Research Article

SART-Type Image Reconstruction from Overlapped Projections

Hengyong Yu,1,2,3 Changguo Ji,3 and Ge Wang2,3

1 Division of Radiologic Sciences, Department of Radiology, Wake Forest University Health Sciences, Winston-Salem, NC 27157, USA
2 Biomedical Imaging Division, VT-WFU School of Biomedical Engineering and Sciences, Wake Forest University Health Sciences, Winston-Salem, NC 27157, USA
3 Biomedical Imaging Division, VT-WFU School of Biomedical Engineering and Sciences, Virginia Tech, Blacksburg, VA 24061, USA

Correspondence should be addressed to Hengyong Yu, hengyong-yu@ieee.org

Received 7 May 2010; Accepted 29 July 2010

Copyright © 2011 Hengyong Yu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To maximize the time-integrated X-ray flux from multiple X-ray sources and shorten the data acquisition process, a promising way is to allow overlapped projections from multiple sources being simultaneously on without involving the source multiplexing technology. The most challenging task in this configuration is to perform image reconstruction effectively and efficiently from overlapped projections. Inspired by the single-source simultaneous algebraic reconstruction technique (SART), we hereby develop a multisource SART-type reconstruction algorithm regularized by a sparsity-oriented constraint in the soft-threshold filtering framework to reconstruct images from overlapped projections. Our numerical simulation results verify the correctness of the proposed algorithm and demonstrate the advantage of image reconstruction from overlapped projections.

1. Introduction

Since the first computed tomography (CT) scanner was made [1], all the commercial scanners have been employing the X-ray source with a single small focal spot, which can be modeled as a point source. In micro-CT and even nano-CT applications, the reduced X-ray focal spot size becomes increasingly a limiting factor to achieve contrast and temporal resolution targets. To address this issue, our group recently proposed to use a line-shaped X-ray source so that more photons can be generated in a given data acquisition interval [2]. In this context, the X-ray source can be modeled as a line-segment, which can be further discretely modeled as an array of points [3]. In single source CT scanners, contrast resolution is limited by the finite focal-spot size necessary to generate a sufficient number of X-ray photons, and temporal resolution is limited by the time taken to acquire sufficiently many projections over a full-scan or half-scan angular range. Since a line source covers a wide angular range per view, the number of photons is increased to radiate an object to be reconstructed. Therefore, use of a line-shaped X-ray source or a multiple source array is a candidate scheme to balance spatial, contrast and temporal resolution.

Interestingly, the technology of field-emission X-ray sources based on carbon nanotubes (CNT) is a recent invention with several intrinsic advantages over conventional X-ray tubes. To maximize the time-integrated X-ray flux from multiple sources and improve the signal-to-noise ratio (SNR), multiplexing was used, where multi-source are excited with different temporal modulations [4]. If two or more sources are simultaneously fired, X-ray photons reach the same detector together for any single measurement, and one would not be able to identify which photon comes from which source. To unmix the signals from various sources, several projections can be collected at different time instants for the same view [5]. Due to the limited readout speed and mixed signals, there seems little advantage to use multiplexing configuration in terms of contrast and temporal resolution. In collaboration with Dr. Otto Zhou's group at University of North Carolina, we are developing a novel multi-source micro-CT system. In our system, we plan to fire a multi-source system simultaneously and acquire multiple projections at any viewing angle on nonoverlapped segments of a shared detector array [6]. With recent developed compressive sampling (CS) techniques [7, 8], we are also working to improve temporal resolution and reduce radiation dose from a limited number of nonoverlapped projections [9].
developed a CS-based algorithm to solve this problem [3]. Assuming that the differences between measured and predicted projections from various source points have equal weights, these differences are then backprojected to different X-ray source points in the SART framework. However, this algorithm converges to the least square solution that is not necessarily the true image. In our simulation, the reconstructed images sometimes suffered from blurring [2]. Then, our group developed a CS-based algorithm to solve this problem [3]. The algorithm is implemented in a projection onto convex sets (POCS) framework and employed a steepest gradient searching strategy. Although this algorithm often converges to the true image, its convergence speed is rather slow. Because the SART framework has an excellent convergence behavior especially when the ordered subset (OS) format is applied, in this paper we will develop a SART-type algorithm for image reconstruction from overlapped projections.

The paper is organized as follows. In the next section, we will formulate the overlapped projection model, which is not a sum of line integrals and quite different from that in the single-source case. In the third section, we will design a SART-type reconstruction algorithm for image reconstruction from overlapped projections. In the fourth section, we will report numerical results. In the last section, we will discuss related issues and conclude the paper.

2. Imaging Model

2.1. Nonoverlapped Projection Model. For CT reconstruction, a two-dimensional digital image can be expressed as \( f = (f_i) \in \mathbb{R}^I \times \mathbb{R}^J \), where the index \( 1 \leq i \leq I \) and \( 1 \leq j \leq J \) are integers. Define

\[
\hat{f}_n = f_{i_j}, \quad n = (i - 1) \times J + j, \tag{1}
\]

with \( 1 \leq n \leq N \) and \( N = I \times J \), we have the image in a vector representation \( \mathbf{f} = [f_1, f_2, \ldots, f_N]^T \in \mathbb{R}^N \). In this paper, we will use both the signs \( f_{i_j} \) and \( f_n \) for convenience. Let \( \mathbf{g} = [g_1, g_2, \ldots, g_M]^T \in \mathbb{R}^M \) be a measured vector with \( M \) being the product of the number of projections and the number of detector elements. They are linked by the following linear system:

\[
\mathbf{g} = \mathbf{A} \mathbf{f}, \tag{2}
\]

where \( \mathbf{A} = (a_{m,n}) \in \mathbb{R}^M \times \mathbb{R}^N \) is a measurement matrix. In a typical fan-beam geometry, the \( n \)th pixel can be viewed as a rectangular region with a constant value \( f_n \), the \( m \)th measured datum \( g_m \) as an integral of partially covered pixel areas by a narrow beam from an X-ray source to a detector element which are weighted by the corresponding X-ray linear attenuation coefficients. Thus, the component \( a_{m,n} \) in (2) denotes the intersection area between the \( n \)th pixel and the \( m \)th fan-beam ray (Figure 1). While the whole matrix \( \mathbf{A} \) represents the forward projection, \( \mathbf{A}^T \) implements the backprojection.

\[2.2. Overlapped Projection Model.\] While the imaging model (2) is valid for a single-source system, it cannot be used for multi-source-generated projections. The reason is that the measured data in (2) has been postprocessed by a logarithmic operation. In other words, we must model the raw data directly. For a multi-source system with \( Q \) sources, assuming that the \( q \)th source emits \( I_q \) photons towards each detector element. According to Beer’s law, we have the following imaging model:

\[
\mathbf{p} = \sum_{q=1}^{Q} I_q e^{-\mathbf{A}^q}, \tag{3}
\]

where \( \mathbf{A}^q \) is the system matrix defined in Section 2.1 for the \( q \)th source, \( s \) a constant to normalize the area model (Figure 1) for the measurement matrix to the line integral model, and \( e^{-\mathbf{A}^q} \) is a vector whose element is the exponential function of the corresponding element of \(-s\mathbf{A}^q\). In practice, we can approximate \( s \) as the reciprocal of the average width of an X-ray path in an object to be reconstructed. If we assume that all the X-ray sources have the same intensities, \( I_1 = I_2 = \cdots = I_q = I_0 \), we have

\[
\hat{\mathbf{p}} = \frac{\mathbf{p}}{I_0} = \sum_{q=1}^{Q} e^{-\mathbf{A}^q}, \tag{4}
\]

where the constant \( s \) is absorbed by \( \hat{\mathbf{p}} = \mathbf{sf} \), and \( 1/I_0 \) absorbed by \( \hat{\mathbf{p}} = \hat{\mathbf{p}}/I_0 \). The key task in this paper is to reconstruct \( \mathbf{sf} \) from \( \hat{\mathbf{p}} \) and known measurement matrices \( \mathbf{A}^q \).

3. Reconstruction Algorithm

3.1. Single-Source Algorithm. In the past decade, our group studied a block-iterative (BI) or ordered-subset (OS) version of a general Landweber scheme [10], of which the SART and OS-SART [11] are special cases, for minimization of a weighted least square functional in the real/complex space, and proved its convergence under quite general conditions [12, 13]. The SART or OS-SART technology has been widely
used in the CT field. Particularly, for the system (2) the SART-type solution is given by [11]

\[
f_n^k = f_n^{k-1} + \lambda^k \frac{1}{a_{+n}} \sum_{m=1}^{M} a_{m,n} (g_m - A_m f_n^{k-1}),
\]

where \( a_{+n} = \sum_{m=1}^{M} a_{m,n} > 0 \), \( a_{m+} = \sum_{n=1}^{N} a_{m,n} > 0 \), \( A_m \) is the \( m \)-th row of \( A \), \( k \) the iteration index, and \( 0 < \lambda^k < 2 \) a free relaxation parameter. Let \( \Lambda^{+N} \in \mathbb{R}^N \times \mathbb{R}^N \) be a diagonal matrix with \( \Lambda_{n,n}^{+N} = 1/a_{+n} \) and \( \Lambda^{M+} \in \mathbb{R}^M \times \mathbb{R}^M \) be a diagonal matrix with \( \Lambda_{m,m}^{M+} = 1/a_{m+} \), (5) can be rewritten as

\[
\tilde{f}^k = f_n^{k-1} + \lambda^k \left( \Lambda^{+N} A^T \Lambda^{M+} (g - A f_n^{k-1}) \right).
\]

Note that for all the iteration index \( k \), \( \Lambda^{M+} \), \( \Lambda^{+N} \) and \( \Lambda \) remain unchanged.

3.2. Multisource Algorithm. Since (4) is a non-linear equation, there is no analytic solution to this problem. Here, we will find a solution in the SART framework summarized in the proceeding subsection. Our strategy is to linearize the equation and approximate the solution successively. Denote the approximate solution for (4) as \( \tilde{f}^k \) after \( k \) iterations, we have

\[
\tilde{p} = \sum_{q=1}^{Q} e^{-A^q \tilde{f}^k} = \sum_{q=1}^{Q} e^{-A^q \tilde{f}^k} = \sum_{q=1}^{Q} (e^{-A^q \tilde{f}^k} \odot e^{-A^q \tilde{f}^k}),
\]

where \( \Delta \tilde{p}^k \) is the error image, and “\( \odot \)” represents the component-wise multiplication of vectors of the same size. Then, \( e^{-A^q \Delta \tilde{p}^k} \) in (7) can be expanded in a Taylor series

\[
e^{-A^q \Delta \tilde{p}^k} = \sum_{l=0}^{\infty} \frac{(-1)^l}{l!} (A^q \Delta \tilde{p}^k)^l,
\]

where \( (A^q \Delta \tilde{p}^k)^l = (A^q \Delta \tilde{p}^k) \odot (A^q \Delta \tilde{p}^k) \cdots \). Substituting the 1st order approximation of \( e^{-A^q \Delta \tilde{p}^k} \) into (7), we have

\[
\tilde{p} \approx \sum_{q=1}^{Q} (e^{-A^q \tilde{f}^k} \odot (1 - A^q \Delta \tilde{f}^k))
\]

where \( \text{E}^{q,k} = (e^{-A^q \tilde{f}^k}) \odot (A^q \Delta \tilde{f}^k) \cdots \). Substituting the 1st order approximation of \( e^{-A^q \Delta \tilde{p}^k} \) into (7), we have

\[
\tilde{g}^k = B^k \Delta \tilde{f}^k,
\]

where

\[
\tilde{g}^k = \sum_{q=1}^{Q} \left( E^{q,k} A^q \right).
\]

Since the projection errors \( \text{E}^{q,k} \) from the involved sources are known, (10) can be understood as a linear system of \( \Delta \tilde{f}^k \) with a measurement matrix \( B^k \).

Because (10) has the same structure as that of (2), we can use the SART-type formula (6) to solve for \( \Delta \tilde{f}^k \). Let \( \Lambda^{+N,k} \in \mathbb{R}^N \times \mathbb{R}^N \) be a diagonal matrix with

\[
\Lambda^{+N,k}_{n,n} = \frac{1}{\sum_{m=1}^{M} b_{m,n}^{k}},
\]

and \( \Lambda^{M+,k} \in \mathbb{R}^M \times \mathbb{R}^M \) be a diagonal matrix with

\[
\Lambda^{M+,k}_{m,m} = \frac{1}{\sum_{n=1}^{N} b_{m,n}^{k}},
\]

we have the SART-type solution for \( \Delta \tilde{f}^k \)

\[
\Delta \tilde{f}^k = \Delta \tilde{f}^{k-1} + \lambda^{k,l} \left( \Lambda^{+N,k} B^k \right)^T \Lambda^{M+,k} \left( \text{E}^{q,k} - B^k \Delta \tilde{f}^{k-1} \right),
\]

where \( l \) is the iteration index and \( \lambda^{k,l} \) the relax parameter. Since

\[
(B^k)^T = \left( \sum_{q=1}^{Q} (E^{q,k} A^q) \right)^T = \sum_{q=1}^{Q} (E^{q,k} A^q)^T
\]

\[
= \sum_{q=1}^{Q} \left( (A^q)^T (E^{q,k})^T \right) = \sum_{q=1}^{Q} ((A^q)^T E^{q,k}),
\]
Equation (14) can be rewritten as

\[
\Delta \tilde{f}^{k,l} = \Delta \tilde{f}^{k,l-1} + \lambda^k \left( \Lambda^{+N,k} \left( \sum_{q=1}^{Q} \left( (A^q)^T E_q^k \right) \Lambda^{M+k} \left( g^k - B^k \Delta \tilde{f}^{k,l-1} \right) \right) \right) \\
= \Delta \tilde{f}^{k,l-1} + \lambda^k \left( \sum_{q=1}^{Q} \left( (A^q)^T g^k \Lambda^{M+k} \left( g^k - B^k \Delta \tilde{f}^{k,l-1} \right) \right) \right) \\
= \Delta \tilde{f}^{k,l-1} + \lambda^k \left( \sum_{q=1}^{Q} \left( (A^q)^T g^k \Lambda^{M+k} \left( g^k - B^k \Delta \tilde{f}^{k,l-1} \right) \right) \right) \\
= \Delta \tilde{f}^{k,l-1} + \lambda^k \left( \sum_{q=1}^{Q} \left( (A^q)^T g^k \Lambda^{M+k} \left( g^k - B^k \Delta \tilde{f}^{k,l-1} \right) \right) \right) \\
\]

(16)

Once we have an approximation solution \( \Delta \tilde{f}^{k,l} \) for \( \Delta \tilde{f}^k \) after \( L \) iterations, we can update the reconstructed image by

\[
\tilde{f}^{k+1} = \tilde{f}^k + \Delta \tilde{f}^{k,l}. 
\]

(17)

For example, we can choose \( \Delta \tilde{f}^{k,0} = 0 \) as the initial image and set \( L = 1 \) for one step iteration to approximate \( \Delta \tilde{f}^k \), which results in a simplified algorithm

\[
\tilde{f}^{k+1} = \tilde{f}^k + \lambda^k \left( \sum_{q=1}^{Q} \left( (A^q)^T g^k \Lambda^{M+k} \left( e^{-A^q \tilde{f}^k} \otimes \tilde{g}^k \right) \right) \right). 
\]

(18)

For numerical implementation, our SART-type reconstruction algorithm for image reconstruction from overlapped projections can be summarized in the following pseudocode:

(S.1.) initialize \( \tilde{f}^k \) and \( \lambda^k \) for \( k := 0 \);
(S.2.) compute \( \tilde{f}^k \), \( A^{+N,k} \) and \( A^{M+k} \);
(S.3.) initialize the error image \( \Delta \tilde{f}^{k,1} := 0 \);
(S.4.) for \( q = 1 \) to \( Q \) backproject the projection error;
  (S.4.1.) weight the projection error \( \tilde{g}^k \) by \( e^{-A^q \tilde{f}^k} \);
  (S.4.2.) weight the projection error \( \tilde{g}^k \) by \( A^{M+k} \);
  (S.4.3.) backproject the weighted projection error towards the \( q \)th X-ray source;
  (S.4.4.) weight the backprojected error image by \( A^{+N,k} \);
  (S.4.5.) add the backprojected error image to \( \Delta \tilde{f}^{k,1} \) by
\[
\Delta \tilde{f}^{k,1} := \Delta \tilde{f}^{k,1} + A^{+N,k} \Lambda^{M+k} \left( e^{-A^q \tilde{f}^k} \otimes \tilde{g}^k \right); 
\]
(S.5.) update the current estimated image by

\[
\tilde{f}^{k+1} := \tilde{f}^k + \lambda^k \Delta \tilde{f}^{k,1}; 
\]
(S.6.) set \( k := k + 1 \);
(S.7.) go to (S.2) until the convergence criteria are satisfied;
(S.8.) scale the reconstructed image \( \tilde{f}^k \) to obtain the final result \( \tilde{f}^* = \tilde{f}^k/s \).

In the above pseudo-code, (S.1) initializes the iteration index \( k \), the relax parameter \( \lambda \), and the initial image \( \tilde{f}^0 \). In our numerical simulation, we always set \( \lambda^k = 1 \) and \( \tilde{f}^0 = 0 \). The outer loop (S.2)-(S.7) solves for \( \tilde{f}^k \) successively.

(S.2) precomputes several important intermediate variables to update an reconstructed image in the iteration step \( k \). (S.3)-(S.4) computes the current error image for one step iteration according to (16). Because the backprojection operation for different X-ray sources in the inner loop (S.4) has the same structure, it can be implemented by calling a common procedure. (S.5) updates an reconstructed image. (S.6) updates the iteration index. In (S.7), the convergence criteria are checked. The stopping criteria for (S.7) can be the maximum iteration number is reached and/or the relative reconstruction error comes under a predefined threshold [14]. Finally, the reconstructed image \( \tilde{f}^k \) is scaled by dividing with the constant \( s \).

3.3. Sparsity Regularization. The conventional data acquisition is based on the Nyquist sampling theory, which states that for accurate reconstruction of a band-limited signal or image the sampling rate must be at least twice the highest frequency of the signal or image. However, the recently developed CS theory shows that a high-quality signal or image can be reconstructed from far fewer measurements than what is usually required by the Nyquist sampling theorem [7, 8]. In light of the work on solving the linear inverse problems with sparsity constraints by Daubechies et al. [14, 15], we recently adapted the single source SART to reconstruct an image from a limited number of projections subject to a sparsity constraint [16], and demonstrated that the sparsity constraint helped improve the quality of reconstructed images effectively and reduce the number of projections significantly. Using the same strategy described in our previous papers [16, 17], here we use the sparsity constraint to regularize the proposed multi-source SART algorithm. This can be done by adding a soft-threshold filtering step between (S.5) and (S.6) in the pseudo-code given in Section 3.2. Particularly, we have the following pseudo-code segment:

(SS.5.1.) compute the sparse transform;
(SS.5.2.) estimate the optimal threshold;
(SS.5.3.) perform the soft-threshold filtering;
(SS.5.4.) perform the inverse sparse transform.
In the above pseudo-code, the sparse transform in (SS.5.1) can be either any invertible lossless compressible transform such as wavelet transform [16] and Fourier transform or uninvertible transforms such as discrete gradient transform (DGT) and discrete difference transform (DDT) [17]. For an uninvertible transform, the inverse sparse transform in (SS.5.4.) is in terms of pseudo inversion as we performed for DGT and DDT [17]. (SS.5.2.) determines an optimal threshold automatically using the projected gradient method for fast convergence [14]. In fact, we can omit (SS.5.2.) and specify any fixed filtering threshold. However, both the convergence speed and final result depend on the choice of the threshold.

4. Numerical Simulation

To verify the proposed SART-type algorithm for image reconstruction from overlapped projections, we implemented it in Matlab on a PC, with the computationally intensive segments coded in C and linked via the MEX mechanism. As illustrated in Figure 2, we simulated a triple-source fan-beam micro-CT system. In the system, the source XS0 is rotated on a circular scanning locus of radius 120 mm. The object was a modified Shepp-Logan phantom in a compact support with a radius of 35 mm. We used an equi-distance detector array of length 120 mm. The detector was perpendicular to the direction from the origin to the X-ray source XS0 that is 40 mm from the system origin. The detector array consisted of 500 elements. On the line through the X-ray source XS0 and parallel to the detector, we put two sources XS1 and XS2 being 25 mm apart from XS0 on its right and left sides, respectively. With this triple-source configuration, we simulated single-source (only XS0 was fired), dual-source (XS1 and XS2 fired simultaneously), and triple-source cases (XS0, XS1 and XS2 fired simultaneously).

For each of our selected numbers of projections over a full-scan range, we first equiangularly acquired the corresponding projection dataset based on the aforementioned projection model in the single, dual, and triple source cases, respectively. Then, we reconstructed the images using our algorithm described in Section 3.2. In our simulation, the parameter $\lambda$ in the SART formula (18) was set to 1.0, and the stopping criterion was defined as reaching the maximum iteration number 5000. Because the Shepp-Logan phantom is a piecewise constant function, its DGT and DDT are sparse. Hence, we also employed the sparsity regularization in terms of total difference minimization [17] and the threshold for filtering was automatically computed using the projected gradient method [14]. Figure 3 shows the reconstructed 256 × 256 images from 9, 11, 13, and 15 projections, respectively. For real-world applications, measurement noise is unavoidable. To test the stability of the proposed algorithm against data noise, we repeated the above reconstructions from projections corrupted by Poisson noise, assuming $I_0 = 5 \times 10^4$ photons per detector element [18]. The results are in Figure 4, which indicate the stability of the proposed algorithm.

5. Discussions and Conclusion

In the CT field, the line integral model along an X-ray path has been widely used in consistency with Beer’s law. However, it does not reflect the divergence due to the combination of the finite detector size and the source focal spot. As shown in Figure 1, we have assumed an area model for the X-ray path and normalized it for the multi-source imaging model. Our area model treats the X-ray path as a narrow fan-beam from the X-ray point source to the detector element, and we believe that the area model works better than the line model. Note that the proposed algorithm views the projection procedure as a matrix transform, both the area and line models can be handled by our algorithm. In other words, the proposed algorithm is independent of the imaging model as long as it is linear or can be transformed into a linear one. Additionally, to simplify the derivation and demonstrate our idea, we have assumed that the photon numbers emitting from all the sources to each detector are the same. In fact, these numbers may be different, and can be easily incorporated into our algorithm.

As far as the convergence of the proposed algorithm is concerned, it should converge to the least square solution in the cases of either noise-free and noisy projections. The reason is that the proposed method is in the framework of the general Landweber scheme, whose convergence has been well studied under quite general conditions [12, 13]. When only a small number of projections are available, we can use some sparse constraints to steer the solution to the truth. However, the convergence speed of the current soft-threshold filtering technology is still slow although it has been accelerated using the projected gradient method [14]. In the future, we will employ more advanced techniques for a faster speed, which may include but not limited to optimizing the code, employing parallel computation, and developing new algorithms.
In conclusion, we have developed a SART-type algorithm for image reconstruction from overlapped projections. The algorithm has been verified and demonstrated in the numerical simulation. Our methodology has a potential to support more flexible designs of multi-source CT/micro-CT systems for better contrast and temporal resolution.

Acknowledgments

This work is partially supported by the NSF/MRI Program (CMMI-0923297) and NIH/NIBIB Grants (EB009275 and EB011785). The authors are grateful for constructive comments from the reviewers.
References


