Introduction: a brief overview of iterative algorithms in X-ray computed tomography

M. Soleimani and T. Pengpen

Engineering Tomography Lab (ETL), University of Bath, Bath, UK

This paper presents a brief overview of some basic iterative algorithms, and more sophisticated methods are presented in the research papers in this issue. A range of algebraic iterative algorithms are covered here including ART, SART and OS-SART. A major limitation of the traditional iterative methods is their computational time. The Krylov subspace based methods such as the conjugate gradients (CG) algorithm and its variants can be used to solve linear systems of equations arising from large-scale CT with possible implementation using modern high-performance computing tools. The overall aim of this theme issue is to stimulate international efforts to develop the next generation of X-ray computed tomography (CT) image reconstruction software.

1. Basic principle of computed tomography

X-ray computed tomography (CT) plays a key role in many scientific discoveries [1]. X-ray CT image reconstruction is an exciting mathematical and computational challenge. There is growing interest in iterative reconstruction algorithms mainly due to recent surge in computational power. In this paper, a brief overview of iterative algorithm in X-ray CT is presented [2–7].

When the monochromatic rays penetrate through a homogeneous object as shown in figure 1a, the ray attenuation can be explained by the Lambert–Beer Law as follows: \( I = I_0 \cdot e^{-\mu x} \), where \( I_0 \) and \( I \) are an incident photon intensity and an intensity of photons after having passed an object with thickness \( x \). The purpose of CT is to determine the linear attenuation coefficient \( \mu \) of the object and convert those values to a grey-scale image. The attenuation coefficient is measured in units per metre. Unfortunately, the body organs which photons pass through are not homogeneous tissues as shown in figure 1b. The Lambert–Beer Law must be modified to \( I = I_0 e^{-(\mu_1+\mu_2+\mu_3+\cdots+\mu_n)x} \). Moreover, the
Figure 1. (a) Attenuation of monochromatic rays running through a homogeneous object; (b) attenuation of monochromatic rays running through an inhomogeneous object with all the same thickness; and (c) attenuation of monochromatic rays running through an inhomogeneous object with different thicknesses. (d) Attenuation of polychromatic rays running through an inhomogeneous object with different thicknesses.

Human organs are not only inhomogeneous tissues but also the tissue thicknesses are different, as illustrated in figure 1c. In this model, the numbers of photons are calculated by

\[ I = I_0 \cdot e^{-\left(\mu_1x_1 + \mu_2x_2 + \mu_3x_3 + \cdots + \mu_nx_n\right)} \]  \hspace{1cm} (1.1)  

\[ I = I_0 \cdot e^{-\sum_{i=1}^{n} \mu_i x_i} \]  \hspace{1cm} (1.2)  

and

\[ I = I_0 \cdot e^{-\int_0^x \mu dx} \]  \hspace{1cm} (1.3)

For clinical CT scanners, an X-ray source produces an X-ray beam with a wide range of photon energies (polychromatic radiation). Therefore, equation (1.3) can be extended to

\[ I = \int_{E_0}^{E_{\text{max}}} I_0(E) \cdot e^{-\sum_{i=1}^{n} \mu_i(E)x_i} dE \]  \hspace{1cm} (1.4)
unfiltered-projection

filtered-projection

Figure 2. (a) X-ray projections, (b) back projection reconstruction and (c) unfiltered and filtered projections.

\[ I = \int_0^{E_{\text{max}}} I_0(E) \cdot e^{-\int_0^x \mu(E)dx} \, dE. \] (1.5)

The summation of the attenuations in the particular X-ray path can be defined as

\[ \ln \frac{I_0(E)}{I} = \int_0^X \mu(E) \, dx. \] (1.6)

The natural logarithm of the ratio of the incident photons and the outgoing photons (\(\ln(I_0/I)\)) is known as ray sums. A set of ray sums at a particular angle is a projection as shown in figure 2a.

A back projection algorithm is typically used for calculating the linear attenuation coefficients \(\mu\) in an image matrix of an object by projecting all projections backwards in the direction of each ray path. In the case of small numbers of projections, the reconstructed image would show a spiky artefact (figure 2b). This is also known as a star artefact. On the other hand, blurring around the edges of the images would be developed from spikiness with more projections used. In order to remove this artefact, a mathematical filtering process can be used. Digital filters are implemented to raw projection data and the results, called filtered projection, are back-projected to the image matrix. Next a brief overview of a number of iterative algorithms will be presented.

2. Algebraic reconstruction methods: ART, SART and OS-SART

An iterative reconstruction algorithm is a technique that uses the differences between the measured data and the calculated data to update an image. In this section, the approaches and definitions for SART and OS-SART in Jiang and Wang are described. For example, an image \(x\) is required which solves

\[ Ax = b, \] (2.1)

where \(x \in \mathbb{R}^N\) is the image value for \(N\) cells and \(b \in \mathbb{R}^M\) is measured data from \(M\) measurements. Each element \(A_{nm}\) of the weight matrix makes \(A \in \mathbb{R}^{M \times N}\) is the length of intersection of the \(n\)th ray with the \(m\)th cell, for example \(\tilde{a}_1\) is the weight matrix of the intersection between the 1st ray and the image matrix 2-by-2 as illustrated in figure 3. For iterative reconstruction, equation (2.3)
Figure 3. Five rays passing through a 2-by-2 image matrix (here, $M = 5$, $N = 4$, $A \in \mathbb{R}^{5 \times 4}$) (simplified illustration since in reality the rays are not necessarily parallel to the planes between cells).

can be solved by

$$x^{(k+1)} = x^{(k)} + \lambda_n \delta x^{(k)},$$

where $\delta x^{(k)}$ is an update, $\lambda_n$ is a relaxation parameter, $k$ is the iteration index and $n$ is the cycle index (one cycle is defined as all available datasets used in one iteration); let

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{R}^N. \quad (2.3)$$

Usually, the initial image values are defined as all zero, then $x^{(0)} = 0$. Also, let

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_M \end{bmatrix} \in \mathbb{R}^{M \times 1}. \quad (2.4)$$

and

$$A = \begin{bmatrix} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \tilde{a}_3^T \\ \vdots \\ \tilde{a}_M^T \end{bmatrix} \in \mathbb{R}^{M \times N}. \quad (2.5)$$

ART updates row-by-row so

$$x^{(k+1)} = x^{(k)} + \lambda_n \tilde{a}_i^T \left( b_i - \tilde{a}_i^T x^{(k)} \right) \tilde{a}_i, \quad (2.6)$$
where $i$ is the row index, and the relaxation parameter $\lambda_n$ is changed by the cycle index $n$. SART iteratively updates for the whole image as

$$x^{(k+1)} = x^{(k)} + \lambda_n V^{-1} A^T W (b - Ax^{(k)}), \quad (2.7)$$

where $V$ is diagonal with $j$th diagonal element

$$V_{jj} = \sum_{i=1}^{M} |A_{ij}|, \quad (2.8)$$

and $W$ is diagonal and with $i$th diagonal element

$$W_{ii} = \frac{1}{\sum_{j=1}^{N} |A_{ij}|}, \quad (2.9)$$

and $A_{ij}$ is the $(i,j)$th element of $A$. However, $b$ can also be divided into $B$ blocks where

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_B \end{bmatrix} \in \mathbb{R}^M \quad \text{and} \quad b_\beta \in \mathbb{R}^{M_\beta}, \beta \in [1,B], \quad (2.10)$$

and $M = \sum_{\beta=1}^{B} M_\beta$. Similarly $A$ can be divided into $B$ blocks where

$$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_\beta \end{bmatrix} \in \mathbb{R}^{M \times N} \quad \text{and} \quad A_\beta \in \mathbb{R}^{M_\beta \times N}. \quad (2.11)$$

The reconstruction method of divided blocks is the so-called ‘OS-SART’ which updates an image block-by-block using

$$x^{(k+1)} = x^{(k)} + \lambda_n V^{-1} A^T W_\beta (b_\beta - A_\beta x^{(k)}), \quad (2.12)$$

where $W_\beta$ is diagonal with $i$th diagonal element

$$[W_\beta]_{ii} = \frac{1}{\sum_{j=1}^{N} |A_{ij}(\beta)|}. \quad (2.13)$$

To counteract the potential divergence of reconstruction after some cycles, in this study a relaxation parameter $\lambda_n$ is scheduled corresponding to the cycle index $n(n \geq 1)$ as

$$\lambda_n = sf \times \frac{1}{n^\alpha}, \quad \text{where} \quad 0 < \alpha \leq 1, \quad (2.14)$$

and $sf$ is the scaling factor. Denote that $\alpha$ is a rate of change of the relaxation parameter. This guarantees the convergence of the algorithm.

### 3. Conjugate gradient least-squares algorithm

The conjugate gradient least squares (CGLS) is an iterative solution used in large sparse least-squares problem of $Ax = b$. The weight matrix ($A \in \mathbb{R}^{M \times N}$) is created using a forward projection programme. The matrix $A$ is a combination of the sub-matrix for each projection. In the case where the matrix $A$ is positive-definite, the matrix $A^T A$ is positive-definite for any matrix $A$. The iterative
method is terminated at \(i\) steps for \(k = 0, 1, 2, \ldots, i\). The residual \((r^k)\) at each step is computed by
\[
r^k = b - Ax^k. \tag{3.1}
\]

Starting with an initial approximation \((x^0)\), then, \(r^0 = b - Ax^0\), \(p^0 = A^T r^0\)
\[
\psi^k = \frac{\|A^T r^k\|^2}{\|Ap^k\|^2}, \tag{3.2}
\]
\[
x^{(k+1)} = x^k + \psi^k p^k, \tag{3.3}
\]
\[
r^{(k+1)} = r^k - \psi^k Ap^k, \tag{3.4}
\]
\[
\beta^k = \frac{\|A^T r^{k+1}\|^2}{\|A^T r^k\|^2} \tag{3.5}
\]
and
\[
p^{k+1} = A^T r^{k+1} + \beta^k p^k. \tag{3.6}
\]

Normally, an initial approximation is \(x^0 = 0\), then \(r^0 = b\) and \(p^0 = A^T b\), where \(p\) is conjugate vectors. In the case of the residual \(r\) being zero, it implies that the problem is solved. Otherwise, if the residual \(r\) is non-zero, the desired solution for \(r^k\) to be minimized, which can be monitored by the L2 norm of the \(r^k\) for each iteration. This can be achieved by updating the residual \(r\) into the problem iteratively. Moreover, regularized CGLS was also studied here for ill-posed problems.

4. Conclusion

A brief overview of a number of iterative algorithm for X-ray tomography is presented in this article. The hope is that the next generation of X-ray tomography algorithms to be further developed in the next few years are better adapted in real application of X-ray tomography devices.

References

5. Herman GT, Meyer LB. 1993 Algebraic reconstruction techniques can be made computationally efficient. *IEEE Trans. Med. Imaging* 12, 600–609. (doi:10.1109/42.241889)