

Markov Chain Monte Carlo-Based L_1/L_2 Regularization and Its Applications in Low-Dose CT Denoising

Shuoqi Yu

School of Mathematics and Statistics, Shandong Normal University, Jinan, China

Email: 2192068511@qq.com

How to cite this paper: Yu, S.Q. (2025) Markov Chain Monte Carlo-Based L_1/L_2 Regularization and Its Applications in Low-Dose CT Denoising. *Journal of Applied Mathematics and Physics*, 13, 419-428.

<https://doi.org/10.4236/jamp.2025.132021>

Received: January 13, 2025

Accepted: February 14, 2025

Published: February 17, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this paper, a low-dose CT denoising method based on L_1/L_2 regularization method of Markov chain Monte Carlo is studied. Firstly, the mathematical model and regularization method of low-dose CT denoising are summarized, and then the theoretical basis of MCMC method and its application in image denoising are introduced. We evaluated the performance of various regularization strategies by comparing the denoising effects of L_1 , L_2 , and L_1/L_2 regularization terms in MCMC sampling at Gaussian noise levels. The experimental results show that L_1/L_2 regularization has the best performance in balancing noise removal and image detail retention, significantly superior to single L_1 and L_2 regularization, which proves its effectiveness for low-dose CT denoising.

Keywords

Low-Dose CT Denoising, Regularization, Statistical Inverse Problem, MCMC Sampling

1. Introduction

Low-dose CT imaging is designed to reduce the radiation dose to the patient by reducing the intensity of the X-rays to obtain CT images [1]. However, this dose reduction can significantly increase the noise in the image, affecting the image quality and diagnostic value [2]. Low-dose CT denoising technology is designed to remove noise from noisy CT images to improve image quality. The inverse problem of image denoising is generally regarded as an ill-posed problem [3], which is characterized by the non-uniqueness of the solution, instability and extreme sensitivity to the input data. For image denoising, small changes in noise can cause instability

in the recovery results. In order to solve the problem of noise reduction, regularization method is introduced to impose appropriate constraints in the process of solving, so as to improve the stability and reliability of the solution. Regularization avoids over-fitting noise by incorporating prior knowledge (such as signal smoothness, sparsity, or other statistical properties) and enhances the model's ability to recover the real image.

However, the traditional regularization method is prone to the accumulation of errors in the process of iterative solution, which leads to the problem of fuzzy boundary in the image after denoising.

Therefore, this paper proposes a low-dose CT denoising method based on L_1/L_2 regularization method of Markov chain Monte Carlo. By constructing a posterior probability distribution of low-dose CT image data with noise, regularization prior is incorporated to control the discomfort. Then, the posterior probability distribution is randomly sampled to ensure the image details while effectively denoising. This method effectively balances the effect of noise reduction and detail preservation in theory and shows significant advantages in practical application. Firstly, the mathematical model and regularization method of image denoising are introduced, and the implementation process of MCMC algorithm is described in detail, including initialization and sampling, and how to optimize the performance of the algorithm by adaptive step size.

Through numerical simulation experiment and real CT image experiment, we compare the denoising effect under different regularization strategies, using peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) as evaluation indexes. By comparing the experimental results of L_1 , L_2 and L_1/L_2 regularization based on Markov chain Monte Carlo, it is observed that L_1/L_2 regularization has the best denoising effect on low-dose CT, especially in preserving image structure and detail.

2. Regularization Model of Low-Dose CT Image Denoising

In the process of low-dose CT denoising, the main factor to be considered is that while removing noise, the integrity of the image itself must be ensured. $u \in \mathbb{R}^n$ is the original image, and the noisy image is $b \in \mathbb{R}^m$. The image denoising problem can generally be expressed as the inverse problem [4] of the following problems:

$$b = Iu + \eta \quad (1)$$

where η is noise and I is the identity matrix. Due to the existence of noise, the inverse problem shows discomfort, mainly reflected in the non-uniqueness and instability of the solution. Regularization is a common method to solve ill-determined problems, which can improve the quality and stability of the solution. The main idea is to add regularization term $\text{Reg}(u)$ to the prior distribution, and after adding regularization term, the denoised CT image is obtained by solving the following minimization problem:

$$u^* = \arg \min_u \left\{ \frac{1}{2} \|Au - b\|_{\ell_2}^2 + \alpha \text{Reg}(u) \right\} \quad (2)$$

where $\|Au - b\|_{\ell_2}^2$ is the fidelity term, $\text{Reg}(u)$ is the regularization term, and α is the regularization coefficient, used to balance the fidelity term and the regularization term.

L_1 regularization [5] and L_2 regularization on a given gradient:

$$\begin{aligned}\|\nabla u\|_1 &= \|\nabla_x u\|_1 + \|\nabla_y u\|_1 \\ \|\nabla u\|_2 &= \sqrt{\|\nabla_x u\|_2^2 + \|\nabla_y u\|_2^2}\end{aligned}$$

L_1/L_2 regularization term:

$$\frac{\|\nabla u\|_1}{\|\nabla u\|_2}$$

where ∇ is the gradient operator, let $D_x u, D_y u \in R^n$ be the first-order forward difference operators of the horizontal and vertical directions of the image respectively, then $D = [D_x u, D_y u] = \nabla$.

3. Solution of Low-Dose CT Denoising Problem and Sampling Method

3.1. Estimation of the Posterior Density Function

The prior density function corresponding to the regularization term is

$$\pi_{\text{pr}}(u) \propto e^{-\alpha \text{Reg}(u)},$$

suppose that the η is a Gaussian random vector with a mean of 0, and the covariance matrix Γ which is positive definite, *i.e.*

$$\eta \sim N(0, \Gamma).$$

The likelihood density function is

$$\pi(b | u) \propto \exp\left\{-\frac{1}{2}(Au - b)^T \Gamma^{-1} (Au - b)\right\}. \quad (3)$$

The posterior probability density of u obtained by Bayes' formula is

$$\pi(u | b) \propto \exp\left\{-\frac{1}{2}(Au - b)^T \Gamma^{-1} (Au - b) - \alpha \text{Reg}(u)\right\}. \quad (4)$$

Assuming that the noise covariance matrix is $\Gamma = \beta^2 I$, the standard form of the posterior probability density function of u is

$$\pi(u | b) \propto \exp\left\{-\frac{1}{2\beta^2} \|Au - b\|_{\ell_2}^2 - \alpha \text{Reg}(u)\right\}, \quad (5)$$

CM method is used to estimate the posterior density function, and the calculation formula of CM method is

$$u_{\text{CM}} = \int_{\mathbb{R}^n} u \pi(u | b) du. \quad (6)$$

3.2. MCMC Sampling Method

The MCMC algorithm is used to sample the posterior density function [6], and sample sequence $u^{(1)}, u^{(2)}, \dots, u^{(M)}$ is obtained. Suppose that the number of samples

in the pre-burning period of the probe posterior probability density function is m_0 , when the number of samples sampled is large enough, the remaining number after removing the samples in the pre-burning period is $M - m_0$, and the above integral can be approximated as the average of $M - m_0$ samples, that is

$$u_{\text{CM}} = \int_{\mathbb{R}^n} u \pi(u|b) du \approx \frac{1}{M - m_0} \sum_{m=m_0+1}^M u^{(m)}. \quad (7)$$

The MCMC sampling algorithm is used to sample the posterior density function, and a proposed distribution $Q(u'|u) = N(u, \sigma^2 I)$ is obtained, where u is the current state, u' is the proposed new state, σ^2 is the variance, I is the identity matrix, new state $u' = u + \delta \times \varepsilon$, $\varepsilon \sim N(0, I)$, the proposed distribution is symmetric, *i.e.* $Q(u'|u) = Q(u|u')$, since the jump probabilities from u to u' and from u' to u are the same. The adaptive step size adjustment is used to make the acceptance rate close to 0.234 [7]. Every 10,000 sampling times, the current acceptance rate is checked and the step size is adjusted accordingly. The rule of step size adjustment can be expressed as

$$\text{stepsize}_{\text{new}} = \begin{cases} \text{stepsize} \times 1.1, & \text{if acceptance rate} > 0.234 \\ \text{stepsize} \div 1.1, & \text{otherwise} \end{cases} \quad (8)$$

in practical applications, 1.1 is a common adjustment factor that provides a reasonable gradual adjustment amplitude (about 10% step change), without causing too drastic a change in sampling step size, and can quickly adapt to the characteristics of the state space, in the experiment, the adjustment factor 1.1 keeps the sampling acceptance rate in the range of 20% - 30%, which is close to the theoretical optimal value of 23.4%.

The acceptance rate is calculated as

$$\alpha = \min \left(1, \frac{Q(u'|b)Q(u|u')}{Q(u|b)Q(u'|u)} \right) = \min \left(1, \frac{Q(u'|b)}{Q(u|b)} \right). \quad (9)$$

Since the sampling method usually deals with the probability density in logarithmic form, in order to simplify the calculation, the posterior probability is logarithmic and set as

$$Q(u|b) = \log \pi(b|u) + \log \pi(u) + C, \quad (10)$$

where C is a normalized constant and is independent of u , so it can be ignored.

The adaptive step size can be dynamically adjusted according to the acceptance rate in the sampling process. When the acceptance rate is higher than the target value, increasing the step size can accelerate the speed of exploring the state space, and thus reach the equilibrium state faster. On the contrary, if the acceptance rate is too low, reducing the step size can improve the acceptance rate, ensure that the algorithm effectively explores the state space, and enable the algorithm to adjust to the scale suitable for the current target distribution more quickly, so as to accelerate the smooth distribution.

We summarize the MCMC:

Step 1: Set the initial value $u^{(1)} = \mathbf{b} = [b_1, b_2, \dots, b_n]$, $m = 1$, the sample of the

combustion period is m_0 , the total number of samples is M , and the initial step size is δ .

Step 2: Update candidate samples $u = u^{(m)} + \delta \cdot \varepsilon$, $\varepsilon \sim N(0, I)$.

Step 3: The logarithmic posterior probability distribution after the regularization term is

$$Q(u|b) = -\frac{1}{2\beta^2} \|Au - b\|^2 - \text{Reg}(u)$$

Calculate the receiving probability α as

$$\alpha = \min \left(1, \frac{Q(u|b)}{Q(u^{(m)}|b)} \right) = \min \left(1, \exp \left(Q(u|b) - Q(u^{(m)}|b) \right) \right).$$

Step 4: If $\alpha \geq t$, $t \sim U(0, 1)$, then $u^{(m+1)} = u$, otherwise reject candidate sample order $u^{(m+1)} = u^{(m)}$.

Step 5: When $m = M$, sampling stops. Otherwise, continue sampling and make $m = m + 1$. Repeat the second and third steps.

Step 6: After every 10,000 samples, adjust Step 1 according to the current acceptance rate δ

$$\delta = \begin{cases} \delta \times 1.1, & \alpha > 0.234 \\ \delta \div 1.1, & \text{otherwise} \end{cases}$$

go to Step 2 update δ .

Step 7: Based on the sampling results, the denoised image estimated by conditional mean (CM) is calculated

$$u_{\text{CM}} \approx \frac{1}{M - m_0} \sum_{m=m_0+1}^M u^{(m)}.$$

In order to compare the denoising effects of L_1 , L_2 regularization terms and L_1/L_2 regularization terms, and compare the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [8] of the denoised images, the higher the PSNR value is, the smaller the error between the denoised image and the original image, and the better the denoised effect is. The closer the SSIM value is to 1, the higher the similarity between the denoised image and the original image. The calculation of PSNR and SSIM is given below:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}_x^2}{\text{MSE}} \right) \quad (11)$$

where $\text{MSE} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (x(i, j) - y(i, j))^2$, where x is the original image, y is the denoised image, m and n are the size of the image, and MAX_x is the maximum possible pixel value of the image.

$$\text{SSIM} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (12)$$

where x represents the original image, y represents the denoised image, μ_x and

μ_x and μ_y are the mean values of the images x and y , σ_x^2 and σ_y^2 are the variances of the images x and y , σ_{xy} is the covariance of the images x and y , $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$ is a small constant to avoid denominators of zero, where L is the pixel value of the dynamic range, and $k_1 = 0.01$ and $k_2 = 0.03$ are the default values.

4. Experiment Results

4.1. Low-Dose Noise Image

In the study of low-dose CT image denoising algorithm, in order to objectively evaluate the effectiveness of the denoising algorithm, people often choose the recognized Sheep-Logan model as the research object.

Noise was added to the projection data of Sheep-Logan head model to mimic low-dose CT images and the noise approximately follows the non-stationary Gaussian distribution with the mean of 0. The noise variance formula [9] is as follows:

$$\sigma_i^2 = f_i \times \exp\left(\frac{\mu_i}{\varphi}\right) \tag{13}$$

where μ_i and σ_i^2 are the projected data mean and noise variance obtained on the i -th detector, f_i and φ are two configuration parameters, set here $f_i = 0.3$, $\varphi = 30$.

Inverse Radon transform is used to invert the projected data after noise, and the noise image is obtained. **Figure 1** shows the comparison between the original image and the noise image.

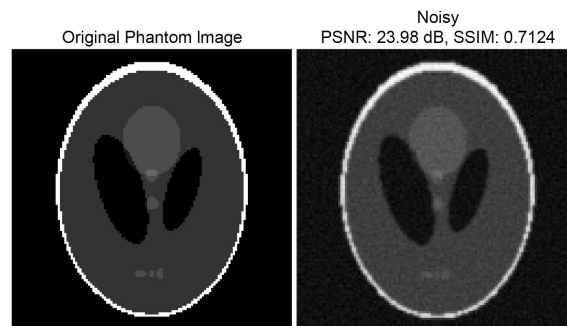


Figure 1. The left image is the original image without noise, and the right image is the noisy image.

4.2. Comparison of the Denoising Effect of L_1 , L_2 and L_1/L_2 Regularization Terms

Next, the regularization terms $\text{Reg}(u) = \|\nabla u\|_1$, $\text{Reg}(u) = \|\nabla u\|_2$ and

$\text{Reg}(u) = \frac{\|\nabla u\|_1}{\|\nabla u\|_2}$ are respectively used to carry out MCMC sampling on the projected data containing Gaussian noise, set the sampling to 400,000 times, $\beta = 0.05$,

$\alpha = 10^5$ and obtain three denoised images, as shown in **Figure 2**.

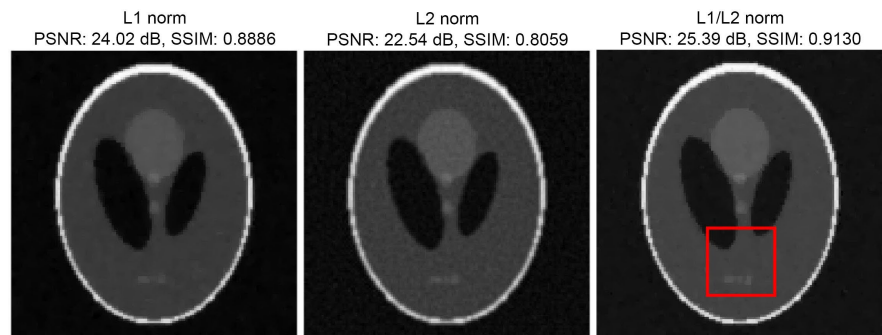


Figure 2. The first one is the denoised image using L_1 regularization, the second one is the denoised image using L_2 regularization, and the third one is the denoised image using L_1/L_2 regularization.

Table 1. Comparison of PSNR and SSIM for noisy images and denoised images using L_1 , L_2 , and L_1/L_2 regularization.

	PSNR	SSIM
Noisy	23.98	0.7124
L_1 regularization	24.02	0.8886
L_2 regularization	22.54	0.8059
L_1/L_2 regularization	25.39	0.9130

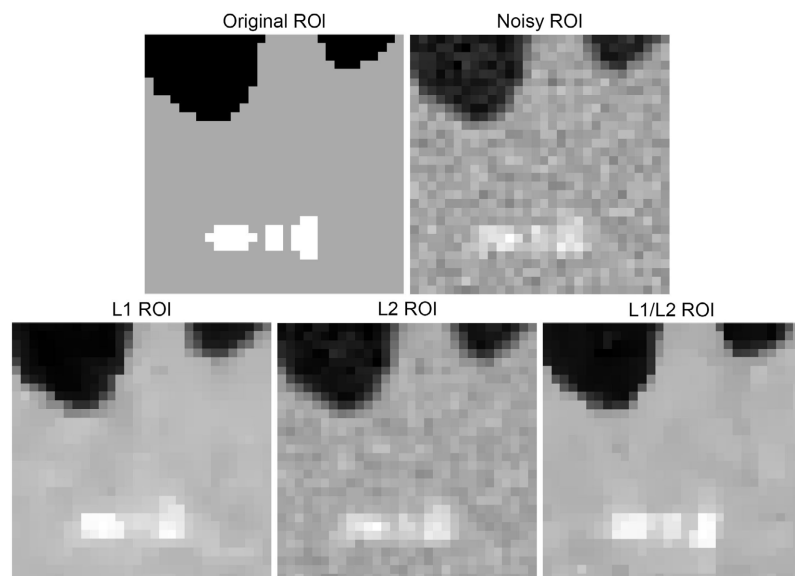


Figure 3. Local comparison results, select the local area in the red box in **Figure 2** (50 - 79 columns, 80 - 109 rows, a total of 30×30 pixels) for analysis.

The PSNR and SSIM of the three images are shown in **Table 1**, it can be seen that L_1/L_2 regularization has the highest PSNR and SSIM values, while L_2 regularization has the lowest PSNR and SSIM values. It can be seen that L_1/L_2 regularization has the best effect on image denoising and detail retention. In comparison, L_1 regularization can also effectively improve image quality. However, it is inferior

to L_1/L_2 regularization in terms of denoising and detail retention, while L_2 regularization performs the worst in terms of denoising and detail retention, and from a local zooming in image, as shown in **Figure 3**, L_1/L_2 is also superior to L_1 and L_2 regularization in terms of detail retention and denoising.

4.3. Explore the Denoising Effect of Real Low-Dose CT Image

In order to further prove the effectiveness of this algorithm in practical application, this experiment selected full dose CT and quarter dose CT images from data set files on the Kaggle platform. Kaggle is a globally renowned data science and machine learning competition platform that provides rich datasets for users to use. In order to study the effectiveness of the algorithm, this paper selected full dose CT and quarter dose chest CT images. The section thickness was 1mm, and the quarter dose CT was low-dose CT. As shown in **Figure 4**, it can be seen that there was obvious noise in the quarter dose CT images.

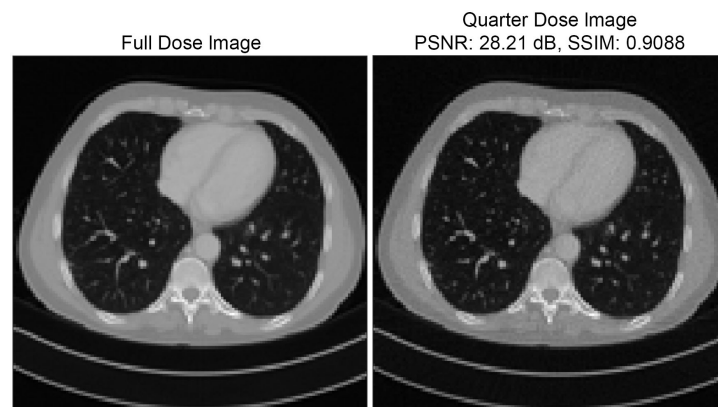


Figure 4. The left image is a full dose CT image, and the right image is a quarter dose image.

The algorithm in this paper was used to de-noise low-dose CT, the sampling times were still set to 400,000 times, and the parameters were set to $\alpha = 0.02$ and $\beta = 10^3$ to obtain three different kinds of CT images after regularization, as shown in **Figure 5**.

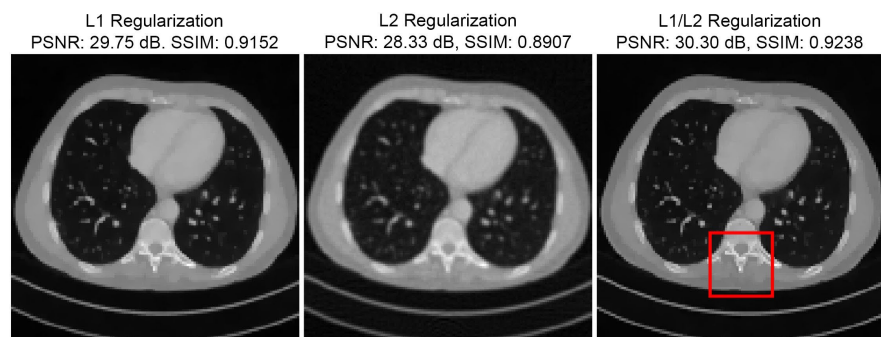


Figure 5. The first one is the denoised image using L_1 regularization, the second one is the denoised image using L_2 regularization, and the third one is the denoised image using L_1/L_2 regularization.

According to the experimental results, this is shown in **Figure 5** and **Table 2**, the PSNR of CT images after L_1 , L_2 and L_1/L_2 regularization has been improved, among which L_1/L_2 regularization has been improved the most, and SSIM has also been improved in addition to L_2 regularization decrease, which is also the most obvious improvement in L_1/L_2 regularization. In addition, by observing the local amplification area, as shown in **Figure 6**, L_1/L_2 regularized and denoised CT images also have the best performance in noise removal and detail retention.

Table 2. Comparison of PSNR and SSIM for noisy images and denoised images using L_1 , L_2 , and L_1/L_2 regularization.

	PSNR	SSIM
Quarter dose	28.21	0.9088
L_1 regularization	29.75	0.9152
L_2 regularization	28.33	0.8907
L_1/L_2 regularization	30.30	0.9238

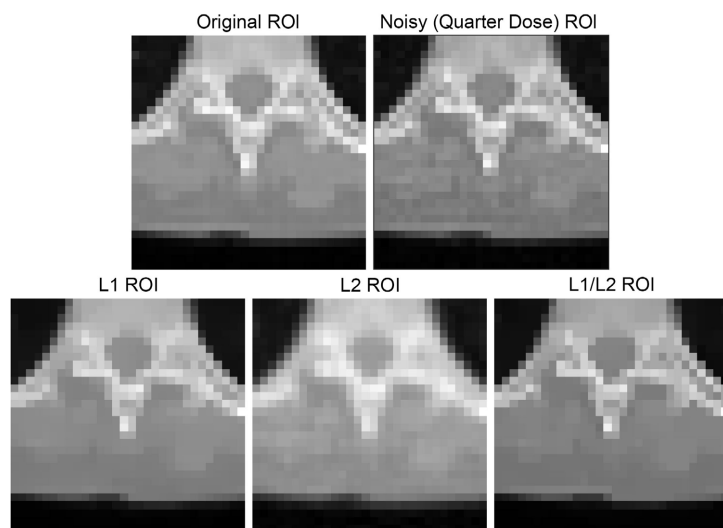


Figure 6. Local comparison results, select the local area in the red box in **Figure 5** for analysis.

Through the real clinical low-dose CT image denoising analysis, the effectiveness and advantages of the L_1/L_2 regularization denoising effect based on MCMC sampling algorithm are further proved.

5. Conclusions

In this paper, the low-dose CT denoising algorithm is analyzed and verified, and the regularization model based on MCMC sampling is introduced in this field. By employing MCMC sampling to process noisy image data, the denoising effect of three regularization strategies— L_1 regularization, L_2 regularization, and L_1/L_2

regularization—was evaluated.

The results demonstrated that L_1/L_2 regularization outperformed both L_1 and L_2 regularizations in multiple evaluation metrics. Particularly, L_1/L_2 regularization excelled in balancing noise removal with the preservation of image details, offering superior denoising results. This highlights the significance of using regularization in denoising models to effectively balance image fidelity and noise suppression. In addition, the adaptive step size in MCMC sampling further improves the stability and convergence speed of the algorithm, making the denoising process more efficient and practical. The experimental results show that the proposed method is not only feasible in theory, but also can achieve the expected denoising performance in practical application.

In summary, this study demonstrated the effectiveness and superiority of combining regularization models with MCMC sampling algorithms for image denoising. Future research could focus on optimizing computational efficiency and exploring its applications in broader image-processing tasks.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Mayo, J.R., Hartman, T.E., Lee, K.S., Primack, S.L., Vedal, S. and Müller, N.L. (1995) CT of the Chest: Minimal Tube Current Required for Good Image Quality with the Least Radiation Dose. *American Journal of Roentgenology*, **164**, 603-607. <https://doi.org/10.2214/ajr.164.3.7863879>
- [2] Haque, A., Wang, A.S. and Imran, A. (2022) Noise2Quality: Non-Reference, Pixel-Wise Assessment of Low Dose CT Image Quality. *Medical Imaging 2022: Image Perception, Observer Performance, and Technology Assessment*, San Diego, 20-24 February 2022, 120351C. <https://doi.org/10.1117/12.2611254>
- [3] Hadamard, J. (1923) Lectures on Cauchy's Problem in Linear Partial Differential Equations. Yale University Press.
- [4] Bertero, M. and Boccacci, P. (1998) Introduction to Inverse Problems in Imaging. IOP Publishing Ltd. <https://doi.org/10.1887/0750304359>
- [5] Liu, J., Huang, J. and Zhang, S. (2012) Limited-Angle CT Reconstruction via the L_1/L_2 Minimization. *Inverse Problems and Imaging*, **6**, 357-375.
- [6] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, **21**, 1087-1092. <https://doi.org/10.1063/1.1699114>
- [7] Roberts, G.O., Gelman, A. and Gilks, W.R. (1997) Optimal Scaling for Various Metropolis-Hastings Algorithms. *Statistical Science*, **7**, 473-483.
- [8] Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P. (2004) Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, **13**, 600-612. <https://doi.org/10.1109/tip.2003.819861>
- [9] Li, T., Li, X., Wang, J., et al. (2004) Nonlinear Sinogram Smoothing for Low-Dose X-Ray CT. *IEEE Transactions on Nuclear Science*, **51**, 2505-2513. <https://doi.org/10.1109/tns.2004.834824>