A Sinogram Restoration Technique for the Limited-Angle Problem in Computer Tomography*

Sze Fong Yau and Kai Kou Yu

Department of Electrical and Electronic Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

Tomographic reconstruction from limited-angle projection data, also know as a sinogram, is required in many fields, including medical imaging, sonar, and radar. We present a sinogram restoration technique that *restores* a complete sinogram from the available incomplete sinogram. Using two-dimensional sampling theory and the result of Rattey and Lindgen, which shows that the spectral support of a sinogram is bowtie-shaped, a matrix formulation is developed. Restoration of the complete sinogram is then posed as a least-squares minimization problem, which is solved by a novel iterative algorithm. Our technique does not require any a priori knowledge of the underlying object and can be applied to any incomplete sinogram. The algorithm can also be regarded as a variation of the well-known projection-onto-convex-set method with improved computational efficiency. Computer simulation results are presented to demonstrate the efficacy of the proposed technique.

Journal of Imaging Science and Technology 40:3 254-264 (1996)

Introduction

Computer tomography (CT) is a technique used to reconstruct a cross-section image of an object from its line-integral projections, which are taken at different angular views. The general method of CT continues to become an important tool in many areas of application, including medical imaging, acoustic imaging, synthetic-aperture radar, and nondestructive evaluation. It is well known that when enough high-quality line-integral projections, or *raysums,* are collected over a total viewing angle of 180°, the cross-section function of the object can be reconstructed by classical reconstruction methods such as convolution back-projection (CBP) and the algebraic reconstruction technique (ART).[1] Under this circumstance, the collection of line-integral projections is referred to as a complete (or full-angle) set of projections, an image of which is called a *sinogram.*

In many practical situations, a complete sinogram is not available. For example, it may not be possible for the x-ray source to move through the full angular range because of obstacles around the scanning object.[2] In cardiac CT imaging, the carriage containing the x-ray emitters and detectors can travel only part of the way through the full angular range before significant heart motion occurs. Without a restoration technique, the "naive" reconstruction* of an incomplete sinogram is usually rife with artifacts and may be unable to provide even coarse structural information about the object.[3] In these cases, by the projection slice theorem,[4] the two-dimensional spectral information of the underlying object is known for only a limited angle of cone. This is the well-known limited-angle or missing-cone problem.[5] It is an inverse problem that is inherently ill conditioned, and the inversion of the radon transform is severely ill posed.

Techniques for dealing with the limited-angle problem have been proposed by many researchers. In general, these techniques can be classified as transform methods that incorporate little a priori information about the underlying image and as finite series expansion techniques that may incorporate a priori information as constraints. Examples of these methods are the Clark-Palmer-Lawrence interpolation (CPL) method,[6] which performs a coordinate-scaling transformation on the available samples; the affine transformation method of Reeds and Shepp,[7] which uses an affine scale change of the image to expand the available angular span; and the method of projections onto convex sets (POCS).[8-11] Oskoui and Stark[12] compared the performance of the above methods and found that the POCS algorithm performed better than the other two methods.

The success of POCS can be attributed to the fact that a priori information of the object, such as amplitude limits, bounded spatial support, and maximum distance from a reference object, is used to compensate for the lack of the full sinogram. Because this knowledge can be associated with convex sets, the iterative method of POCS can make full use of it to produce reconstruction with improved quality.

Whereas the performance of POCS is superior to that of other existing methods, it is computationally intensive. This is because most of the available a priori information is associated with the object itself, and the available data comprise the incomplete sinogram. The POCS method therefore requires iterations between the object and its sinogram. The computation required for each iteration is therefore that needed for image reconstruction (e.g., by CBP), projection onto the convex sets, and numerical tomography projection of the reconstructed image to go back to the sinogram domain.† Because

*A "naive" reconstruction assumes the missing projections to be identically zero.

† The word *projection* carries double meaning here. In *projection* onto convex sets, the operation of *projection* refers to a mathematical operation of finding a point that lies in a convex set and yet is nearest to another (given) point generally not lying in the set. In tomography *projection,* projection refers to a line integral. A sequence of such line integrals is called a sinogram.

Original manuscript received December 2, 1995. Revised April 11, 1996.

* Supported in part by RGC Earmarked Research Grant HKUST 113/92E and HKUST678/95E and RGC Direct Allocation Grant DAG91/92.EG14.

numerical and systematic errors are invariably introduced in each reconstruction and numerical tomography projection, convergence of the POCS method may not be achieved in practice, and in some cases the reconstruction quality may even deteriorate with more iterations.[8]

The iterative sinogram restoration algorithm (ISRA) proposed in this article can be regarded as a variation of the POCS method. The constraints used in our method belong to a subset of those used by POCS. We require the reconstructed image (1) to conform with the available incomplete sinogram and (2) to have a bounded spatial support. The second requirement poses no limitation in practice because all objects to be reconstructed must be of finite size and the spatial support of the cross section can be obtained by direct measurement. Thus it is fair to say that ISRA is a general restoration method *without* any a priori assumption about the underlying object. Although not implemented in this study, other a priori information, such as a reference image, can easily be incorporated in the proposed method as well, but this will require additional a priori assumptions on the underlying object.

Unlike the POCS method, which uses the information on bounded spatial support of an object directly, in ISRA we translate this information to the sinogram domain. More specifically, given the knowledge that an object is of bounded spatial support, Rattey and Lindgren[13] show that the spectral support of its sinogram is bowtie-shaped. Using this result, information on the spatial support of the object is translated to information on its sinogram.

With the above translation, we present an iterative method that iterates in the sinogram domain *only.* This method eliminates the expensive image reconstruction and numerical tomography projection in iterations of the conventional POCS algorithm, resulting in reduced computational requirements. Further, because numerical and systematic errors are eliminated with the removal of intermediate reconstruction and numerical tomography projection, we have observed a slightly improved image quality in the final reconstruction.

The proposed ISRA is applicable to parallel-beam projections with a uniform sensor array. If fan-beam data are collected, they can be transformed into parallel-beam projections and then the proposed technique may be applied.

This article is organized as follows. We discuss two-dimensional sampling theory and the special signal structure of sinograms and their application to formulation of the limited-angle restoration problem. A system of linear equations that must be solved is established. The problem is then posed as an optimization problem. In the next section, an iterative algorithm is presented to solve the optimization problem. Proof of the convergence of this algorithm is also presented. Computer simulation results are presented to demonstrate the efficacy of the proposed algorithm.

Throughout this discussion, we use the following notations:

I	=	$\sqrt{-1}$
$\delta(\bullet)$	=	Dirac delta function
$(\bullet)^T$	=	transpose
$(\bullet)^*$	=	Hermitian transpose
$\boldsymbol{A} \odot \boldsymbol{B}$	=	Schur-Hadamard (element-by-element) matrix product
$\boldsymbol{A} \otimes \boldsymbol{B}$	=	Kronecker product[14]
A_{ij}, $(\boldsymbol{A})_{ij}$	=	the i, j element of $\boldsymbol{A}$
$\lVert A \rVert_F$	=	$\sqrt{\sum (A_{ij})^2}$ (Frobenious norm of $\boldsymbol{A}$)
vec $(\boldsymbol{A})$	=	the ordered stack of columns of $\boldsymbol{A}$[14]
$\overline{\boldsymbol{A}}$	=	the complex conjugate of $\boldsymbol{A}$
Re$\{\boldsymbol{A}\}$	=	the real part of $\boldsymbol{A}$, defined as $(\boldsymbol{A} + \overline{\boldsymbol{A}})/2$
$\boldsymbol{I}_{n\times n}$	=	$n \times n$ identity matrix
$\mathbf{1}_{m\times n}$	=	$m \times n$ matrix of ones
$\phi_{m\times n}$	=	$n \times n$ matrix of zeros

When no confusion will occur, the dimensions of $\boldsymbol{I}_{n\times n}$, $\mathbf{1}_{m\times n}$ and $\phi_{m\times n}$ are omitted.

Problem Formulation

The central idea of ISRA is to *restore* a complete sinogram, on a *different* sampling lattice, from the observed incomplete sinogram. Many of the existing limited-angle techniques use the available data either to recover the missing projections or to recalculate a whole sinogram with the sampling lattice of the available data. In this article, given a limited-angle sinogram, ISRA restores a complete sinogram on another sampling lattice, which has a *lower* resolution than the incomplete data. We believe this to be a more reasonable approach, because an incomplete sinogram inherently provides less information than a complete one. Simulation results show that restoration on a lower resolution lattice in general provides better overall reconstruction quality and is more robust to noise.

Consider a cross-section function $f(t_1, t_2)$ of an object. Let $F(\omega_{t_1}, \omega_{t_2})$ be its two-dimensional Fourier transform:

$$F(\omega_{t_1}, \omega_{t_2}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t_1, t_2)\exp[-2I\pi(t_1\omega_{t_1} + t_2\omega_{t_2})]\,dt_1 dt_2 \quad (1)$$

Suppose that $f(t_1, t_2)$ is spatially limited to a disk with radius R_m in the t_1–t_2 plane and is essentially bandlimited‡ in the frequency domain to a disk with radius W_M, as shown in Fig. 1. The continuous sinogram, or the radon transform of $f(t_1, t_2)$, is a continuous two-dimensional signal, defined as

$$x(\phi, \rho) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t_1, t_2)\partial(\rho - t_1\cos\phi - t_2\sin\phi)\,dt_1 dt_2. \quad (2)$$

‡ It is well known that a spatially limited signal cannot be bandlimited. A spatially limited function is *essentially bandlimited* to S if there exists a spectral support S in the frequency domain such that most of the signal energy (say 99%) is concentrated in S.

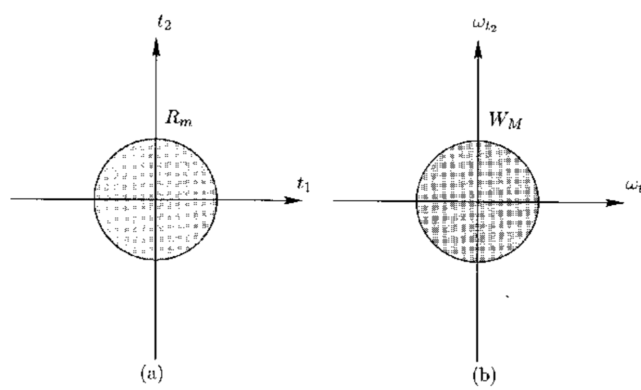

Figure 1. (a) Spatial support of function $f(t_1, t_2)$; (b) Support of $F(\omega_{t_1}, \omega_{t_2})$

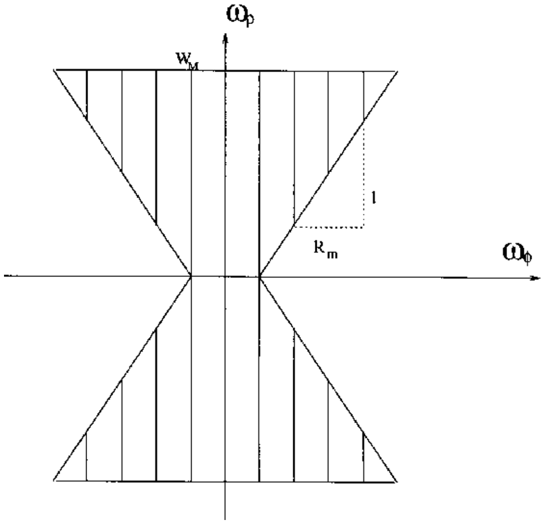

Figure 2. Spectral support of finite-length bowtie.

Let $X(\omega_\phi,\omega_\rho)$ be the two-dimensional Fourier transform of $x(\phi,\rho)$:

$$X(\omega_\phi,\omega_\rho)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x(\phi,\rho)\exp(-I2\pi(\phi\omega_\phi+\rho\omega_\rho))d\phi d\rho, \quad (3)$$

then approximately 99% of the total energy of $X(\omega_\phi,\omega_\rho)$ is within finite-length $R_m W_M$ bowtie,[13] as illustrated in Fig. 2. Hence, $X(\omega_\phi,\omega_\rho)$ is essentially bandlimited to the finite-length $R_m W_M$ bowtie.

When a continuous sinogram is sampled with sampling period $\boldsymbol{T} = [T_\phi, T_\rho]$ satisfying the sampling requirements[13]

$$T_\phi < \frac{\pi}{K-1}, \quad T_\rho < \frac{2R_m}{L-1}, \quad (4)$$

$$K \geq R_m W_M + 2, \quad (5)$$

and

$$L \geq \frac{2}{\pi} R_m W_M + 1, \quad (6)$$

the original continuous signal can be reconstructed by the interpolation formula[15]:

$$x(\phi,\rho) = \sum_{n_\phi=-\infty}^{\infty}\sum_{n_\rho=-\infty}^{\infty} x_d(n_\phi,n_\rho)\,\mathrm{sinc}\left(\frac{\phi-\mathrm{n}_\phi T_\phi}{T_\phi}\right)\mathrm{sinc}\left(\frac{\rho-n_\rho T_\rho+R_m}{T_\rho}\right), \quad (7)$$

$x_d(n_\phi, n_\rho)$ *is* the sampled version of the sinogram given by

$$x_d(n_\phi, n_\rho) = x(n_\phi T_\phi,\ n_\rho T_\rho - R_m) \quad (8)$$

for $-\infty < n_\phi, n_\rho < \infty$, and

$$\sin(a) = \frac{\sin(\pi a)}{\pi a}. \quad (9)$$

For signal $f(t_1, t_2)$ with finite support as shown in Fig. 1 (a), its sinogram will also be of finite support; hence,

$$x(\phi,\rho) = \sum_{n_\phi=0}^{N_\phi-1}\sum_{n_\rho=0}^{N_\rho-1} x_d(n_\phi,n_\rho)\mathrm{sinc}\left(\frac{\phi-\mathrm{n}_\phi T_\phi}{T_\phi}\right)\mathrm{sinc}\left(\frac{\rho-n_\rho T_\rho+R_m}{T_\rho}\right), \quad (10)$$

where

$$(N_\phi - 1)T_\phi = 2p \text{ and } (N_\rho - 1)T_\rho = 2R_m. \quad (11)$$

Suppose that instead of sampling at $\boldsymbol{T} = [T_\phi, T_\rho]$, the sinogram $x(\phi, \rho)$ is in fact sampled at a lower sampling period $\tilde{\boldsymbol{T}} = [\tilde{T}_\phi, \tilde{T}_\rho]$ (i.e., at a higher sampling rate):

$$\tilde{T}_\phi < T_\phi \text{ and } \tilde{T}_\rho < T_\rho. \quad (12)$$

The sampled sinogram at sampling period $\tilde{\boldsymbol{T}}$ is then

$$x_l(m_\phi, m_\rho) = x(m_\phi \tilde{T}_\phi, m_\rho \tilde{T}_\rho - R_m) \quad (13)$$

for $m_\phi = 0, 1, \ldots M_\phi - 1$, and $m_\rho = 0, 1, \ldots, M_\rho - 1$. Note that $(M_\phi - 1)\tilde{\boldsymbol{T}}_\phi, = (N_\phi - 1)T_\phi = 2\pi$ and $(M_\rho - 1)\tilde{\boldsymbol{T}}_\rho = (N_\rho - 1)T_\rho = 2R_m$. As shown in Fig. 3, the available samples of sinograms have a higher resolution or spatial frequency content than $x_d(n_\phi, n_\rho)$. Combining Eqs. 10 and 13, we have

$$\sum_{n_\phi=0}^{N_\phi-1}\sum_{n_\rho=0}^{N_\rho-1} x_d(n_\phi,n_\rho)\,\mathrm{sinc}\left(\frac{m_\phi\tilde{T}_\phi-n_\phi T_\phi}{T_\phi}\right)\mathrm{sinc}\left(\frac{m_\rho\tilde{T}_\rho-n_\rho T_\rho}{T_\rho}\right) = x_l(m_\phi,m_\rho) \quad (14)$$

for $m_\phi = 0, 1, \ldots, M_\phi - 1$, and $m_\rho = 0, 1, \ldots, M_\rho - 1$.

Define matrices $\boldsymbol{X} \in \boldsymbol{R}^{\boldsymbol{N}_\phi \times N_\rho}$ and $\tilde{\boldsymbol{X}}_{\boldsymbol{l}} \in \boldsymbol{R}^{\boldsymbol{M}_\phi \times M_\rho}$ such that

$$(\boldsymbol{X})_{ij} = x_d(i-1, j-1) \quad (15)$$

$$(\tilde{\boldsymbol{X}}_l)_{ij} = x_l(i-1, j-1), \quad (16)$$

and interpolation matrices $\boldsymbol{S}_1 \in \boldsymbol{R}^{\boldsymbol{M}_\phi \times N_\phi}$ and $\boldsymbol{S}_2 \in \boldsymbol{R}^{\boldsymbol{M}_\rho \times N_\rho}$ such that

$$(\boldsymbol{S}_1)_{ij} = \mathrm{sinc}\left(\frac{(i-1)\tilde{T}_\phi-(j-1)T_\phi}{T_\phi}\right) = \mathrm{sinc}\left(\frac{(i-1)(N_\phi-1)}{M_\phi-1}-j+1\right), \quad (17)$$

$$(S_2)_{ij} = \mathrm{sinc}\left(\frac{(i-1)\tilde{T}_\rho-(j-1)T_\rho}{T_\rho}\right) = \mathrm{sinc}\left(\frac{(i-1)(N_\rho-1)}{M_\rho-1}-j+1\right). \quad (18)$$

then Eq. 14 can be written more compactly as

$$\boldsymbol{S}_1\boldsymbol{X}\boldsymbol{S}_2^* = \tilde{\boldsymbol{X}}_l. \quad (19)$$

However, as shown in Fig. 4, the sinogram $x(\phi, \rho)$ is sampled only at viewing angles $0 \leq \phi \leq \Theta$, where $0 < \Theta < \pi$. Thus, when sampled at period $\tilde{\boldsymbol{T}}$, $x_l(m_\phi, m_\rho)$ is measured for $\mathrm{m}_\phi = 0,1,\ldots, P-1$, and $m_\rho = 0,1,\ldots, M_\rho - 1$, where $P < \frac{M_\phi}{2}$. In addition, because $x(\phi + \pi, \rho) = x(\phi, -\rho)$, we can get $x_l(m_\phi\, m_\rho)$ for $m_\phi = \frac{M_\phi}{2}, \ldots, \frac{M_\phi}{2} + P - 1$ and $m_\rho = 0,1, \ldots, M_\rho - 1$ by

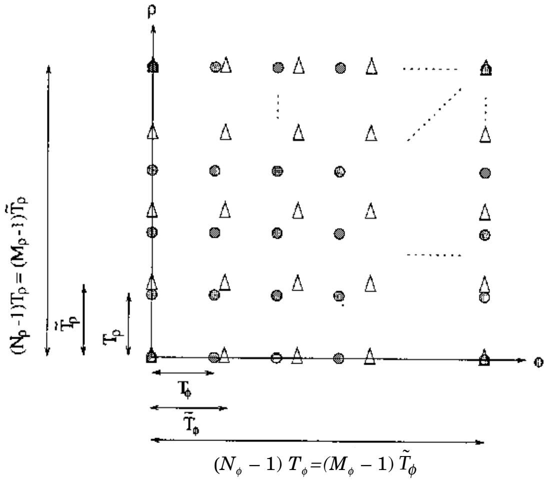

Figure 3. Sampling lattices of finite support sinogram: • represents the samples $x_l(m_\phi,m_\rho)$ and Δ represents samples $x_d(n_\phi,n_\rho)$.

symmetry. Despite this, we cannot have all equations in Eq. 14 for all m_ϕ and m_ρ. Thus, the available equations in the limited-data case should be written as

$$(\boldsymbol{S}_1\boldsymbol{X}\boldsymbol{S}_2)\odot\boldsymbol{Z}=\boldsymbol{X}_l, \tag{20}$$

where $\boldsymbol{X}_l$ is the data maxtrix defined as

$$(x_l)_{ij}=\begin{cases}x_l(i-1,j-1) \text{ if } x_l(i-1,j-1) \text{ is available}\\ 0 \ \text{ otherwise,}\end{cases} \tag{21}$$

and $\boldsymbol{Z}\in\boldsymbol{R}^{M_\phi\times M_\rho}$ is an indicator matrix such that

$$(\boldsymbol{Z})_{ij}=\begin{cases}1 \ \text{ if } x_l(i-1,j-1) \text{ is available}\\ 0 \ \text{ otherwise,}\end{cases}$$

$$=\begin{cases}1 \ \text{ for } i=1,\ldots,P,\dfrac{M_\phi}{2}+1,\ldots,\dfrac{M_\phi}{2}+P \text{ and for all } j\\ 0 \ \text{ otherwise.}\end{cases}, \tag{22}$$

From the above discussion, we understand the following:

1. If a *complete* sampled sinogram $\boldsymbol{X}$ with sampling period $\boldsymbol{T}$ is available, high-quality reconstruction of $f(t_1,t_2)$ can be obtained.
2. We have access to incomplete data $\boldsymbol{X}_l$, which is related to $\boldsymbol{X}$ through Eq. 20.

The limited-angle problem can therefore be solved if we can compute the complete sampled sinogram $\boldsymbol{X}$ from $\boldsymbol{X}_l$ alone. This computation, however, is not possible in practice, because there are $N_\phi N_\rho$ unknowns in $\boldsymbol{X}$, which are usually more than the $2PM_\rho$ nontrivial equations available in Eq. 20. Thus, additional information is needed to compensate this situation. Besides, even if we have enough equations, the system of equations in Eq. 20 is still ill-conditioned and needs to be regularized.

Fortunately, inspection of the spectral support of $X(\omega_\phi,\omega_\rho)$ in Fig. 2 reveals that supplementary information on $\boldsymbol{X}$ is available. Because the spectral support of the original sinogram $x(\phi,\rho)$ is bowtie-shaped, it is straightforward to translate this information into additional constraint. Let $X_d(k_\phi,k_\rho)$ be the two-dimensional discrete Fourier transform of $x_d(n_\phi,n_\rho)$; then

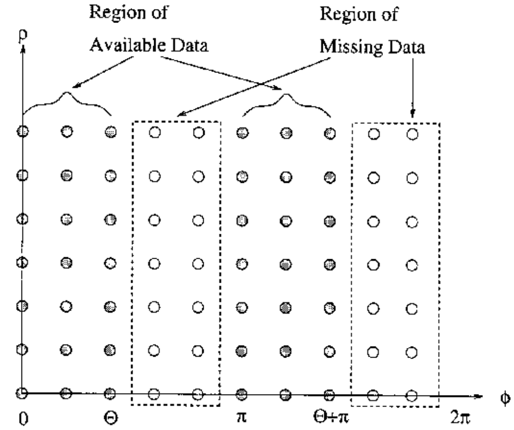

Figure 4. Regions of available and missing data.

$$X_d(k_\phi,k_\rho)=\sum_{n_\phi=0}^{N_\phi-1}\sum_{n_\rho=0}^{N_\rho-1}x_d(n_\phi,n_\rho)\exp(-I2\pi(\frac{n_\phi k_\phi}{N_\phi}+\frac{n_\rho k_\rho}{N_\rho}))$$
$$=0 \tag{23}$$

for appropriate values of k_ϕ and k_ρ. Defining the discrete Fourier transformation matrices $\boldsymbol{F}_1\in\boldsymbol{R}^{N_\phi\times N_\phi}$ and $\boldsymbol{F}_2\in\boldsymbol{R}^{N_\rho\times N_\rho}$ as

$$(\boldsymbol{F}_1)_{ij}=\exp(-2\pi I\frac{(i-1)(j-1)}{N_\phi}\Big), \tag{24}$$

$$(\boldsymbol{F}_2)_{ij}=\exp(-2\pi I\frac{(i-1)(j-1)}{N_\rho}\Big), \tag{25}$$

and an indicator matrix $\boldsymbol{U}\in\boldsymbol{R}^{N_\rho\times N_\rho}$ such that

$$(\boldsymbol{U})_{ij}=\begin{cases}1 \ \text{ if } X_d(i,j)=0\\ 0 \quad \text{otherwise}\end{cases}, \tag{26}$$

then Eq. 23 can be written more compactly as

$$(\boldsymbol{F}_1\boldsymbol{X}\boldsymbol{F}_2^T)\odot\boldsymbol{U}=0. \tag{27}$$

The indicator matrix $\boldsymbol{U}$ in Eq. 26 depends on the spatial extent of the function. In practice, we can set $2R_m$ equal to the maximum extent of the sampling window and specify $\boldsymbol{U}$ accordingly.

Equations 20 and 27 form a set of simultaneous linear equations. Because $\boldsymbol{F}_1$, $\boldsymbol{F}_2$, $\boldsymbol{S}_1$, and $\boldsymbol{S}_2$ are all full-rank matrices, it can be shown that if the number of active equations is greater than the number of unknowns, i.e.,

$$2PM_\rho+\|\boldsymbol{U}\|_F^2>N_\phi N_\rho, \tag{28}$$

then the system is overdetermined. Because the signal under consideration is actually only essentially bandlimited, Eqs. 20 and 27 hold only approximately. The problem of determining $\boldsymbol{X}$ can therefore be posed as the following least-squares optimization problem.

Problem Formulation for Sinogram Restoration. The restored complete sinogram is given by $\hat{\boldsymbol{X}}$, which is defined as

$$\hat{\boldsymbol{X}} = \arg\min_{\boldsymbol{X}} J(\boldsymbol{X}), \tag{29}$$

where $J(\boldsymbol{X})$ is the cost function given by

$$J(\boldsymbol{X}) = \lambda_l \left\| \boldsymbol{S}_1 \boldsymbol{X} \boldsymbol{S}_2^* - \boldsymbol{X}_l \right\|_F^2 + (1-\lambda_l) \left\| (\boldsymbol{F}_1 \boldsymbol{X} \boldsymbol{F}_2^T) \odot \boldsymbol{U} \right\|_F^2 \tag{30}$$

and the parameter $0 < \lambda_l < 1$ is a regularization parameter between Eqs. 20 and 27.

The selection of an *optimal* regularization parameter λ_l is a difficult task. Generally speaking, the best choice is not known a priori for a given ill-posed problem. Reeves and Merserau[16] presented a method for obtaining an optimal estimate of the regularization parameter by using generalized cross-validation. However, the generation of the estimate requires the computation of the system eigenvalues. This is difficult in our application because of the large size of the problem. The parameter λ_l is assigned to achieve a compromise between two constraints that hold only approximately. Because the errors in these equations are not known, the best assignment is not provided at this moment. Heuristic approaches are therefore used in choosing these parameters. For example, if the given sinogram is noisy, a smaller λ_l will give a better result.

An Iterative Algorithm

As discussed in the previous section, the limited-angle restoration of a sinogram can be considered as the solution of $\hat{X}$ defined by Eq. 29. Equivalently, $J(\boldsymbol{X})$ as defined by Eq. 30 needs be minimized. Although different methods exist to solve this multidimensional minimization problem, most of them require reorganizing the matrix $\boldsymbol{X}$ into a column vector. To be more specific, we can rewrite Eqs. 29 and 30 into vector form:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \left\| \boldsymbol{H}\boldsymbol{x} - \boldsymbol{y} \right\|_F^2, \tag{31}$$

where $\boldsymbol{x} = \mathrm{vec}(\boldsymbol{X})$.[14] Matrix $\boldsymbol{H}$ and data vector $\boldsymbol{y}$ are formed by rearranging the matrices in concert. The optimal solution is then given by

$$\hat{\boldsymbol{x}} = (\boldsymbol{H}^* \boldsymbol{H})^{-1} \boldsymbol{H}^* \boldsymbol{y}. \tag{32}$$

Unfortunately, there is a major difficulty to this approach. Due to the irregular structure of the indicator matrices $\boldsymbol{Z}$ and $\boldsymbol{U}$ in Eq. 30, formation of $\boldsymbol{H}$ and $\boldsymbol{y}$ is difficult. Moreover, the matrix $\boldsymbol{H}$ involved in solving Eq. 29 is of dimension $(2PM_\rho + \|\boldsymbol{U}\|_F^2) \times N_\phi N_\rho$. For example, in Experiment 1 in the following section, a 56 × 56-pixel sinogram is to be restored from a sinogram with 23 available views and 64 raysums per view. Because approximately half of the DFT samples of a complete sinogram are outside its bowtie-shaped spectral support, the total number of nontrivial equations (i.e., number of rows in $\boldsymbol{H}$) is approximately 2 × 23 × 64 + 56 × 56/2 = 4512. The number of columns of $\boldsymbol{H}$ is 56 × 56 = 3136. Note further that $\boldsymbol{H}$ is *not* a sparse matrix—it has about 14 million entries. Such a large matrix is difficult to store in digital computers, and computation of its pseudoinverse is even more difficult.

An iterative method proposed by Kaczmarz[17] is the row action projection (RAP) algorithm, which involves iterative orthogonal projections onto the hyperplanes specified by each active equation in Eq. 32. The update equation in this method is given by

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \gamma_k \frac{y_{i_k} - (\boldsymbol{H})_{i_k,:}^T \boldsymbol{x}_k}{\left\| (\boldsymbol{H})_{i_k,:} \right\|_F^2} (\boldsymbol{H})_{i_k,:}^T, \tag{33}$$

where γ_k is the relaxation factor, y_i is the i-th element of $\boldsymbol{y}$ and $(\boldsymbol{H})_{i,:}$ is the i-th row of $\boldsymbol{H}$. The variable i_k is the cyclic control parameter defined by $i_k = k (\mathrm{mod}\ L)$, where L is the number of rows in $\boldsymbol{H}$. The RAP algorithm was shown to converge to the least-squares solution[18] and is more computationally attractive than other methods.[19] The RAP algorithm is the basis for ART[1] and has been applied to resolution enhancement in CT.[20] Inspection of Eq. 33 indicates that although the RAP algorithm makes it possible to solve the minimization problem in Eq. 31, it is still computationally intensive. Besides, RAP assumes no special structure in matrix $\boldsymbol{H}$, which may be invaluable in devising more computationally efficient algorithms.

In this section, we utilize the special structure of the cost function $J(\boldsymbol{X})$ to derive an iterative algorithm that is based on the method of alternating projection. We will show that, independent of initialization, the algorithm will converge to the desired solution. In addition, the algorithm does not require formation of matrix $\boldsymbol{H}$, matrix-to-vector conversion, and matrix inversion for each iteration. Fast implementation of the proposed algorithm will also be discussed. To our best knowledge, we have not been able to find a similar algorithm in the literature.

We first note that the cost function $J(\boldsymbol{X})$ is a quadratic (and hence convex) function in $\boldsymbol{X}$. Hence, any descending algorithm on $J(\boldsymbol{X})$ will lead to the correct unique solution of the problem of Eq. 29.[21,22] Next, we prove the following theorem.

Theorem 1. *Let $\boldsymbol{A} \in \boldsymbol{C}^{P\times Q}$ and $\boldsymbol{B} \in \boldsymbol{C}^{R\times S}$ be full-rank matrices such that $PR \geq QS$ and $\boldsymbol{Y} \in \boldsymbol{C}^{P\times R}$ are given. Let $\hat{\boldsymbol{X}} \in \boldsymbol{R}^{Q\times S}$ be defined as*

$$\hat{\boldsymbol{X}} = \arg\min_{\boldsymbol{X}} \left\| \boldsymbol{A}\boldsymbol{X}\boldsymbol{B}^* - \boldsymbol{Y} \right\|_F^2; \tag{34}$$

then

$$\hat{\boldsymbol{X}} = Re\{(\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*\boldsymbol{Y}\boldsymbol{B}(\boldsymbol{B}^*\boldsymbol{B})^{-1}\}. \tag{35}$$

Proof: See Appendix A.

We can state the optimization problem of Eq. 29 in a form similar to Eq. 34. Define the matrices

$$\boldsymbol{A} = \begin{bmatrix} \sqrt{\lambda_l}\boldsymbol{S}_1 \\ \sqrt{1-\lambda_l}\boldsymbol{F}_1 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} \sqrt{\boldsymbol{S}_2} \\ \overline{\boldsymbol{F}}_2 \end{bmatrix}, \boldsymbol{Y} = \begin{bmatrix} \sqrt{\lambda_l}\boldsymbol{X}_l & 0 \\ 0 & \boldsymbol{Y}_2 \end{bmatrix}, \tag{36}$$

and a combined indicator matrix

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{1} - \boldsymbol{Z} & \boldsymbol{1} \\ \boldsymbol{1} & \boldsymbol{1} - \boldsymbol{U} \end{bmatrix}, \tag{37}$$

Then $J(\boldsymbol{X})$ as defined in Eq. 30 can be rewritten as

$$J(\boldsymbol{X}) = \left\| (\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}^* - \boldsymbol{Y}) \odot (\boldsymbol{1} - \boldsymbol{W}) \right\|_F^2. \tag{38}$$

Because Eq. 38 is of similar form to Eq. 34, one might expect Theorem 1 to be helpful for finding $\hat{X}$. However, Theorem 1 cannot be directly applied to minimize Eq. 38 because there is an element-by-element matrix multiplication in the cost function. The following theorem introduces a dummy matrix in another optimization problem that has the same unique solution as Eq. 29.

Theorem 2. *Let $\hat{X}$ be defined as in Eq. 29 and $\boldsymbol{X}^+$, $\boldsymbol{R}^+$ be defined as*

$$\{\boldsymbol{X}^+, \boldsymbol{R}^+\} = \arg\min_{\boldsymbol{X}, \boldsymbol{R}} \left\| \boldsymbol{AXB}^* - \boldsymbol{Y} - \boldsymbol{W} \odot \boldsymbol{R} \right\|_F^2, \tag{39}$$

*where **A, B, Y** and **W** are as defined in Eqs. 36 and 37.*

Assume that $\hat{X}$ as defined in Eq. 29 is unique; then $\boldsymbol{X}^+$ as defined in Eq. 39 is unique and further

$$\hat{\boldsymbol{X}} = \boldsymbol{X}^+. \tag{40}$$

Proof: See Appendix B.

Using Theorems 1 and 2, the following iterative algorithm based on the principle of alternating projection can be used to solve the optimization problem (Eq. 29).

Theorem 3. *Let $\boldsymbol{P}_A$, $\boldsymbol{P}_B$ be pseudoinverses of **A** and **B** defined as*

$$\boldsymbol{P}_A = (\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*, \tag{41}$$

$$\boldsymbol{P}_B = (\boldsymbol{B}^*\boldsymbol{B})^{-1}\boldsymbol{B}^*, \tag{42}$$

respectively. Let a sequence of matrices $\boldsymbol{X}_k$ and $\boldsymbol{R}_k$ be defined by the iterative equations

$$\boldsymbol{X}_{k+1} = Re\{\boldsymbol{P}_A(\boldsymbol{Y} + \boldsymbol{W} \odot \boldsymbol{R}_k)\boldsymbol{P}_B^*\} \tag{43}$$

and

$$R_k = (\boldsymbol{AX}_k\boldsymbol{B}^* - \boldsymbol{Y}) \odot \boldsymbol{W}, \tag{44}$$

where $\boldsymbol{X}_0 \in \boldsymbol{R}^{N_\phi \times N_\rho}$ is arbitrary. Assume that $\hat{\boldsymbol{X}}$ as defined in Eq. 29 is unique, then

$$\hat{\boldsymbol{X}} = \lim_{k\to\infty} \boldsymbol{X}_k. \tag{45}$$

Proof: See Appendix C.

Independent of initialization, Theorem 3 provides an iterative method to locate the global minimum of the optimization problem of Eq. 29. Although Theorem 3 shows that the sequence of matrix $\boldsymbol{X}_k$ always converges to the optimal solution, the convergence rate is sometimes slow. Because the feasible region of solution and the cost function are both convex, the convergence rate of the algorithm can be increased by introducing a relaxation parameter. We state the modified algorithm in the following corollary.

Corollary 1 (Iterative Sinogram Restoration Algorithm). *Let β be a relaxation parameter with $0 < \beta < 2$. Let a sequence of matrices $\boldsymbol{X}_k$ and $\boldsymbol{R}_k$ be defined by the iterative equations*

$$\boldsymbol{X}_{k+1} = \beta\, Re\{\boldsymbol{P}_A(\boldsymbol{Y} + \boldsymbol{W} \odot \boldsymbol{R}_k)\boldsymbol{P}_B^*\} + (1-\beta)\boldsymbol{X}_k \tag{46}$$

and

$$\boldsymbol{R}_k = (\boldsymbol{AX}_k\boldsymbol{B}^* - \boldsymbol{Y}) \odot \boldsymbol{W}, \tag{47}$$

where $\boldsymbol{X}_0 \in \boldsymbol{R}^{N_\phi \times N_\rho}$ is arbitrary. Assume that $\hat{\boldsymbol{X}}$ as defined in Eq. 29 is unique, then

$$\hat{\boldsymbol{X}} = \lim_{k\to\infty} \boldsymbol{X}_k. \tag{48}$$

Proof: Let the iteration process of Eqs. 43 and 44 in Theorem 3 be written as $\boldsymbol{X}_{k+1} = P\boldsymbol{X}_k$. Then, by Theorem 3,

$$\hat{\boldsymbol{X}} = \lim_{k\to\infty} P^k \boldsymbol{X}_0. \tag{49}$$

On the other hand, the operation in Eq. 46 can be stated as $\boldsymbol{X}_{k+1} = \{\beta P + (1-\beta)\boldsymbol{I}\}\boldsymbol{X}_k$.

Because the solution space and the cost function are both convex, by the result in Ref. 23, Theorem 2.4-1, for $0 < \beta < 2$,

$$\hat{\boldsymbol{X}} = \lim_{k\to\infty} P^k \boldsymbol{X}_k \Rightarrow \hat{\boldsymbol{X}} = \lim_{k\to\infty} \{\beta P + (1-\beta)\boldsymbol{I}\}^k \boldsymbol{X}_0. \tag{50}$$

With a suitable choice of β, the rate of convergence of ISRA can be accelerated. When $\beta = 1$, the iterative equations, Eqs. 46 and 47, are equivalent to Eqs. 43 and 44. In this situation, ISRA is said to be *unirelaxed.*

In actual implementation of the proposed iterative algorithm, an effective stopping mechanism is required. Because the cost function $J(\boldsymbol{X})$ is convex and is descending in every iteration, we define the following cost function ratio for the k-th iteration:

$$g(k) = \frac{J(\boldsymbol{X}_k)}{J(\boldsymbol{X}_0)} \times 100\%, \tag{51}$$

which is always positive and less than 100%. Using this ratio, a stopping criterion is proposed as

$$g(k-1) - g(k) < \delta \tag{52}$$

for $0 < \delta < 100\%$. To ensure the convergence of the iterative algorithm, a very small value of δ should be used.

Substituting Eq. 47 and the corresponding matrices in $J(\boldsymbol{X})$ into Eq. 46, the ISRA for the limited-angle problem is summarized in Table I. In Table I, we have replaced matrix multiplications involving $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ by fast Fourier transform (FFT). This operation significantly reduces the computation time in each iteration. In Eq. 56, we have used $\mathcal{F}_1\{\bullet\}$ to denote the one-dimensional FFT operator on column vectors of a matrix and $\mathcal{F}_2\{\bullet\}$ to denote the two-dimensional FFT operator. It is also worth noting that in actual implementation $\boldsymbol{P}_A$ and $\boldsymbol{P}_B$ need to be calculated only once. The sizes of matrices $\boldsymbol{P}_A$ and $\boldsymbol{P}_B$ are also more manageable: $\boldsymbol{P}_A$ is $N_\phi \times (M_\phi + N_\phi)$ and $\boldsymbol{P}_B$ is $N_\rho \times (M_\rho + N_\rho)$. Thus, the dimensions of $\boldsymbol{P}_A$ and $\boldsymbol{P}_B$ in Experiment 1 in the next section are both 56×120.

Experimental Results

In this section we present some simulation results to demonstrate the performance of ISRA in solving the limited-angle problem. Before applying the ISRA, we must decide several parameters. The first one is the extent of the bowtie-shaped spectral support of the sinogram. This region not only specifies the positions where the discrete Fourier transforms of a sinogram are known to be zeros but also determines the validity of Eq. 28 in the application of ISRA. For all experiments in this study, we assume that the largest spatial extent of the object equals the width of the sampling window. Thus, we set the edges of the bowtie (see Fig. 2) to have a slope of unity. Consequently, this region occupies approximately 50% of the whole spectrum and the number of zeros in the discrete Fourier transform is $\|U\|_F^2 \approx N_\phi N_\rho / 2$ in Eq. 28. In fact, for the head phantom used, the designed spectral support contains more than 99.99% of the spectral energy of its sinogram.

TABLE I. Iterative Sinogram Restoration Algorithm (ISRA)

1. Let $\boldsymbol{S}_1$, $\boldsymbol{S}_2$, $\boldsymbol{F}_1$, $\boldsymbol{F}_2$, $\boldsymbol{X}_l$, $\boldsymbol{Z}$ and $\boldsymbol{U}$ be as defined in Eqs. 17, 18, 24, 25, 16, 22, and 26, respectively. Let $k = 0$, $\boldsymbol{X}_0 = \mathbf{0}_{N_\phi \times N_\rho}$,

$$\boldsymbol{P}_A = (\lambda_l \boldsymbol{S}_1^* \boldsymbol{S}_1 + (1-\lambda_l) N_\phi \boldsymbol{I})^{-1} [\sqrt{\lambda_l} \boldsymbol{S}_1^* \quad \sqrt{1-\lambda_l}\, \boldsymbol{F}_1^*], \tag{53}$$

(54)

$$\boldsymbol{W} = \begin{bmatrix} \mathbf{1} - \mathbf{Z} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} - \mathbf{U} \end{bmatrix}. \tag{55}$$

2. At iteration $k + 1$,

$$\boldsymbol{X}_{k+1} = (1-\beta)\boldsymbol{X}_k + \beta\, \mathrm{Re}\left\{ \boldsymbol{P}_A \begin{bmatrix} \boldsymbol{X}_T & \sqrt{\lambda_1}\, \boldsymbol{S}_1 (F_1\{X_k^T\})^T \\ \sqrt{1-\lambda_l}\, F_1\{\boldsymbol{X}_k\} \boldsymbol{S}_2^* & \sqrt{1-\lambda_l}\, (F_2\{\boldsymbol{X}_k\}) \odot (1-\boldsymbol{U}) \end{bmatrix} \boldsymbol{P}_B^* \right\} \tag{56}$$

where

$$\boldsymbol{X}_T = \sqrt{\lambda_l}\, [\boldsymbol{X}_l \odot Z + (\boldsymbol{S}_1 \boldsymbol{X}_k \boldsymbol{S}_2^*) \odot (1-Z)]. \tag{57}$$

3. If the stopping criterion of Eq. 52 is true, then stop; otherwise $k = k + 1$ and go to Step 2.

Although the optimal choice of regularization parameter λ_l is a difficult task, we have been able to find a proper value for satisfactory results. Generally speaking, results obtained from different regularization parameters within a small range are visually alike. In Experiments 1, 2, and 4, we have used $\lambda_l = 0.75$. In Experiment 3, because there is noise, a smaller value, $\lambda_l = 0.6$, is used. In all simulations we set the initial sinogram matrix $\boldsymbol{X}_0 = \mathbf{0}_{N_\phi \times N_\rho}$. As mentioned in the previous section, the solution provided by ISRA is unique to the problem formulation and is independent of the relaxation parameter β. However, our experience shows that in most cases a large value of β can reduce the number of iterations needed to meet the convergence requirement. The relaxation parameter β was set to 1.9 in all the experiments.

To evaluate the performance of the proposed sinogram restoration technique, visual evaluation is, of course, the most straightforward method. To this end, we display the original sinograms, available incomplete sinograms, and sinograms restored by ISRA. Corresponding reconstructions from these sinograms are displayed as well.

To evaluate the performance of the proposed method quantitatively, we define the relative error of a complete sinogram as

$$\mathrm{RE_S} = \frac{\|\boldsymbol{X}^o - \hat{\boldsymbol{X}}\|_F}{\|\boldsymbol{X}^o\|_F} \times 100\%, \tag{58}$$

where $\boldsymbol{X}^o$ is the original sinogram and $\hat{\boldsymbol{X}} = \boldsymbol{X}_k$ is the sinogram restored by ISRA at the k-th iteration. Similarly the relative error of an estimated image in object space is defined as

$$\mathrm{RE_O} = \frac{\|f - \hat{\boldsymbol{f}}\|_F}{\|f\|_F} \times 100\%, \tag{59}$$

where f is the original image and $\hat{f}$ is the CBP reconstruction from the restored sinogram $\hat{\mathbf{X}}$.

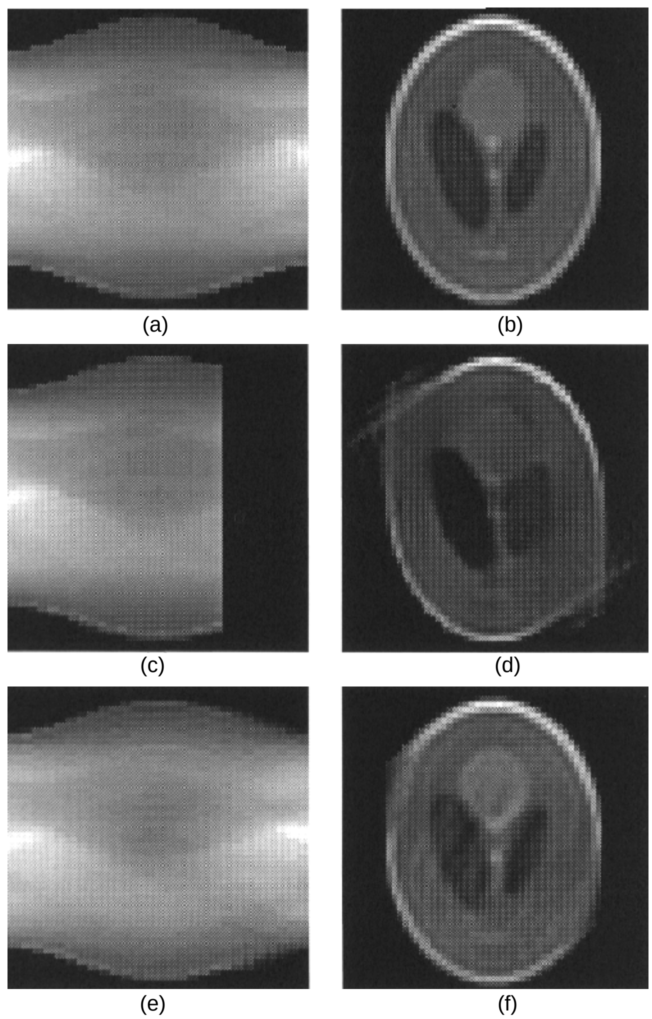

Figure 5. Experiment 1. A limited-angle problem of 129°: (a) original sinogram; (b) original image; (c) available sinogram; (d) "naive" reconstruction from (c); (e) sinogram restored by ISRA; and (f) image reconstructed from restored sinogram.

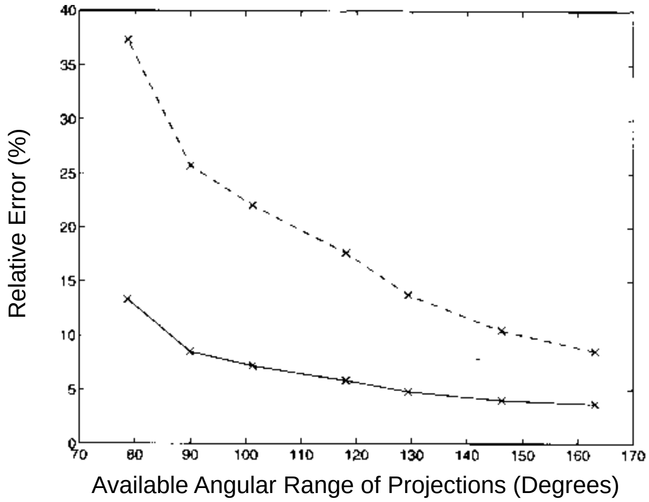

Figure 6. Experiment 2. Relative error curves of the restored sinogram (*solid line*) and of the reconstructed image (*dashed line*) for different available angular ranges.

Experiment 1. Limited-Angle Reconstruction of Head Phantom. The object used in this simulation is the Shepp–Logan head phantom. Figure 5(b) shows a 56×56-pixel square image of the head phantom, which is reconstructed by CBP from the complete sinogram in Fig. 5(a), which has $N_\phi/2 = 28$ views (from 0 to π radian) and $N_\rho = 56$ raysums per view. Suppose the sinogram is actually sampled with $M_\phi/2 = 32$ views (from 0 to π radian) and $M_\rho = 64$ raysums per view but only $P = 23$ views (i.e., ($\Theta = 129°$ angular view) are available. Figure 5(c) shows the available sinogram and Figure 5(d) shows the "naive" CBP reconstruction assuming zero entries in the missing part. Whereas the boundary of the object suffers distortion, the ellipse on the right-hand side of the phantom is severely blurred and the gray level of the whole picture is obviously not accurate. To enhance the observed image, ISRA is applied to restore a 56×56-pixel complete sinogram from the limited-angle data. The restored complete sinogram is shown in Fig. 5(e), where it can be seen that the projections in the missing region are recovered. Figure 5(f) displays a reconstruction of the head phantom from the restored sinogram. Visual examination shows that the right ellipse is recovered with better quality than that of the limited-angle data.

Experiment 2. Different Angular Range of Missing Region. This experiment demonstrates the performance of the ISRA in solving the limited-angle problem with different angular missing range. Here the simulation setting is the same as in Experiment 1, with the exception that different values of P, which indicates the available region of projection data, are used. We repeated Experiment 1 with $P = 29$, 26, 23, 21, 18, 16 and 14, which are equivalent to available angular views of $\Theta = 163°$, 146°, 129°, 118°, 101°, 90°, and 78°. Figure 6 shows the relative error curves of the restored sinogram and reconstructed image in this experiment. It is found that the relative error curves behave linearly for $90° < \Theta < 180°$ but increase rapidly for $\Theta < 90°$. Figure 7 shows the results of two selected runs, $\Theta = 146°$ and 90°. Whereas in both cases the right ellipses are successfully recovered, the restored image for $\Theta = 146°$ is of superior quality.

We now consider a more difficult situation in CT. When a limited-angle sinogram corrupted by noise is available, the "naive" CBP reconstruction is degraded and features are difficult to identify. Experiment 3 demonstrates the performance of ISRA in this situation.

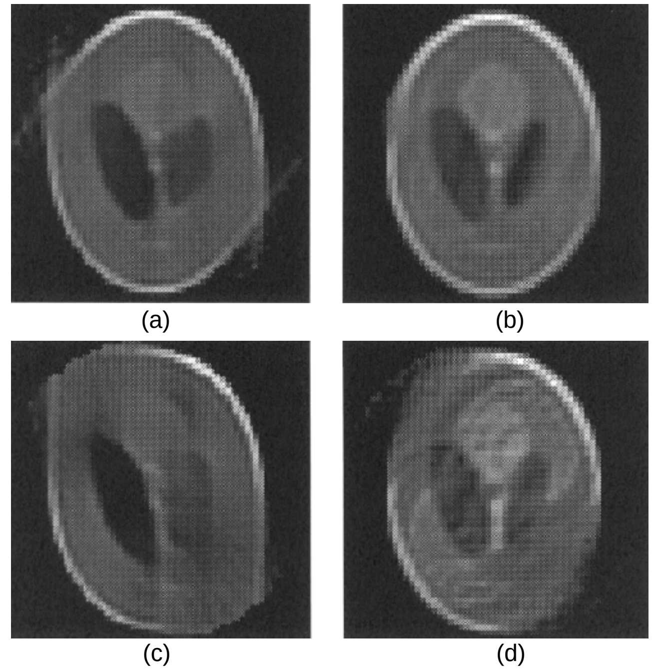

Figure 7. Experiment 2. Images in different available angular ranges: (a) "naive" reconstruction and (b) reconstruction using sinogram restored by ISRA for $\Theta = 146°$; (c) "naive" reconstruction and (d) reconstruction using sinogram restored by ISRA for $\Theta = 90°$.

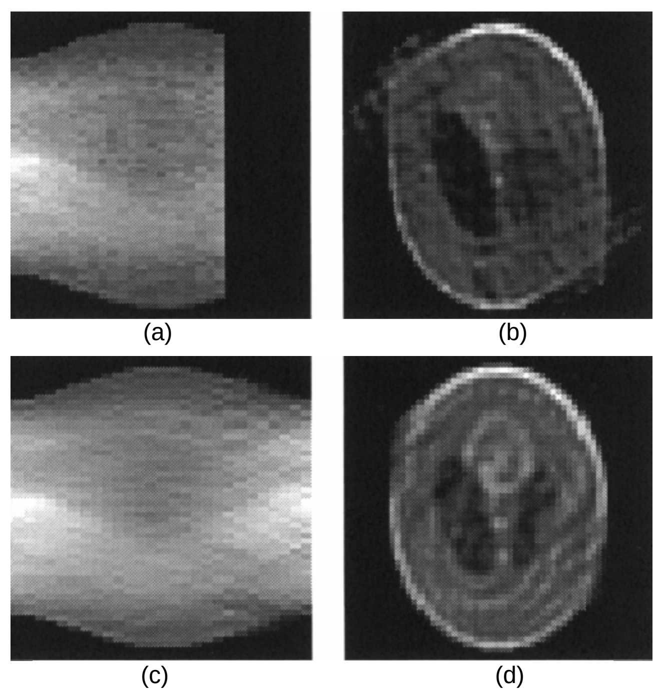

Figure 8. Experiment 3. Limited-angle restoration with SNR = 20 dB: (a) available sinogram; (b) "naive" reconstruction; (c) sinogram restored by ISRA; and (d) reconstruction from restored sinogram.

To simulate a noisy limited-angle sinogram, we first add independent zero-mean Gaussian noise with variance σ^2 to each element of the true sinogram. The resulting observed sinogram has a signal-to-noise ratio (SNR) defined as

$$\mathrm{SNR} = 10 \log_{10} \frac{\frac{1}{N_p} \sum_{i,j} (x(i,j) - \eta)^2}{\sigma^2}, \tag{60}$$

where $x(i,j)$ is the noise-free sinogram, η is its mean value, and $N_\rho = 2PM_\rho$ is the number of samples available.

Experiment 3. Noisy Limited-Angle Sinogram. This experiment simulates the case when the sinogram in Fig. 5(c) of Experiment 1 is corrupted by noise. Figure 8(a) displays the noise-corrupted limited-angle sinogram with SNR = 20 dB, and Fig. 8(b) shows the "naive" reconstruction, which suffers from severe artifacts. After ISRA is applied to the observation, a smooth sinogram is restored and the missing part is recovered, as shown in Fig. 8(c). The image reconstructed from the restored sinogram is shown in Fig. 8(d), in which all features are obviously better than in the "naive" reconstruction in Fig. 8(b). As a quantitative measure, the example has relative errors $\mathrm{RE_S}$ = 6.07% and $\mathrm{RE_O}$ = 16.97%.

In the presence of noise, there exists an optimal value of the regularization parameter λ_l (see, for example, Ref. 24), to achieve best results. Alternatively, if a suboptimal value of λ_l is used, there exists an optimal number of iterations for best result. Note however, that the minimum of the cost function is not achieved at this optimal number of iterations. Unfortunately, as discussed in the previous section, we are not aware of any systematic method to determine the optimal value of λ_l. The value of λ_l (=0.6) used in this experiment has been determined by trial and error to achieve the best result.

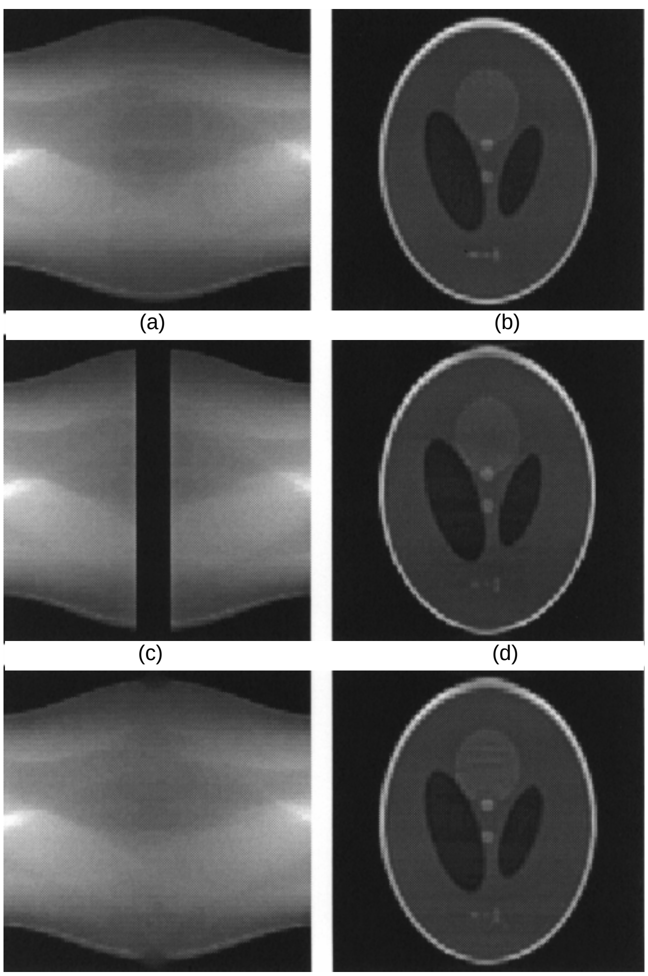

Figure 9. Experiment 4. Result obtained by using ISRA: (a) original sinogram; (b) original image; (c) available sinogram; (d) "naive" reconstruction; (e) sinogram restored by ISRA; and (f) image reconstructed from restored sinogram.

Experiment 4. Comparison with Other Methods. Oskoui and Stark[12] made a comparative study of CPL, ATM, and POCS reconstruction methods for the limited-angle CT problem. Simulation was performed to reconstruct a head phantom image from a set of limited-angle projections spanning a 160° angular range using the different techniques. To compare the performance of ISRA with these techniques, ISRA was applied to the same limited-angle problem as that used by Oskoui and Stark.

In Ref. 12, Oskoui and Stark used a 129 × 129-pixel Shepp–Logan phantom as the testing object [Figs. 9(a) and (b)]. Figure 9(c) shows a sinogram with 99 view angles available over the 160° range, with 129 raysums per view. The "naive" reconstruction is shown in Fig. 9(d). From the available sinogram in Fig. 9(c), we restored a complete sinogram, as shown in Fig. 9(e), which has 116 view angles over a 180° range at 100 raysums per view. On this sampling pattern, the CBP reconstruction is shown in Fig. 9(f). The improvement in image quality is not visually significant, but the sinogram is evidently recovered in the missing region. Because ISRA obtains an image with a smaller number of pixels than the other three methods, our result is compared with a full-angle image of the same pixel size. Table II shows the relative errors of the image by using CLP, ATM, POCS, and ISRA methods. It can be seen that ISRA performs better than CLP and ATM, and it is comparable to POCS. Because POCS utilized additional information on the object (which included an amplitude limit constraint and an energy constraint), its performance is expected to be better than that of the ISRA.

Conclusion

We have developed a sinogram space restoration method for the limited-angle problem. The algorithm uses the spectral consistency of the sinogram and closeness to observed data to restore a complete sinogram, which is then used to reconstruct an object via CBP. A computationally efficient iterative algorithm was developed to solve the optimization problem. The proposed method can be regarded as a variation of the POCS method with all iterations performed in the sinogram domain. Thus, the algorithm is much more computationally efficient, because computationally expensive reconstruction and numerical tomography projection required by conventional POCS are eliminated. Unlike the conventional POCS method, the proposed method does not use any unknown a priori information on the underlying object. We have also proposed the idea of restoring a complete sinogram on a sampling lattice with lower resolution than the original sampling grid, thereby increasing the robustness of the ISRA in the presence of noise. Experimental results show significant improvement in the quality of the reconstructed images from restored sinograms.

Appendix A. Proof of Theorem 1

Consider a system of equations $\boldsymbol{Hx} = \boldsymbol{y}$ where $\boldsymbol{H} \in \mathbf{C}^{M\times N}$ and $\boldsymbol{y} \in \mathbf{C}^{M\times 1}$ are given. If $M > N$, the least-squares solution for $\boldsymbol{x} \in \boldsymbol{R}^{N\times 1}$ is

$$\boldsymbol{x} = \mathrm{Re}\{(\boldsymbol{H}^*\boldsymbol{H})^{-1}\boldsymbol{H}^*\boldsymbol{y}\}. \tag{A1}$$

Because Eq. 34 can be transformed into vector form,[14]

$$\mathrm{vec}(\hat{\boldsymbol{X}}) = \arg\min_{\boldsymbol{X}} \left\| (\overline{\boldsymbol{B}} \otimes \boldsymbol{A})\mathrm{vec}(\boldsymbol{X}) - \mathrm{vec}(\boldsymbol{Y}) \right\|_F^2 \tag{A2}$$

and $(\overline{\boldsymbol{B}} \otimes \boldsymbol{A})$ is a full-rank matrix, the least-squares solution is

Table II. Numerical Results for the Four Methods in the 160° Limited-Angle Case in Experiment 4

	RE_S (whole image)	RE_O (zeroing background)
CPL	31.1%	22.6%
ATM	35.0%	26.7%
POCS	7.4%	6.0%
ISRA	9.8%	8.6%

$$\begin{aligned}\mathrm{vec}(\hat{X}) &= \mathrm{Re}\{((\bar{B}\otimes A)^*(\bar{B}\otimes A))^{-1}(\bar{B}\otimes A)^*\mathrm{vec}(Y)\}\\ &= \mathrm{Re}\{((B^T\bar{B})\otimes(A^*A))^{-1}(B^T\otimes A^*)\mathrm{vec}(Y)\}\\ &= \mathrm{Re}\{((B^T\bar{B})^{-1}B^T)\otimes((A^*A)^{-1}A^*)\mathrm{vec}(Y)\}.\end{aligned} \tag{A3}$$

Then we can rearrange Eq. A3 into matrix form:

$$\begin{aligned}(\hat{X}) &= \mathrm{Re}\{(A^*A)^{-1}A^*Y((B^T\bar{B})^{-1}B^T)^T\}\\ &= \mathrm{Re}\{(A^*A)^{-1}A^*YB(B^*B)^{-1}\}.\end{aligned} \tag{A4}$$

Appendix B. Proof of Theorem 2

By inspection, it can be seen that for *any* fixed X, the corresponding R that *minimizes* $\|AXB^*-Y-W\odot R\|_F^2$ is given by

$$R = (AXB^*-Y)\odot W. \tag{B1}$$

Hence,

$$R^+ = (AX^+B^*-Y)\odot W \tag{B2}$$

and

$$\begin{aligned}&\min_{X,R}\|AXB^*-Y-W\odot R\|_F^2\\ &= \|AX^+B^*-Y-W\odot R^+\|_F^2\\ &= \|(1-W)\odot(AX^+B^*-Y\|_F^2\\ &= J(X^+)\\ &\geq \min_X J(X),\end{aligned} \tag{B3}$$

where $J(\bullet)$ is as defined in Eq. 38 and the second equality is obtained by using Eq. B2. This leads to the result

$$J(\hat{X}) = \min_X J(X) \leq \min_{X,R}\|AXB^*-Y-W\odot R\|_F^2. \tag{B4}$$

On the other hand, by letting

$$\hat{R} = (A\hat{X}B^*-Y)\odot W, \tag{B5}$$

we have

$$\begin{aligned}&J(\hat{X})\\ &= \|(A\hat{X}B^*-Y-W\odot\hat{R}\|_F^2\\ &\geq \min_{X,R}\|AXB^*-Y-W\odot R\|_F^2.\end{aligned} \tag{B6}$$

Combining Eqs. B4 and B6, we see that

$$\min_{X,R}\|AXB^*-Y-W\odot R\|_F^2 = J(\hat{X}) \tag{B7}$$

and hence $X^+ = \hat{X}, R^+ = (A\hat{X}B^*-Y)\odot W$ is *a* solution to Eq. 39. To prove that $X^+ = \hat{X}$ is the *unique* minimum point for Eq. 39, suppose that $\{\tilde{X},\tilde{R}\}$ is another solution such that $\hat{X}\neq\tilde{X}$. Then it is clear that

$$\|A\hat{X}B^*-Y-W\odot R^+\|_F^2 = \|A\tilde{X}B^*-Y-W\odot\tilde{R}\|_F^2, \tag{B8}$$

and by Eq. B1,

$$\tilde{R} = (A\tilde{X}B^*-Y)\odot W. \tag{B9}$$

Thus, substituting Eqs. B2 and B9 into Eq. B8, we have

$$J(\hat{X}) = J(\tilde{X}). \tag{B10}$$

However, because X is the *unique* global minimum of Eq. 30 by assumption, $\tilde{X} = \hat{X}$. This is a contradiction. Hence, X^+ is unique.

Appendix C. Proof of Theorem 3

By inspection, Eq. 39 is a quadratic programming problem in X and R. Hence Eqs. 43 and 44 are merely alternating projection formulas based on the principle of minimizing the function with respect to one variable while keeping the other fixed. Let

$$\hat{J}(k) = \|AX_kB^*-Y-W\odot R_k\|_F^2. \tag{C1}$$

Then it is easy to see that

$$\begin{aligned}\hat{J}(k) &= \|AX_kB^*-Y-W\odot R_k\|_F^2\\ &\geq \|AX_{k+1}B^*-Y-W\odot R_k\|_F^2\\ &\geq \|AX_{k+1}B^*-Y-W\odot R_{k+1}\|_F^2\\ &= \hat{J}(k+1),\end{aligned} \tag{C2}$$

where the first inequality is obtained by using Theorem 1 and Eq. 43, and the second inequality is obtained by using Eqs. B1 and 44. Thus, our algorithm is a descending algorithm.

Further, if $X_k \neq X^+$ and hence, by Eq. 44 $R_k \neq R^+$, then all inequalities in Eq. C2 are strict. Thus, by the global convergence theorem[22], $X_k \rightarrow X^+ = \hat{X}$.

References

1. G. T. Herman, *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography*, Academic Press, New York; 1980.
2. K. C. Tam, J. W. Eberhard, and. K. W. Mitchell, Incomplete-data CT image reconstructions in industrial applications, *IEEE Trans. Nucl. Sci.* **37**: 1490–1499 (1990).
3. B. P. Medoff, Image reconstruction from limited data: theory and applications in computerized tomography, in *Image Recovery: Theory and Applications*, H. Stark, Ed., Academic Press, Orlando, 1984, pp. 321–368.
4. A. Macovski, *Medical Imaging*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
5. D. A. Hayner and W. K. Jenkins, The missing cone problem in computer tomography, *Advances in Computer Vision and Image Processing*, **1**: 83–144 (1984).
6. J. J. Clark, M. R. Palmer, and P. D. Lawrence, A transformation method for the reconstruction of functions from nonuniformly spaced samples, *IEEE Trans. Acous. Speech Sig. Proc*, **ASSP-33**, (4): 1151–1165 (1985).

7. J. A. Reeds and L. A. Shepp, Limited angle reconstruction in tomography via squashing, *IEEE Trans. Med. Imaging* **MI-6**: 89–97 (1987).
8. M. I. Sezan and H. Stark, Tomographic image reconstruction from incomplete view data by convex projections and direct Fourier inversion, *IEEE Trans. Med. Imaging* **MI-3:** 91–98 (1984).
9. P. Oskoui-Fard and H. Stark, Tomographic image reconstruction using the theory of convex projections, *IEEE Trans. Med. Imaging* **MI-7**: 45–58 (1988).
10. D. C. Youla and H. Webb, Image restoration by the method of convex projections: Part 1—Theory, *IEEE Trans. Med. Imaging* **MI-1**: 81–94 (1982).
11. M. I. Sezan and H. Stark, Image restoration by the method of convex projections: Part 2—Applications and numerical results, *IEEE Trans. Med. Imaging* **MI-1**: 95–101 (1982).
12. P. Oskoui and H. Stark, A comparative study of three reconstruction methods for a limited view computer tomography problem, *IEEE Trans. on Medical Imaging* **8:** 43–49 (1989).
13. P. A. Rattey and A. G. Lindgren, Sampling the 2-D radon transform, *IEEE Trans. Acoustics, Speech, Signal Processing* **ASSP-29**: 994–1002 (1981).
14. A. Graham, *Kronecker products and matrix calculus with applications*, Ellis Horwood Ltd., London, 1981.
15. A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ (1989).
16. S. J. Reeves and R. M. Merserau, Optimal estimation of the regularization parameter and stabilizing functional for regularized image restoration, *Optical Engineering* **29**: 446–454 (1990).
17. S. Kaczmarz, Angenaherta auflosung von systemen linearer gleichungen, *Bull. Acad. Polon. Sci. Letters* **A35**: 355–357 (1937).
18. P. Eggermont, G. Herman, and A. Lent, Iterative algorithm for large partitioned linear systems with applications to image reconstruction, *Linear Algebra and Its Applications* **40**: (1981).
19. T. Huang, Restoring images with shift-varying degradation, in *Proc. SPIE/OSA, Image Processing,* 1976, pp. 149–151.
20. S. Kuo and R. J. Mammone, Resolution enhancement of tomographic images using the row action projection method, *IEEE Trans. Med. Imaging* 1 **MI-10**: 593–601 (991).
21. A. V. Fiacco and G. P. McCormick, *Nonlinear programming: sequential unconstrained minimization techniques,* Wiley, New York, 1968.
22. D. G. Luenberger, *Introduction to linear and nonlinear programming,* Addison Wesley, Reading, MA, 1984.
23. D. C. Youla, Mathematical theory of image restoration by the method of convex projections, in *Image Recovery: Theory and Applications,* H. Stark, Ed., Academic Press, Orlando, 1984, pp. 29–77.
24. K. Miller, Least squares methods for ill-posed problems with a prescribed bound, *J. Opt. Soc. Am.* **1**: 52–74 (1970).