Convergence of Iterative Method as Discretization of Continuous-Time Image Reconstruction System for Computed Tomography

Kiyoko Tateishi¹, Ken‘ichi Fujimoto² and Tetsuya Yoshinaga²

¹ Graduate School of Health Sciences and ² Institute of Health Biosciences, The University of Tokushima
3-18-15 Kuramoto, Tokushima 770-8509, Japan
E-mail: k-tateishi@blue.plaha.or.jp, [fujimoto,yoshinaga]@medsci.tokushima-u.ac.jp

Abstract We present a nonnegatively constrained iterative method formulated by discretizing nonlinear differential equations in a continuous-time image reconstruction (CIR) system for computed tomography (CT). The method of using the discretization has a simple structure, and is based on the continuous approach for an ill-posed inverse problem; therefore, we expect that the method of using the discretization produces better-quality images quickly and easily against the conventional methods. We give proof of the convergence of a desired solution in the discretized CIR system, theoretically. The theory is illustrated through experiments with a simulated phantom and projection data acquired from an X-ray CT scanner.

Keywords: computed tomography, image reconstruction, continuous approach, discretization, convergence of solution

1. Introduction

Iterative image reconstruction [1, 2] and filtered back-projection are the most common methods of image reconstruction for computed tomography (CT). The iterative methods have advantages over the filtered back-projection procedure, which is generally faster and can be implemented in hardware, in reducing artifacts under noisy or truncated projection data. Although the iterative methods are software-based and slower, much research [3, 4, 5] has been done on improving the iterative deblurring procedures because of the high quality of these reconstructions.

Among iterative reconstruction methods, the algebraic reconstruction technique (ART) [1] and ordered-subset expectation-maximization (OSEM) [3] method are well known for emission CT in nuclear medicine. Unfortunately, each method has some inappropriate behavior [6]. When there are no exact solutions to the CT inverse problem, ART may produce a negative solution and converge to not an isolated solution but a limit cycle. The OSEM without the subset-balanced condition, which is quite restrictive, usually fails to converge, even if there exist exact nonnegative solutions.

Some of the authors have proposed a method of solving the inverse problem in order to reconstruct tomographic images. The method consists of a block continuous-time image reconstruction (CIR) system [7, 8] described by a switched nonlinear system with a piecewise smooth vector field. Since a numerical integration to solve the differential equation in the large state space using, for example, the Range-Kutta method is slow, we have also proposed its implementation in an analog electronic circuit [9]. Although the hardware can yield fast image reconstructions, a software-based iterative method without such special customized hardware is also useful. Thus, in this paper, we present a nonnegatively constrained iterative method formulated by discretizing the nonlinear differential equation in the block CIR system. The method of using the Euler discretization has a simple structure, and is based on the continuous approach for an ill-posed inverse problem; therefore, we expect that it will produce good-quality images quickly and easily.

We succeeded in obtaining proof of the convergence of a desired solution in the discretized CIR system, theoretically. Namely, the sequence of state variables converges to a nonnegatively constrained solution of the inverse problem by minimizing the cost function. Then, the theory was demonstrated through numerical analyses for examples with simulated phantom data and projection data acquired from an X-ray CT scanner.

2. CIR System

The problem of CT is to reconstruct an image with pixel values \( x \in R^n \) using data acquired from projections \( y \in R^m \) and a projection operator \( A \in R^{m \times n} \), where \( I \) and \( J \) denote the number of projections and pixel values, respectively, and \( R^+ \) indicates the set of nonnegative real num-
bers. Since projection data are, in practice, noisy and inconsistent, we assume that the projections are described as a statistical model by the following algebraic system of equations.

\[ y = Ax + n \]  

(1)

Here, \( n \in \mathbb{R}^l \) is a noise vector. Hence we should treat it as an ill-posed problem, which means that its solution is not unique, does not exist, or does not depend continuously on the data.

For solving \( x \) in Eq. (1), we consider the optimization problem described as

\[ \min_{x \in \mathbb{R}^l} V(x(t)), \quad t \in \mathbb{R} \]

\[ V(x) := KL(x, e) \]  

(2)

where \( e \in \mathbb{R}^l \) is a locally unique solution to Eq. (1) without the noise term, and \( KL \) is the generalized Kullback–Leibler divergence defined by

\[ KL(x, e) = \sum_{j=1}^J e_j \log \frac{e_j}{x_j} + x_j - e_j \]  

(3)

To obtain a local minimum of the objective function \( V(x) \), some of us have proposed a continuous dynamical method [7, 8] as an initial value problem in the following form:

\[ \frac{dx}{dt} = XB_m (e) (z_m - B_m x) \]

(4)

\[ t - kT \in [m-1, m), \quad t \in \mathbb{R}, \]

\[ x(0) = x_0 \in \mathbb{R}^{l_+}_1 \]

for a series of times \( 0 = t_0 < t_1 < t_2 < \ldots < t_M = \tau \), nonnegative integer \( k \), and an initial state \( x_0 \in \mathbb{R}^{l_+}_1 \) with \( R^{l_+}_1 \) denoting the set of positive real numbers, where \( X := \text{diag}(x) \) indicates the diagonal matrix of order \( J \times J \) in which the corresponding diagonal elements are \( x_i \) and \( B_m \in \mathbb{R}^{l_j \times l_+} \) and \( z_m \in \mathbb{R}^{l_j} \) denote, respectively, a submatrix consisting of partial rows of \( A \) and a subvector of \( y \) with the same corresponding rows of \( B_m \), for \( m = 1, 2, \ldots, M \), such that there exists an elementary matrix \( P \) satisfying

\[ P \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_M \end{bmatrix} = A, \quad \text{and} \quad P \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_M \end{bmatrix} = y \]  

(5)

Equation (4) is a switched nonlinear system with a piecewise smooth vector field when \( M \geq 2 \) and is an autonomous system when \( M = 1 \).

The following can be proved theoretically [7, 8]. Let \( \phi(t, x_0) \) be a solution to Eq. (4) with initial value \( x_0 \). Then, in Eq. (4) on the state space \( \mathbb{R}^l \), when we choose \( x_0 \in \mathbb{R}^{l_+}_1 \), \( \phi(t, x_0) \) behaves as \( \mathbb{R}^{l_+}_1 \), for all \( t \in \mathbb{R} \). If there exists a locally unique equilibrium \( e \in \{0\} \), then \( V(\phi(t, x_0)) \) decreases along the solution \( \phi(t, x_0) \) in \( t \in \mathbb{R} \) through the positive initial state \( x_0 \in \mathbb{R}^{l_+}_1 \) at \( t = 0 \). Although the zero vector, which is unexpected for image reconstruction, is a trivial solution of Eq. (4), the equilibrium is locally unstable, i.e., none of the trajectories converge to it. Moreover, the point \( e \) is a common equilibrium for all subsystems in Eq. (4), and \( V(x) \) decreases in time along the solution. Namely, the equilibrium \( e \) in the switched system is uniformly asymptotically stable, which can be ensured by a common Lyapunov function for all switching modes under an arbitrary switching law.

3. Discretized CIR System

We present an iterative method by discretizing the CIR system in Eq. (4). We use a simple forward Euler method to approximate the solution of the differential equation in Eq. (4). Then we obtain

\[ x^{k+1} = g(x^k), \quad k = 0, 1, 2, \ldots, \quad x^0 = x_0 \]  

(6)

where the map \( g \) is defined by the composition of submaps:

\[ g = g_M \circ g_{M-1} \circ \cdots \circ g_1 \]  

(7)

with

\[ g_m(x) = x + \gamma_m \Phi(x) B_m (e) (z_m - B_m x) \]  

for \( m = 1, 2, \ldots, M \), where \( \gamma_m > 0 \) is a parameter, and \( \Phi(x) = X/\|X\|_x \). The corresponding step size for the discretization is \( \gamma_m/\|X\|_x \). This enables us to consider conditions on an upper bound of the parameter \( \gamma_m \) for the non-negativity and convergence of the solutions. Note that although the system in Eqs. (6)–(8) is formally derived on the basis of the method of numerical discretization, the purpose is to obtain not the exact reproduction of the solution to Eq. (4) but a fast convergence speed by designing an iterative algorithm. The Euler method is effective in this case, because its associated computational cost for each iteration.

4. Theory

Some theoretical results concerning the convergence of solutions in the discretized CIR system of Eqs. (6)–(8) are given in this section. Firstly, we show the property that any solution emanating from a positive initial state is positive, which means that the discretized CIR system does not produce images with unphysical negative pixel values.

**Proposition 1** Let \( \gamma_m \) satisfy \( 1 - \gamma_m B_{mj} (e) B_{mu} > 0 \) for \( m = 1, 2, \ldots, M \), and \( j = 1, 2, \ldots, J \), where \( u = (1, 1, \ldots, 1)^T \in \mathbb{R}^J \). If the system \( y = Ax \) is consistent or there exists a solution \( x \in \mathbb{R}^{l_+}_1 \), then the solution \( x^\theta \) with an initial state \( x^0 \in \mathbb{R}^{l_+}_1 \) for any \( k = 1, 2, \ldots, \), is in \( \mathbb{R}^{l_+}_1 \).

**Proof 1** For each \( k \), the \( j \)th element of \( x^{k+1} \)

\[ x^{k+1}_j = x^k_j \left( 1 + \frac{\gamma_m}{\|X\|_x} B_{mj} (e) (z_m - B_m x^k) \right) \]

\[ = x^k_j \left( 1 + \frac{\gamma_m}{\|X\|_x} B_{mj} (e - x^k) \right) \]
where \( B_{m}^{j} \) denotes the \( j \)-th column of \( B_{m} \), satisfies

\[
x_{j}^{k+1} = x_{j}^{k} \left( 1 - \gamma_{m} B_{m}^{j} B_{m} \right) x_{j}^{k} + \frac{\gamma_{m} B_{m}^{j} B_{m} x_{j}^{k}}{\| x_{j}^{k} \|_{2}} + \frac{\gamma_{m} B_{m}^{j} B_{m} e}{\| x_{j}^{k} \|_{2}}
\]

\[
\geq x_{j}^{k} \left( 1 - \gamma_{m} B_{m}^{j} B_{m} + \frac{\gamma_{m} B_{m}^{j} B_{m} e}{\| x_{j}^{k} \|_{2}} \right)
\]

\[
> x_{j}^{k} \frac{\gamma_{m}}{\| x_{j}^{k} \|_{2}} B_{m}^{j} B_{m} e
\]

\[
> 0
\]

Therefore, we obtain \( x^{k+1} \in R_{+}^{k} \) for any \( k \).

Next, we show a convergence property related to the subsystem in each \( m \) block, by letting \( L_{m} = \rho(B_{m}^{T} B_{m}) \) be the spectral radius, or the largest eigenvalue, of the matrix \( B_{m}^{T} B_{m} \) for \( m = 1, 2, \ldots, M \).

**Proposition 2** We assume \( 0 < \gamma_{m} < 2/L_{m} \) and \( 1 - \gamma_{m} B_{m}^{j} B_{m} > 0 \) for \( m = 1, 2, \ldots, M \) and \( j = 1, 2, \ldots, J \). If there exists a solution \( x = e \in R_{+}^{k} \) to the equation \( y = A x \), then

\[
\| e_{m} B_{m} x_{m} \|_{2}^{2} - \| e_{m} B_{m} x_{m}^{k+1} \|_{2}^{2} \geq 0
\]

**Proof 2** From Proposition 1, \( x^{k} \in R_{+}^{k} \) is satisfied, and according to the fact that the matrix \( \Phi_{x}^{k} \) is an \( L_{m} \) by \( L_{m} \) diagonal matrix in \( (0, 1) \) and non-diagonal element of zero, we have the following:

\[
\| e_{m} B_{m} x_{m} \|_{2}^{2} - \| e_{m} B_{m} x_{m}^{k+1} \|_{2}^{2} = \| e_{m} B_{m} x_{m} \|_{2}^{2} - \gamma_{m} B_{m} \Phi_{x}^{k} B_{m}^{T} \left( e_{m} B_{m} x_{m} - e_{m} B_{m} x_{m}^{k+1} \right) \|_{2}^{2}
\]

\[
= 2 \gamma_{m} \| B_{m} \Phi_{x}^{k} B_{m}^{T} \|_{2}^{2} - \| e_{m} B_{m} x_{m} \|_{2}^{2} - \gamma_{m} \| e_{m} B_{m} \Phi_{x}^{k} B_{m} \|_{2}^{2} - \| e_{m} B_{m} x_{m}^{k+1} \|_{2}^{2}
\]

Since

\[
\rho((B_{m} \Phi_{x}^{k} B_{m}^{T}) (B_{m} \Phi_{x}^{k} B_{m}^{T})) \leq L_{m}
\]

it follows that

\[
L_{m} \| B_{m} \Phi_{x}^{k} B_{m}^{T} \|_{2}^{2} - \| e_{m} B_{m} x_{m} \|_{2}^{2} \geq 0
\]

Therefore, we have

\[
\| e_{m} B_{m} x_{m} \|_{2}^{2} - \| e_{m} B_{m} x_{m}^{k+1} \|_{2}^{2} \geq 0
\]

**Corollary 1** When \( M = 1 \), under the assumption in Proposition 2, the sequence \( \{x^{k}\} \) converges to the solution of \( y = A x \) with decreasing \( \| y - A x \|_{2} \).

Thus, by selecting \( M = 1 \), we can obtain the convergence of the sequence minimizing \( \| y - A x \|_{2} \) subject to nonnegativity constraints for an appropriately chosen parameter \( \gamma = \gamma_{1} \) and a given positive initial state.

We should note that, for \( M = 1 \), it is difficult to find the largest eigenvalue \( \lambda_{1} \) of the matrix \( A^{T} A \), because \( A \) is large. However, upper bounds for \( \lambda_{1} \) can be obtained in terms of the degree of sparseness of matrix \( A \) as follows [10].

**Proposition 3 (Byrne [10])** Let \( A \) be an \( I \) by \( J \) matrix. For each \( m = 1, 2, \ldots, M \), let \( h_{m} = \sum_{j=1}^{J} A_{m,j}^{2} > 0 \). For each \( n = 1, 2, \ldots, N \), let \( \sigma_{n} = \sum_{m=1}^{M} e_{mn} \nu_{n} \), where \( e_{mn} = 1 \) if \( A_{mn} \neq 0 \) and \( e_{mn} = 0 \) otherwise. Let \( \sigma \) denote the maximum of \( \sigma_{n} \). Then the eigenvalues of matrix \( A^{T} A \) do not exceed \( \sigma \). If \( A \) is normalized so that the Euclidean length of each of its rows is one, then the eigenvalues of \( A^{T} A \) do not exceed \( \sigma \), the maximum number of nonzero elements in any column of \( A \).

5. Experiment

Experiments were examined in order to illustrate the theoretical results using two examples: a numerically simulated phantom and a projection acquired from an x-ray CT scanner.

The simulated phantom image \( e \) is made of 87 \times 87 pixels \((J = 7, 569)\) of the well-known Shepp-Logan phantom, each of which has a value between 0 and 1, as shown in Fig. 1(a). Reconstruction was done using 128 detectors per projection and 180-degree scanning with sampling every 5 degrees \((
\lambda = 4, 608)\). The projection data \( y = A e + n \) were created both with noise and with added Gaussian noise \( n = 0.03 W \) (the signal-noise ratio is around 15 dB), where \( W \) is the standard normal distribution random noise, and with truncation to positive values. The sinograms of the projections are shown in Fig. 1(b). In the sinogram, the horizontal axis corresponds to the projection angle and the vertical axis to the detector channel.

![Fig. 1 (a) Phantom and (b) sinograms without noise (left panel) and with noise (right panel)](image-url)

For the second example, we used X-ray CT projection data from a real multidetector-row CT scanner. We set the scanning conditions of tube voltage, tube current, exposure time, and slice thickness to 120 kV, 250 mA, 750 msec, and 1 mm, respectively. The sinogram is shown in Fig. 2(a). The numbers of projections and pixels of the
reconstructed image were $I = 28,710$ (957 bins and 30 directions) and $I = 454,276$ (674 × 674 pixels), respectively. Figure 2(b) shows an image reconstructed by the filtered back-projection procedure, which is a transform method.

Fig. 2 (a) Sinogram and (b) reconstructed image of X-ray CT by filtered back-projection

In our numerical experiments, we treated the iterative method in Eqs. (6)-(8) with $M = 1$, where each element of the initial value vector $x_0$ was set to 1. According to the conditions in Proposition 2, the values of $\gamma$ for the first and second examples respectively were fixed as 0.8 and $1.9 \times 10^{-5}$, which were obtained using the constants shown in Table 1. The convergence can be measured in terms of the distance norm defined as $d := |y - Ax|_2$. Figures 3(a) and 3(b) show graphs of the actual CPU time versus the distance for the first and second examples, respectively. We see that every distance decreases monotonically with iteration, even though the problem is ill-posed according to the projection noise.

| example | $L$ | $\max_j |A_j^t A u|$ |
|---------|-----|------------------------|
| first   | 1   | 1.1628                 |
| second  | $1.037 \times 10^5$ | $2.2675 \times 10^4$ |

The reconstructed images are shown in Fig. 4. No pixel value in the images was negative, and whitish broken-line artifacts did not appear, unlike those from the filtered back-projection method shown in Fig. 2 (b).

6. Conclusion

We proposed a nonnegatively constrained iterative method formulated by discretizing nonlinear differential equations in a CIR system, and gave theoretical proof of the nonnegativity and the convergence of the solutions. The theoretical results were demonstrated through experiments with a simulated phantom and projections acquired from an X-ray CT scanner.

Acknowledgment

This research was partially supported by KAKENHI (No. 21560449), and the Aihara Project, the FIRST program from JSPS, initiated by CSTP.

References


Kiyoko Tateishi received her Bachelor degree in health science from National Institution for Academic Degrees and University Evaluation in 2002. She has worked for St. Marianna University School of Medicine Hospital as a radiological technologist since 1995 and has been a Master’s student of Graduate School of Health Sciences in the University of Tokushima since 2011. Her research interests are methods of image reconstruction for computed tomography.

Ken’ichi Fujimoto received his Master degree in information science and engineering from Tokyo Institute of Technology in 1997, and his Ph.D. degree in system engineering from the University of Tokushima in 2000. He is currently an Assistant Professor at the University of Tokushima. His research interests are basic research on the dynamics of nonlinear dynamical systems and their applications to medical image engineering.

Tetsuya Yoshinaga received his Master degree from the University of Tokushima, Tokushima, Japan, in 1986, and his Ph.D. degree from Keio University, Yokohama, Japan, in 1992, all in electronic engineering. Presently, he is a Professor at the Institute of Health Biosciences, the University of Tokushima. His research interests are the qualitative theory of nonlinear dynamics, and its applications to medical imaging and bioengineering.

(Received May 12, 2012; revised July 17, 2012)