HIROSAKI UNIVERSITY

Radiation Environment and Medicine 2019 Vol.8, No.2 105–112

REM
Radiat Environ Med

Regular Article

Extended Ordered-subsets Expectation-maximization Algorithm with Power Exponent for Noise-robust Image Reconstruction in Computed Tomography

Yusaku Yamaguchi[1], Moe Kudo[2], Takeshi Kojima[3], Omar Mohammad Abou Al-Ola[4] and Tetsuya Yoshinaga[3]

[1]*National Hospital Organization, Shikoku Medical Center for Children and Adults, 2-1-1 Senyu, Zentsuji, Kagawa 765-8507, Japan*
[2]*Graduate School of Health Sciences, Tokushima University, 3-18-15 Kuramoto, Tokushima 770-8509, Japan*
[3]*Institute of Biomedical Sciences, Tokushima University, 3-18-15 Kuramoto, Tokushima 770-8509, Japan*
[4]*Faculty of Science, Tanta University, El-Giesh St., Tanta, Gharbia 31527, Egypt*

Received 21 February 2019; revised 25 May 2019; accepted 24 June 2019

The maximum-likelihood expectation-maximization (ML-EM) algorithm is the most popular iterative reconstruction method in emission-computed tomography with a noise model based on the Poisson distribution. The ordered-subsets EM (OS-EM) algorithm is known owing to accelerating the convergence of the ML-EM algorithm with the drawback of slow convergence. In this paper, we propose an extended OS-EM algorithm with a power exponent. We theoretically prove the asymptotic stability of an equilibrium corresponding to the solution of the nonlinear hybrid dynamical system whose numerical discretization based on multiplicative calculus coincides with the extended OS-EM algorithm. We provide a numerical experiment to demonstrate the effectiveness of the proposed system and confirm the acceleration of the proposed method and the robustness against noise. The reconstruction of high-quality images made by the method even when the projection data is noisy allows patient dose reduction in clinical practice.

Key words: Computed tomography, Iterative reconstruction, Ordered-subsets expectation-maximization algorithm, Continuous-time image reconstruction, Noise-robust image reconstruction, Dose reduction

1. Introduction

Iterative reconstruction (IR)[1, 2] is a method of obtaining tomographic images from projections in computed tomography (CT) and has advantages regarding the quality of reconstructed images over the filtered back-projection (FBP) procedure[3, 4]. The maximum-likelihood expectation-maximization (ML-EM) algorithm is the most popular IR method in emission CT with a noise model based on the Poisson distribution[5]. In addition, the ML-EM method is based on the principle of the optimization problem and has a feature that a relatively high quality image can be obtained even in the case when using a measured projection with a lower signal-to-noise ratio (SNR) or higher noise levels. However, the ML-EM algorithm has a disadvantage that it requires a huge amount of discrete computation. The ordered-subsets EM (OS-EM) algorithm[6], in which the projection is divided into subsets or blocks where the EM iteration is

*Yusaku Yamaguchi : National Hospital Organization, Shikoku Medical Center for Children and Adults, 2-1-1 Senyu, Zentsuji, Kagawa 765-8507, Japan
E-mail: yamaguchi.yusaku.sf@mail.hosp.go.jp

performed in each subset, is known owing to accelerating the convergence of the ML-EM algorithm with the drawback of slow convergence[6-9]. In the OS-EM method, to further accelerate the convergence rate, a larger power factor was introduced[10-14] while suppressing the divergence in the iterative process. It was shown that the power-based ML-EM algorithm can maximize the likelihood value as well as accelerate the convergence rate[13]. It is known that the ML-EM algorithm decreases an objective function with small iterative steps and increases it with large steps in the case of noisy projection. Therefore, it is necessary to terminate the iteration at an early step to obtain a solution with an image with sufficient quality, and an ML-EM algorithm with an exponent for controlling a noise model in forward projection has also been proposed[15]. For solving the tomographic inverse problem, we have presented a continuous-time image reconstruction (CIR) system[16-19], which is based on the approach of continuous-time dynamical optimization. The CIR system is described by a nonlinear differential equation with piecewise smooth vector fields, and the system is theoretically guaranteed to have asymptotic stability to the ideal solution while satisfying the pixel value constraint for the consistent inverse problem. Recently, we have also proposed a continuous analog to the power-based accelerated OS-EM[20] and an approach to derive an IR algorithm by discretizing the continuous-time dynamical system using additive or multiplicative calculus[21, 22]. The effect of the power factor for accelerating the OS-EM is easily understood by recognizing that it corresponds to the step size derived from the process via the discretization of the continuous analog[20].

In this paper, we propose an extended OS-EM algorithm with a power exponent. The parameter of the power exponent added to the original power-based OS-EM is not the power factor as a step size but the exponent associated with the ratio of the measured projection to the estimated forward projection. Note that when the value of the exponent is one, it corresponds to the conventional power-based OS-EM algorithm. For a consistent CT inverse problem, as a means to guarantee the convergence of the extended OS-EM algorithm to an exact solution, we use the dynamical system theory for a continuous analog of the iterative algorithm. Namely, we theoretically prove the asymptotic stability of an equilibrium corresponding to the solution of the nonlinear hybrid dynamical system whose numerical discretization based on multiplicative calculus coincides with that of the extended OS-EM algorithm. We provide a numerical experiment to demonstrate the effectiveness of the proposed system and confirm the acceleration of the OS-EM method and the robustness against noise.

2. System Description

Image reconstruction in CT is a problem to obtain the unknown variable $x \in R_+^J$ for pixel values satisfying

$$y = Ax, \tag{1}$$

where $y \in R_{++}^I$ and $A \in R_+^{I\times J}$ respectively denote the projection and a projection operator representing the discretization of the Radon transform (and R_+ and R_{++} respectively indicate the set of non-negative and positive real numbers). If the system in Eq. (1) has a non-negative solution, it is consistent; otherwise, it is inconsistent. The problem can be reduced to finding x using an optimization scheme that minimizes an objective function with respect to the system of Eq. (1).

Before describing the proposed system, we will provide a number of definitions and notations. Let $y^m \in R_+^{I^m}$ and $A^m \in R_+^{I^m\times J}$ be, respectively, a subvector consisting of I^m partical elements of y and a submatrix of A with the same corresponding rows of y^m for $m = 1, 2, \ldots, M$ with M indicating the total number of divisions. We also define

$$\lambda_j := \left(\sum_{i=1}^{I} A_{ij}\right)^{-1} \text{ and } \lambda_j^m := \left(\sum_{i=1}^{I^m} A_{ij}^m\right)^{-1}, \ j = 1, 2, \ldots, J,$$

where A_{ij} is the (i, j) element of A and $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_J)^\top$. To simplify the description, we denote vector-valued functions $\mathrm{Log}(\beta) := (\log(\beta_1), \log(\beta_2), \ldots, \log(\beta_L))^\top$ of each element in vector $\beta = (\beta_1, \beta_2, \ldots, \beta_L)^\top$ and $\mathrm{Exp}(\gamma) := (\exp(\gamma_1), \exp(\gamma_2), \ldots, \exp(\gamma_L))^\top$ of $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_L)^\top$, respectively. The objective function considered in this paper is expressed in the generalized Kullback–Leibler (KL) divergence[23]

$$\begin{aligned} \mathrm{KL}(\beta, \gamma) &= \sum_{\ell=1}^{L} \beta_\ell \log\frac{\beta_\ell}{\gamma_\ell} + \gamma_\ell - \beta_\ell \\ &= \beta^\top(\mathrm{Log}(\beta) - \mathrm{Log}(\gamma)) + v^\top(\gamma - \beta), \end{aligned} \tag{2}$$

for the given non-negative vectors β and γ, where v indicates an all-ones vector. The divergence KL (β, γ), known as the Csiszár's I-divergence measure, is non-negative with KL $(\beta, \gamma) = 0$ if and only if $\beta = \gamma$.

The proposed IR algorithm for solving the tomographic inverse problem is described by the iterative formula as follows.

$$\begin{aligned} z_j(n+1) &= z_j(n)\left(\lambda_j^m A_j^{m\top} \mathrm{Exp}\left(\alpha\left(\mathrm{Log}(y^m) - \mathrm{Log}(A^m z(n))\right)\right)\right)^{h^m} \\ &= z_j(n)\left(\lambda_j^m \sum_{i\in B^m} A_{ij}\left\{\frac{y_i}{(Az(n))_i}\right\}^\alpha\right)^{h^m}, \end{aligned} \tag{3}$$

$$n = 0, 1, 2, \ldots, m = (n \bmod M) + 1, z(0) = x^0,$$

with a parameter $\alpha > 0$, where the collection $\{B^1, B^2, \ldots, B^M\}$ is a partition or a subset of the index set $\{i = 1, 2, \ldots, I\}$; the definition is the same as the derivation of A^m and y^m for $m = 1, 2, \ldots, M$.

The proposed discrete system in Eq. (3) has a continuous analog as an initial value problem of the switched nonlinear dynamical system, which consists of a family of M subsystems

$$\frac{dx(t)}{dt} = X(t)\,\mathrm{Log}\left(\Lambda^m A^{m\top}\,\mathrm{Exp}\left(\alpha\left(\mathrm{Log}(y^m) - \mathrm{Log}(A^m x(t))\right)\right)\right), \quad (4)$$
$$t - k\tau \in [t^{m-1}, t^m), \quad t \geq 0, \quad x(0) = x^0 \in R^J_{++}$$

for $m = 1, 2, \cdots, M$, $k = 0, 1, \cdots$, and a series of times $0 = t^0 < t^1 < t^2 < \cdots < t^M = \tau$, where $X(t) := \mathrm{diag}\,(x(t))$ and $\Lambda := \mathrm{diag}(\lambda)$. Note that each subsystem is described by autonomous differential equations with a sufficiently smooth vector field. The solutions to the hybrid dynamical system are constructed through the connection of the last state of the previous mth subsystem and the initial state of the next $(m + 1)$th subsystem at every t^m, for $m = 1, 2, \ldots, M$, and $M + 1$ is set to be 1 for the cyclic switching process.

An integration of the continuous-time system in Eq. (4) by discretization via the multiplicative calculus reduces it to the IR algorithm in Eq. (3). The following shows the geometric multiplicative first-order expansion of a nonlinear function that appeared in the vector field. We describe the jth equation for the mth subsystem in Eq. (4) as

$$\frac{dx_j}{dt}(t) = x_j(t) f_j^m(x(t))$$

at $t \in [t^{m-1} + k\tau, t^m + k\tau)$ for $m = 1, 2, \ldots, M$ and for non-negative integer k. When applying the multiplicative Euler method to the mth subsystem, we obtain a single step at a time:

$$x_j(t + h^m) = x_j(t)\exp\left(h^m f_j^m(x(t))\right),$$

where $h^m > 0$ denotes the step size depending on m. For the block continuous-time hybrid system, by choosing $t^m - t^{m-1} =: h^m$ and by connecting solutions to the subsystems at the discrete time t_n, where $t_0 = 0$ and

$$t_n = \sum_{\ell=1}^{n-kM} h^\ell + k\sum_{\ell=1}^{M} h^\ell$$

for $n = 1, 2, \ldots$, with k being the floor of n/M, we have the block-iterative form

$$x_j(t_{n+1}) = x_j(t_n)\exp\left(h^m f_j^m(x(t_n))\right), \quad (5)$$
$$n = 0, 1, 2, \ldots, m = (n \bmod M) + 1, x(t_0) = x^0.$$

Through this discretization procedure, the discrete-time t_n is identical to the switching time $t = t^m + k\tau$ with $\tau = \sum_{\ell=1}^{M} h^\ell$ in the hybrid dynamical system under the corresponding special switching signals. By applying the same procedure to the continuous-time system in Eq. (4) and substituting $x(t_n)$ in Eq. (5) with $z(n)$, we obtain the discrete-time system in Eq. (3). The step size h^m in Eqs. (5) and (3) corresponds to the scaling parameters. Note that because the step size is derived in the discretization procedure of the hybrid dynamical system, its value does not affect the theoretical results of the solutions for the continuous-time system in Eq. (4).

3. Theoretical Results

In this section, the theoretical results of the solutions to the hybrid dynamical system in Eq. (4) are given.

We first show that a solution can be made to stay positive. This property is preserved whether the inverse problem is consistent and indicates that the CIR system does not produce images with unphysical negative pixel values.

Proposition 1. *If the initial value x^0 of the switched dynamical system in Eq. (4) is chosen in R^J_{++}, its solution $\phi(t, x^0)$ behaves in R^J_{++} for all $t \geq 0$.*

Proof. The vector field of the jth element of the mth subsystem can be rewritten as $dx_j/dt = x_j f_j^m(x)$ with the function f_j^m. Therefore, $d\phi_j/dt \equiv 0$ holds for any j on the subspace satisfying $x_j = 0$, which means the subspace becomes invariant, and, on the basis of the uniqueness criteria of solutions to the initial value problem, no flows can pass through every invariant subspace. This leads to the proof.

Assuming that the individual subsystems have the common equilibrium e satisfying $y^m = A^m e$ for any $m = 1, 2, \ldots, M$, we can prove the existence of a Lyapunov function for all subsystems in Eq. (4), which guarantees that the corresponding switched system has a stable equilibrium e.

The OS-EM algorithm was introduced to reduce the computation time of ML-EM. However, even in the consistent case, the convergence proof[6] requires the values of λ_j^m to be independent of the subset m, which is referred to as the "subset balance." This condition is restrictive in practical applications. Byrne[9] presented a more general sufficient condition called "a generalized subset balance," meaning that λ_j^m is separable and positive values δ^m exist such that

$$\lambda_j^m = \delta^m \lambda_j, \quad j = 1, 2, \ldots, J, \quad (6)$$

where $m = 1, 2, \ldots, M$. Our convergence proof on the stability of an equilibrium observed in the switched system in Eq. (4) needs the same condition as Eq. (6).

One of the main results is a stability theorem for the continuous OS-EM system as follows.

Theorem 1. *If there exists a unique solution $e \in R^I_{++}$ to the system $y = Ax$ with the matrix A satisfying the condition of Eq. (6), the equilibrium e of the dynamical system in Eq. (4) with $\alpha \in (0, 1]$ is asymptotically stable.*

Proof. This follows from Lyapunov's stability theorem. We define a possible candidate for a Lyapunov function as a weighted KL divergence,

$$\begin{aligned} V(x) &= \sum_{j=1}^{J} \lambda_j^{-1} \mathrm{KL}(e_j, x_j) \\ &= \sum_{j=1}^{J} \lambda_j^{-1} \left(e_j \log \frac{e_j}{x_j} + x_j - e_j \right), \end{aligned} \tag{7}$$

which is positive definite and is well-defined via Proposition 1 when initial value x^0 is chosen in R^I_{++}. It can be written as

$$V(x) = \sum_{j=1}^{J} \lambda_j^{-1} \int_{e_j}^{x_j} \frac{s - e_j}{s} ds.$$

Using the concavity of the log function and Jensen's inequality, we then calculate its derivative with respect to the dynamical system in Eq. (4) with $\alpha \in (0, 1]$ as

$$\begin{aligned} \left.\frac{dV}{dt}(x)\right|_{(4)} &= \sum_{j=1}^{J} \frac{x_j - e_j}{\lambda_j x_j} \frac{dx_j}{dt} \\ &= -\sum_{j=1}^{J} \frac{e_j}{\lambda_j} \log\left(\lambda_j^m A_j^{m\top} \mathrm{Exp}(\alpha(\mathrm{Log}(y^m) - \mathrm{Log}(A^m x)))\right) \\ &\quad + \sum_{j=1}^{J} \frac{x_j}{\lambda_j} \log\left(\lambda_j^m A_j^{m\top} \mathrm{Exp}(\alpha(\mathrm{Log}(y^m) - \mathrm{Log}(A^m x)))\right) \\ &\le -\sum_{j=1}^{J} \frac{e_j}{\lambda_j} \left(\lambda_j^m A_j^{m\top} \alpha(\mathrm{Log}(y^m) - \mathrm{Log}(A^m x))\right) \\ &\quad + \sum_{j=1}^{J} \frac{x_j}{\lambda_j} \left(\lambda_j^m A_j^{m\top} \mathrm{Exp}(\alpha(\mathrm{Log}(y^m) - \mathrm{Log}(A^m x))) - 1\right) \\ &= -\delta^m \sum_{i=1}^{I^m} y_i^m \left(\alpha(\log(y_i^m) - \log((A^m x)_i))\right) \\ &\quad + \delta^m \sum_{i=1}^{I^m} (A^m x)_i \left(\exp(\alpha(\log(y_i^m) - \log((A^m x)_i))) - 1\right) \\ &\le -\delta^m \sum_{i=1}^{I^m} y_i^m \left(1 - \exp(\alpha(\log((A^m x)_i) - \log(y_i^m)))\right) \\ &\quad + \delta^m \sum_{i=1}^{I^m} (A^m x)_i \left(\exp(\alpha(\log(y_i^m) - \log((A^m x)_i))) - 1\right) \\ &= \delta^m \sum_{i=1}^{I^m} \{(A^m x)_i\}^{\alpha} \{y_i^m\}^{1-\alpha} + \{(A^m x)_i\}^{1-\alpha} \{y_i^m\}^{\alpha} - ((A^m x)_i + y_i^m) \\ &\le 0 \end{aligned} \tag{8}$$

where $m = 1, 2, \ldots, M$, and x is in R^I_{++}. In the last inequality, we used the fact that the Heinz mean is less than or equal to the geometric mean when $\alpha \le 1$. The derivative equals zero at $x = e \in R^I_{++}$. Consequently, the hybrid system consisting of the family of subsystems in Eq. (4) has a common Lyapunov function defined by Eq. (7), so the equilibrium e of the system is asymptotically stable under arbitrary switching signals.

The proof of Theorem 1 can be rephrased as follows. If the system $y = Ax$ has a unique solution $e \in R^I_{++}$, the objective function $V(x)$ in Eq. (7) decreases monotonically in time for the solution to the system in Eq. (4) with $x^0 \in R^I_{++}$.

We use another Lyapunov function as an objective function that has to be minimized.

Theorem 2. *If the system $y = Ax$ possesses a unique solution $e \in R^I_{++}$, the following objective function $W(x(t))$ decreases monotonically in time t for solutions of the system in Eq. (4) with $M = 1$ and $x^0 \in R^I_{++}$:*

$$W(x) = \frac{1}{1-\alpha} y^{\top} (v - \mathrm{Exp}((1-\alpha)(\mathrm{Log}(Ax) - \mathrm{Log}(y)))) + v^{\top}(Ax - y)$$

where v indicates an all-ones vector and α is positive and not equal to 1.

Proof. We have

$$\begin{aligned} \sum_{i=1}^{I} \int_{y_i}^{A_i x} 1 - \left(\frac{y_i}{w}\right)^{\alpha} dw &= \sum_{i=1}^{I} \left[w - \frac{1}{1-\alpha} y_i^{\alpha} w^{1-\alpha} \right]_{y_i}^{A_i x} \\ &= \sum_{i=1}^{I} \frac{1}{1-\alpha} y_i \left(1 - \left(\frac{A_i x}{y_i}\right)^{1-\alpha}\right) + A_i x - y_i \\ &= W(x) \end{aligned}$$

for $\alpha > 0$ and $\alpha \neq 1$. Then, we see that the function $W(x)$, which is a generalized Hellinger distance[24], is positive definite for $x \in R^I_{++}$. We can obtain the derivative with respect to the system in Eq. (4) with $M = 1$ using the inequalities $(s-1)/s \le \log(s) \le s - 1$ for $s > 0$ as follows:

$$\begin{aligned} \left.\frac{dW}{dt}(x)\right|_{(4)} &= (v - \mathrm{Exp}(\alpha(\mathrm{Log}(y) - \mathrm{Log}(Ax))))^{\top} \\ &\quad \cdot AX \, \mathrm{Log}\left(\Lambda A^{\top} \mathrm{Exp}(\alpha(\mathrm{Log}(y) - \mathrm{Log}(Ax)))\right) \\ &\le -\mathrm{Exp}(\alpha(\mathrm{Log}(y) - \mathrm{Log}(Ax)))^{\top} \\ &\quad \cdot AX\left(u - \mathrm{Exp}\left(-\mathrm{Log}\left(\Lambda A^{\top} \mathrm{Exp}(\alpha(\mathrm{Log}(y) - \mathrm{Log}(Ax)))\right)\right)\right) \\ &\quad + v^{\top} AX\left(\Lambda A^{\top} \mathrm{Exp}(\alpha(\mathrm{Log}(y) - \mathrm{Log}(Ax))) - u\right), \end{aligned} \tag{9}$$

where $u \in R^J$ and $v \in R^I$ indicate all-ones vectors. By putting

$$p := \mathrm{Exp}(\alpha(\mathrm{Log}(y) - \mathrm{Log}(Ax)))$$

for simplicity, Eq. (9) can be written as

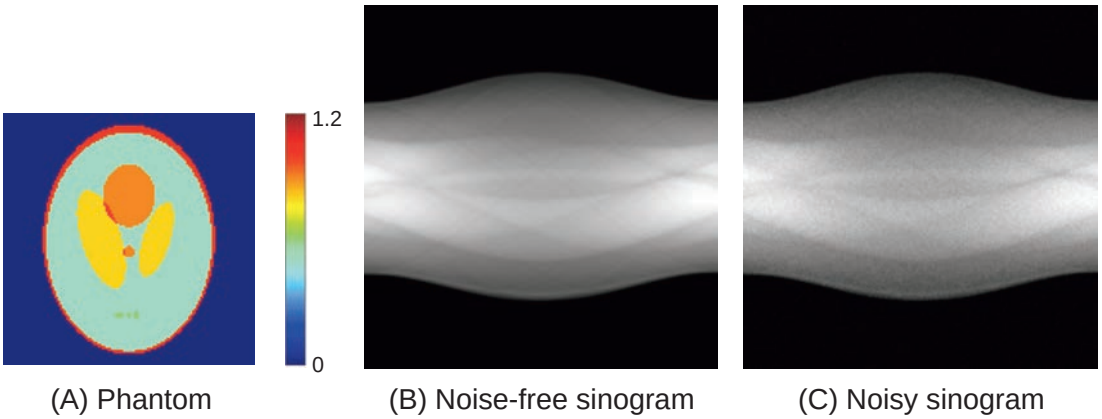


Fig. 1. Simulated phantom image and sinograms. The simulated phantom image shown in (A) is made of 128 × 28 pixels and consists of a range of pixel values between 0 and 1. (B) and (C) show noise-free and noisy sinograms, respectively, and sinogram (C) was created by adding white Gaussian noise such that the SNR is 30 dB.

$$\left.\frac{dW}{dt}(x)\right|_{(4)} \le -p^{\top}AXu + p^{\top}AX\ \mathrm{Exp}\left(-\mathrm{Log}\left(\Lambda A^{\top}p\right)\right) + v^{\top}AX\Lambda A^{\top}p - v^{\top}AXu. \tag{10}$$

For the first and third terms of the right-hand side in Eq. (10), one gets

$$\begin{aligned} -p^{\top}AXu + v^{\top}AX\Lambda A^{\top}p &= -p^{\top}AXu + p^{\top}A\Lambda XA^{\top}v \\ &= p^{\top}AX\ (-u + \Lambda A^{\top}v) \\ &= 0 \end{aligned}$$

and obtains for the second and fourth terms

$$\begin{aligned} p^{\top}AX\ \mathrm{Exp}\left(-\mathrm{Log}\left(\Lambda A^{\top}p\right)\right) - v^{\top}AXu &= u^{\top}X\Lambda^{-1}u - v^{\top}AXu \\ &= u^{\top}XA^{\top}v - v^{\top}AXu \\ &= 0. \end{aligned}$$

Therefore,

$$\left.\frac{dW}{dt}(x)\right|_{(4)} \le 0$$

and the Lyapunov function $W(x)$ decreases along the flow. This concludes the proof.

4. Numerical Experiment

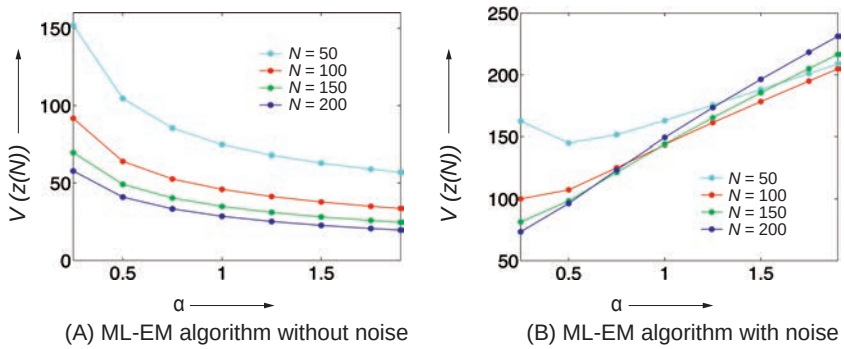
In this section, we present an experiment to illustrate the effectiveness of the proposed system in Eq. (3) with the power exponent α as a chosen parameter. The performance of image reconstruction by the proposed method was evaluated using a numerically simulated phantom.

Figure 1(A) shows the phantom $e \in R^{J}_{+}$ consisting of 128 × 128 pixels (J = 16,384), which has a range of pixel values between 0 and 1. The projection $y \in R^{I}_{++}$ (I = 33,300) was obtained by simulating a 180 degree scan at every 1 degree by 185 detectors per projection view and created by using the model $y = Ae + \sigma$ with $\sigma \in R^{I}$ denoting the white Gaussian noise such that the signal-to-noise ratio (SNR) was 30 dB unless otherwise specified. Noise-free and noisy sinograms are shown in Figures 1(B) and (C), respectively. In the experiment, we set the initial values to

$$z_j(0) = \frac{\sum_{i=1}^{I} y_i}{\sum_{i=1}^{I}\sum_{j'=1}^{J} A_{ij'}}, \quad j = 1, 2, \ldots, J. \tag{11}$$

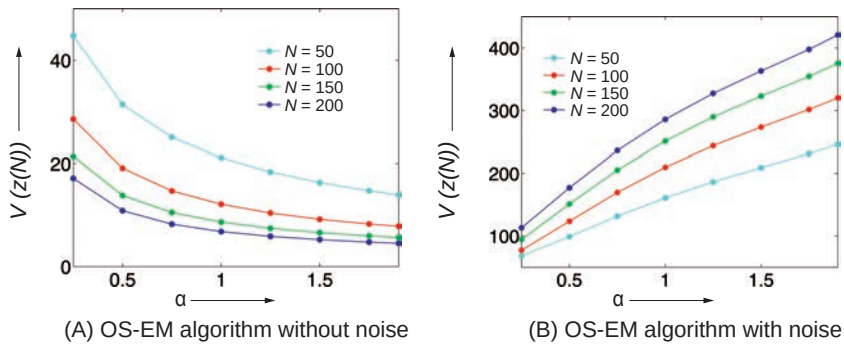
We used unblocked and blocked systems derived from Eq. (3) with M = 1 and M = 6, respectively, and the step size h^m is set to one for any m. We examined the case where the values of α in Eq. (3) were changed between 0.25 and 1.90. For evaluating the quality of reconstructed images and the property of convergence, we used the weighted KL divergence between the phantom e and the reconstructed image z as defined in Eq. (7). Since the KL divergence is a Lyapunov function of the proposed continuous-time system in Eq. (4), it is suitable for a function to evaluate the performance.

Figure 2 shows the values of the evaluation function $V(z(N))$ of the difference between the phantom image e and a reconstructed image $z(N)$ obtained by using the formula in Eq. (3) with M=1 at the Nth iteration starting from the common initial value $z_j(0)$ in Eq. (11) and variations in the parameter α, where the cyan, red, green, and blue curves indicate the results at the 50th, 100th,



(A) ML-EM algorithm without noise (B) ML-EM algorithm with noise

Fig. 2. Comparison of α and evaluation function $V(z)$ at each iteration in ML-EM algorithm. The figure shows the values of evaluation function $V(z)$ using the proposed algorithm in Eq. (3) with $M = 1$ at the Nth iteration while varying α with (A) noise-free and (B) noisy projections, where the cyan, red, green, and blue curves indicate the results at the 50th, 100th, 150th, and 200th iterations, respectively.



(A) OS-EM algorithm without noise (B) OS-EM algorithm with noise

Fig. 3. Comparison of α and evaluation function $V(z)$ at each iteration in OS-EM algorithm. The figure shows the values of evaluation function $V(z)$ using the proposed algorithm in Eq. (3) with $M = 6$ at the Nth iteration while varying α with (A) noise-free and (B) noisy projections, where the cyan, red, green, and blue curves indicate the results at the 50th, 100th, 150th, and 200th iterations, respectively.

150th, and 200th iterations, respectively. Similarly, Figure 3 shows the result for $M=6$. We observe that there exists an appropriate value of α that minimizes $V(z(N))$ at each number of iteration depending on the noise levels. In the simulation under noise-free projection data, we found that ML-EM ($M = 1$) and OS-EM ($M = 6$) give the same tendency that the evaluation $V(z(N))$ decreases as α increases, as shown in Figures 2(A) and 3(A). By contrast, for noisy projection, it was shown that the values of the evaluation $V(z(N))$ decrease with decreasing value of α. That is, we see that an effect of adopting the power exponent α is to vary the convergence rate according to the noise levels.

The values of $V(z(N))$ at $N = 100$ with variation of SNRs are shown in Figures 4(A) and (B) for ML-EM and OS-EM algorithms with α = 0.5, 1, and 1.5. Each graph makes it easy to find a suitable value of α depending on the noise levels of the projection data. That is, we see that the value of the evaluation function for each algorithm with $\alpha \neq 1$ becomes smaller than that of the conventional method ($\alpha = 1$) by setting α to be greater than 1 for higher SNR values and setting α to be less than 1 for lower SNR values; the values of around 40 to 50 dB are the boundary.

The reconstructed images are shown in Figures 5 and 6 to emphasize the property of Figure 4. Figures 5 and 6 present the reconstructed images using the sinograms shown in Figures 1(B) and (C), respectively, for $N = 100$ in the ML-EM and OS-EM algorithms with α = 0.5, α = 1, and α = 1.5. In Figure 5, it can be seen that a high-quality reconstructed image is obtained by increasing the value of α in a situation where there is no noise. It is clear that this tendency is more pronounced in the ML-EM method than in the OS-EM method with fast convergence. This means that the convergence rate is accelerated by increasing the value of α in a situation where there is low noise or no noise; as a result, it is possible to make the solution closer to the optimal solution. Furthermore, in the case where noise is added, as shown in Figure 6, we found that it is possible to suppress the noise component of the reconstructed image by reducing α. In contrast

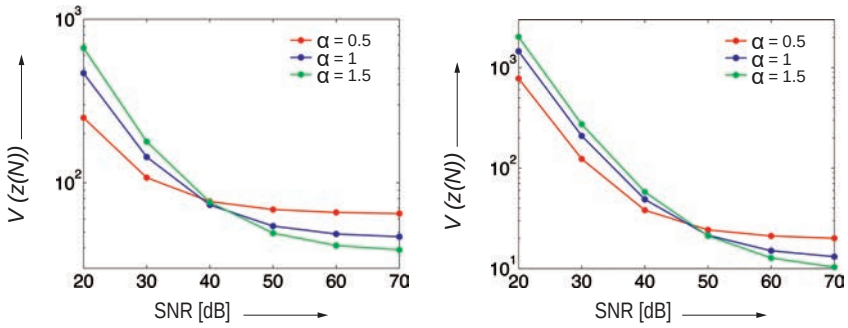


Fig. 4. Comparison of power-based OS-EM at $N = 100$ with power exponent α while varying SNR with (A) $M = 1$ and (B) $M = 6$, where red, blue, and green curves indicate $\alpha = 0.5$, $\alpha = 1$, and $\alpha = 1.5$, respectively.

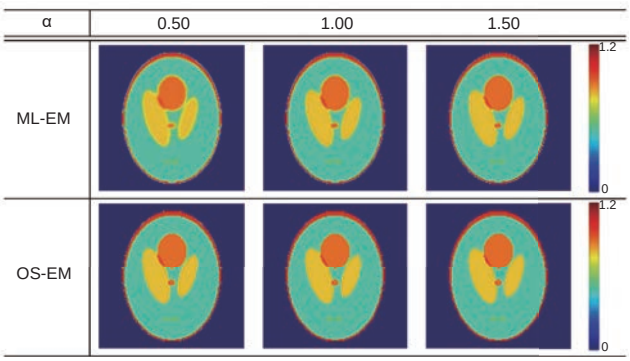


Fig. 5. Comparison of reconstructed images for noise-free projection. Each image indicates reconstructed images using noise-free projection data shown in Figs. 1(B) for $N = 100$ in the ML-EM and OS-EM algorithms with $\alpha = 0.5$, $\alpha = 1$, and $\alpha = 1.5$.

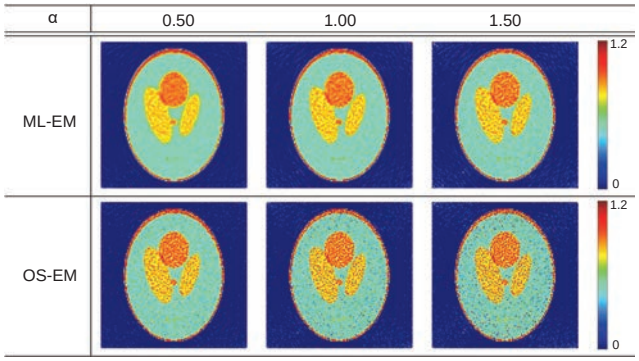


Fig. 6. Comparison of reconstructed images for noisy projection. Each image indicates reconstructed images using noisy projection data with an SNR of 30 dB, as shown in Figs. 1(C) for $N = 100$ in the ML-EM and OS-EM algorithms with $\alpha = 0.5$, $\alpha = 1$, and $\alpha = 1.5$.

to the noise-free situation, by decreasing the value of α, since the convergence rate is decelerated and the divergence of the solution is suppressed, the effect of suppressing the noise component can be seen. We have confirmed that the results shown above do not change even at large image sizes such as 512×512 pixels. In other words, it has been confirmed that the proposed method does not depend on the image size. These results suggest that the quality of the reconstructed image can be controlled by adjusting α according to the X-ray dose used for imaging in clinical applications. As a specific example of clinical application, it is considered that a noise-suppressed reconstructed image can be obtained by reducing the value of α when radiography is performed with a low dose to reduce the radiation exposure dose in X-ray CT. On the other hand, when radiography is performed with a sufficiently high dose, we consider clinical applications where higher quality reconstructed images can be obtained by increasing the value of α. Since our proposed method is an extended system that introduces the power exponent into the conventional OS-EM algorithm, we insist that it is a useful algorithm that does not have the drawbacks that would make it inferior to the conventional OS-EM algorithm.

5. Concluding remarks

We proposed an extended OS-EM algorithm with a power exponent α of the ratio of the measured projection to the estimated forward projection in each iterative step. For the continuous analog of the extended OS-EM, we theoretically prove the stability of an equilibrium by using Lyapunov's theorem with weighted KL divergence as a Lyapunov function for the case $\alpha \leq 1.0$. Moreover, for a continuous-time ML-EM system that is a non-ordered subset system with $M = 1$, it was also possible to prove the stability using the generalized Hellinger distance as the Lyapunov function by expanding the case where the value of exponent α is any positive number. We found that the characteristics of robustness for measurement

noise are shown by setting the exponent $\alpha < 1$ through numerical experiment. The use of the extended OS-EM algorithm with an appropriate exponent depending on the noise levels gives a better quality image. The reconstruction of high-quality images made by our method even for noisy projection data allows patient dose reduction in clinical practice.

Acknowledgements

This research was partially supported by JSPS KAKENHI Grant Number 18K04169.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

1. Gordon R, Bender R, Herman GT. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. J Theor Biol. 1970;29(3):471–81.
2. Badea C and Gordon R. Experiments with the nonlinear and chaotic behaviour of the multiplicative algebraic reconstruction technique (MART) algorithm for computed tomography. Phys Med Biol. 2004;49:1455–74.
3. Kak AC and Slaney M. Principles of computerized tomographic imaging. New Jersey: IEEE Press; 1988.
4. Stark H. Image recovery: Theory and applications. Florida: Academic Press; 1987.
5. Shepp LA and Vardi Y. Maximum likelihood reconstruction for emission tomography. IEEE Trans Med Imag. 1982;1(2):113–22.
6. Hudson HM and Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. IEEE Trans Med Imaging. 1994;13(4):601–9.
7. Byrne C. Block-iterative methods for image reconstruction from projections. IEEE Trans Image Process. 1996;5(5):792–94.
8. Byrne C. Accelerating the EMML algorithm and related iterative algorithms by rescaled block- iterative methods. IEEE Trans Image Process. 1998;7(1):100–9.
9. Byrne C. Block-iterative algorithms. Int Trans Oper Res. 2009;16(4):427–63.
10. Vishampayan S, Stamos J, Mayans R, Koral K, Clinthorne N, Rogers WL. Maximum likelihood image reconstruction for SPECT. J Nuc Med (Abs Book). 1985;26(5):20.
11. Tanaka E, Nohara N, Tomitani T, Yamamoto M. Utilization of non-negativity constraints in reconstruction of emission tomograms. In: Stephen LB, editor. Information Processing in Medical Imaging. Netherlands: Springer; 1986. vol.1: p.379–93.
12. Tanaka E. A fast reconstruction algorithm for stationary positron emission tomography based on a modified EM algorithm. IEEE Trans Med Imag. 1987;6(2):98–105.
13. Hwang D and Zeng GL. Convergence study of an accelerated ML-EM algorithm using bigger step size. Physics in Medicine and Biology. 2006;51(2):237–52.
14. Huang HM and Hsiao IT. Accelerating an ordered-subset low-dose X-ray cone beam computed tomography image reconstruction with a power factor and total variation minimization. PLoS ONE. 2016;11(4):e0153421.
15. Zeng GL. The ml-em algorithm is not optimal for poisson noise. In: 2015 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). IEEE; 2015. p.1–3.
16. Fujimoto K, Abou Al-Ola OM, Yoshinaga T. Continuous-time image reconstruction using differential equations for computed tomography. Commun Nonlinear Sci Numer Simul. 2010;15(6):1648–54.
17. Abou Al-Ola OM, Fujimoto K, Yoshinaga T. Common Lyapunov function based on Kullback–Leibler divergence for a switched nonlinear system. Mathematical Problems in Engineering. 2011; Article ID 723509.
18. Fujimoto K, Abou Al-Ola OM, Yoshinaga T. Implementation of continuous-time image reconstruction system in analog electronic circuit. Far East J Electron Commun. 2011;6(1):13–25.
19. Yamaguchi Y, Fujimoto K, Abou Al-Ola OM, Yoshinaga T. Continuous-time image reconstruction for binary tomography. Commun Nonlinear Sci Numer Simul. 2013;18(8):2081–7.
20. Tateishi K, Yamaguchi Y, Abou Al-Ola OM, Yoshinaga T. Continuous analog of accelerated os-em algorithm for computed tomography. Mathematical Problems in Engineering. 2017; Article ID 564123.
21. Aniszewska D. Multiplicative Runge–Kutta methods. Nonlinear Dyn. 2007;50(1–2):265–72.
22. Bashirov AE, Kurpinar EM, Oezyapici A. Multiplicative calculus and its applications. Journal of Mathematical Analysis and Applications. 2008;337(1):36–48.
23. Kullback S and Leibler RA. On information and sufficiency. Ann Math Stat. 1951;22(1):79–86.
24. Basu A, Shioya H, Park C. Statistical Inference: The Minimum Distance Approach. Chapman and Hall: CRC Press; 2011.