) is conjugate transpose and \( \otimes \) is the outer product operation.

Each of the \( K^2 \) discrete Fourier transforms of equation 56 can be computed independently; i.e., each sum does not need the result or input of another sum [36] [37]. The SVD of the blocks of \( D \) also do not have input/output dependencies with other blocks and can be computed without explicit synchronization. Therefore, parallel processing implementations of the block circulant singular value decomposition algorithm are possible.

Two parallel versions [38] of the BCSVD algorithm were implemented and tested on a Cray-2 supercomputer using macrotasks [35]. The \( \Theta(\Theta \log \Theta) \) grain size of FFT tasks is extremely small. For the example, it takes about 0.45 ms [39] [40]. This is comparable to the 0.31 ms necessary to synchronize with a server process and is much smaller than the 2.63 ms necessary to create a new process. It is thus advantageous to increase the grain size of FFT tasks by computing \( K \) FFTs per task. The resulting granularity of \( \Theta(\Theta \log \Theta) \) is about 29 ms. The task granularity of an SVD process is \( \Omega(K^3) \) which is 428 ms for the example problem.

A prescheduled algorithm was implemented by creating one process for each of the \( K \) FFT tasks and another process for each of the \( \Theta \) SVD tasks. The parent task starts \( n \) processes with either an FFT or an SVD task. All of the \( n \) processes run to completion before another \( n \) processes are started. This method is very easy to implement because all synchronization is implicit in the fork and join like paradigm [41].

To overcome the process creation overhead, a self-scheduling algorithm was constructed [38]. This method is more complex than the prescheduled algorithm but has a smaller time overhead. It requires explicit synchronization between server processes and a task manager. \( n \) server processes are created and each waits for a start signal after initial setup of local state information. After receiving the start signal from the task manager, a server checks what part of the matrix it is to work on next. When finished the server sends a ready signal to the hibernating manager. The manager then reassigns each of the server processes until the task queue is empty.

Figure 13 shows the computation time for different sizes of input matrices. The speedup of the algorithm, shown in figure 14, increases as the size of \( \Theta \) and \( K \) are increased. The prescheduled algorithm is faster for very small matrix sizes because the self-scheduled algorithm server processes have a larger startup overhead than a process started by the prescheduled algorithm. The self-scheduled algorithm is faster for medium sized problems that have small grain sizes but the prescheduled algorithm again approaches the speedup of self-scheduling as the problem size increases.

\(^4\)The matrix \( A \) of equation 53 is less general than the algorithm will accept since it is symmetric with square blocks and real elements instead of having rectangular blocks and complex elements.
Figure 13: Computation time versus problem size with four tasks and four processors available to service tasks.
Figure 14: Speedup versus problem size with four tasks and four processors available to service tasks.
The efficiency, shown in figure 15, does not approach unity as quickly as expected. This might be attributed to the timesharing scheduling algorithm used by the CTSS operating system and not to synchronization overhead because the overhead, shown in table 4, is less than 1.0% for Θ and K larger than 64 [42] [43]. It was not possible to verify this conjecture by using the machine without other users present.

![Multitasking efficiency versus Matrix size](image)

**Figure 15:** Efficiency versus problem size with four tasks and four processors available to service tasks.

The process creation time was found to be 2.63 ms. Task synchronization in the self-scheduling algorithm was 0.31 ms. A typical procedure call was measured to take 4.7 μs. Self-scheduling has less time overhead than prescheduling but is still 66 times more expensive than a procedure invocation.

Data memory usage and overhead is shown in table 5. Very little memory is necessary
for the synchronization of tasks. Each of the processes needs some local working storage for computing FFTs and SVDs. Code memory usage and overhead is shown in Table 6. The code space sharing was small due to a problem in the Fortran compiler that made code replication necessary.

Dynamic memory allocation costs are basically independent of the block size being allocated for small blocks. The cost depends almost entirely on the number of blocks being allocated. Each block takes approximately 0.68 ms to allocate. The server processes of the self-scheduling algorithm avoid this overhead by reusing their local storage during each activation. The prescheduling algorithm originally allocated local storage blocks within each child process. This was deemed to be unsatisfactory and another parameter with working storage was passed to each child to avoid the overhead of dynamic memory allocation.
The BCSVD algorithm provides orders of magnitude speedup by utilizing the circulant structure of matrices. A further speedup was obtained using macrotasking. This does not reduce central processing unit charges because time on all processors is billed to the job [43]. However, a substantial savings in memory charges is achieved because the program memory residency time is reduced by the multiprocessor speedup [44] [45]. For typical problems $\Theta$ and $K$ are approximately 256. This requires approximately 800 megabytes of memory which can be quite costly to use.

Self-scheduling is useful when the task granularity is small. As the task granularity increases, prescheduling overhead becomes less important. Prescheduling is much easier to implement and debug. There are no explicit synchronizations to consider since the operating system handles the process allocation and scheduling. The parent only has to wait for the operating system to signal that the child has finished. Self-scheduling needs explicit synchronization with the server tasks and is therefore more difficult to implement and debug.

The Fortran compiler does not allocate local variables on the stack properly. It puts some local variables into static storage. Thus, code sharing is not possible for the Fortran subroutines. Each process must have a separate copy of the code and local data space. This was done by creating copies of the subroutines and giving each copy a unique name space by appending the process number to the name of the subroutine and all of its descendants.

4 Conclusions

The inverse problem requires the solution of large systems of linear and non-linear equations. For example, the Donner 600-Crystal Positron Tomograph takes 120,200 projection measurements and the resulting linear system is $120,200 \times 120,200$. The computational complexity of configuration space methods based on orthonormal natural pixels has led to the use of a distributed computing environment in which workstations are used to analyze results from our identification, estimation, and optimization algorithms running on supercomputers in a multitasking environment with priority scheduling. Several of the algorithms have been implemented using large grain parallel processing and also remote procedure calls.

The block circulant singular value decomposition (BCSVD) algorithm uses discrete Fourier transforms to rotate the blocks of a block circulant matrix into block diagonal form. Each block on the diagonal is then factored using a general singular value decomposition (SVD) algorithm. The BCSVD algorithm provides orders of magnitude speedup over general SVD algorithms. For a $64^2 \times 64^2$ block circulant matrix, computation time decreased from 12 hours to 23 seconds on a Cray-2 (approximately 1 hour on a SPARCstation II) . Because the BCSVD algorithm is easy to partition, a further speedup can be achieved using parallel processing. The orthogonality properties of multidimensional fast Fourier transforms (FFT) allows the FFT portion of the algorithm to partition into macrotasks. The SVD of the blocks of the block diagonal matrix can be computed independently and a macrotask can be assigned to each SVD. A multiprocessor speedup of 3.06 was achieved for prescheduling and for self-scheduling a multiprocessor speedup of 3.25 was observed using four processors on a Cray-2. Relative time overhead was 0.5% for the prescheduled algorithm and 0.07% for the self-scheduled algorithm. Relative memory overhead was 4% for both cases. Self-scheduling is useful when the task granularity is small. As the task granularity increases, preschedul-
ing overhead becomes less important. The prescheduled algorithm is satisfactory for most emission tomography problems because all the dimensions of the matrices are greater than 64 and the task granularity will therefore be large when compared to the synchronization overhead.

Multitasking the block circulant singular value decomposition algorithm decreases overall computation costs by reducing the time large sections of memory are in use. Little or no gain comes from reduced central processing unit charges since processing time on all processors is charged to a job.

The data acquisition model presented in this work provides a mathematical framework to incorporate detailed knowledge about the response function of a tomography system and the statistical properties of the signals acquired using that system. Using this mathematical forward problem model, it is easy to represent systems of varying complexity; e.g., simple spatially invariant systems, systems with spatially variant response that have symmetries, and the most general linear case, a spatially variant system without symmetry. The novel singular value decomposition of the projection formation operator used in the data acquisition model is a powerful mathematical description of a tomography system and is fundamental to the estimation (inverse problem) methods presented here. While the results presented in the examples are not inconsequential, these techniques should find the most utility in modeling the spatial sampling of each unique tomograph. Extensions of the model to include sampling in three spatial dimensions as well as time should be straightforward; but, in practice, higher dimensional applications may be limited by computational tractability without using special computing technology.

Because the data acquisition process is represented as a linear map from a continuous domain object space to a discrete domain observation space, it is a more physically realistic model of many systems than approximations using continuous-continuous maps or discrete-discrete maps. Thus, the validity of many results that were obtained using these approximations, e.g., angular and lateral sampling density in emission tomography, may need to be reexamined using the new, more robust techniques presented in this work. While the verification of old results is worthwhile, it is the unanswered questions such as the efficacy of iterative algorithms and stopping rules, the formulation of Bayesian estimators that use basis functions in the null space of the projection formation operator, the representation of object functions that are convex cones, and the efficiency of algorithm implementations that provide challenging new research opportunities for the application of the mathematical methods presented here to characterize and solve inverse problems in tomography.

The computational complexity of reconstruction algorithms can be reduced by using orthonormal natural pixels because this basis preserves symmetries in the data acquisition process. The present state of the art in computing hardware can be expected to accomplish a properly composed orthonormal natural pixel least squares reconstruction of data taken with a modern rotationally invariant tomography system, such as the Donner 600-Crystal Positron Tomograph, in 13 minutes per tomograph slice. While substantial performance enhancements are often realized using improved hardware, specialized thought in algorithm development can lead to more significant performance increases allowing the solution of previously intractable problems like those found in tomography.
Generalized Approach to Inverse Problems in Tomography

References


Generalized Approach to Inverse Problems in Tomography


