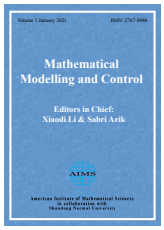
Mathematical Modelling and Control

https://www.aimspress.com/journal/mmc



MMC, 5(2): 216–235.
DOI: 10.3934/mmc.2025016
Received: 10 September 2024
Revised: 24 October 2024
Accepted: 28 November 2024
Published: 09 June 2025

Research article

Restoring medical images with combined noise base on the nonlinear inverse scale space method

Chenwei Li and Donghong Zhao*

School of Mathematics and Physics, University of Science and Technology Beijing, Beijing 100083, China

* **Correspondence:** Email: zdh751111@ustb.edu.cn.

Abstract: This paper proposes a model that eliminates combined additive and multiplicative noise by merging the Rudin, Osher, and Fatemi (ROF) and I-divergence data fidelity terms. Many important techniques for denoising and restoring medical images were included in this model. The addition of the I-divergence fidelity term increases the complexity of the model and difficulty of the solution in comparison with the ROF. To solve this model, we first proposed the generalized concept of the maximum common factor based on the inverse scale space algorithm. Different from general denoising algorithms, the inverse scale space method exploits the fact that u starts at some value c_0 and gradually approaches the noisy image f as time passes which better handles noise while preserving image details, resulting in sharper and more natural-looking images. Furthermore, a proof for the existence and uniqueness of the minimum solution of the model was provided. The experimental findings reveal that our proposed model has an excellent denoising effect on images destroyed by additive noise and multiplicative noise at the same time. Compared with general methods, numerical results demonstrate that the nonlinear inverse scale space method has better performance and faster running time on medical images especially including lesion images, with combined noises.

Keywords: inverse scale space method; total variation; medical image denoising; combined noises

1. Introduction

During image acquisition, storage, processing, and transmission, the sensor is easily affected by various factors. Besides, external factors may cause interference, resulting in some noise during image transmission. The presence of noise results in a blurred image and loss of intricate details from the original image, complicating any subsequent image analysis. The aim is to remove the effects of noise on the quality of the image, a standard problem in image restoration. Image denoising and restoration are some of the major image problems. The mathematical model builds on prior knowledge to degrade the image, and then uses the reverse process of image degradation to restore the image [1].

The additive noise model has been extensively studied by scholars. Most literature studies additive Gaussian noise, assuming that a given original image u is destroyed by additive noise $\eta \sim N(\mu, \sigma^2)$. The task is to recover u from the observed image f:

$$f = u + \eta. \tag{1.1}$$

Many effective methods have been developed to deal with this problem, and well-known techniques include wavelet approaches [2], learning-based methods [3–5], and variational approaches [6]. One of the learning-based approaches is the one that scholars are now committed to pursuing. In the literature [7], the authors proposed a new medical image denoising method to denoise images with Gaussian and impulse noise using end-to-end deep learning. Learning-based denoising methods can generally be used to remove specific types of noise. Variational-based methods are also important methods used for image denoising. Based on [6], the researchers proposed some improved methods: high-order total variation [8], fractional-order total variation [9], and Weberized total variation [10–

12].

In addition, there is another kind of noise called *multiplicative noise*. In a multiplicative noise model, f represents a noisy image, the original image u is destroyed by multiplicative noise η:

$$f = u\eta. \tag{1.2}$$

There are many possible distributions of multiplicative noise. Assume that the multiplicative noise in this paper follows a Gamma distribution with mean one with its probability density function given by

$$G(\eta) = \begin{cases} \frac{L^L \eta^{L-1}}{\Gamma(L)} e^{-L\eta}, & \eta > 0, \\ 0, & \eta \leq 0, \end{cases}$$

where $1/L$ is the variance. Generally, L takes an integer coefficient. The smaller the L is, the more severe the original image is destroyed by noise. $\Gamma(\cdot)$ is a Gamma function, $\Gamma(L) = (L-1)!$. Recently, there is a lot of interest in multiplicative noise removal. Different from additive noise, the distribution of multiplicative noise is usually non-Gaussian, and it has a multiplicative relationship with the original image. Due to its characteristics, it is more difficult and demanding to remove than additive noise. At present, scientists propose many techniques for eliminating multiplicative noise, such as the Lee method [13], scale threshold method [14], and variational methods [15–22].

However, the true observation image has been contaminated by many different types of noise. The above-mentioned denoising method can only remove a certain noise, or the effect of removing other types of noise is not obvious. In this paper we will focus on removing combined additive and multiplicative noise. Most learning-based denoising methods can remove a specific type (additive or multiplicative) of noise in most cases; whereas, variational-based methods have been studied for combined additive and multiplicative noise. There are variational models available for restoring images affected by combined noises. In [23], the author provided restoration of images corrupted with additive, multiplicative, and mixed noise by a new total least squares-based image denoising algorithm. In [24], the authors developed a variational model and an efficient nonlinear multigrid approach via a robust fixed-point smoother. Other scholars in [25] proposed a variational model combining a fractional-order total variation-norm with the fields of experts-image prior [26], employing a fast alternating minimization algorithm, Newton method, and a non-convex optimization iPiano algorithm [27].

In [28], the authors derived a scale space method for inverse problems, which usually starts from a noisy image and then gradually smooths the image. Different from other methods, the inverse scale space method starts from $u(x,0) = c_0$ and gradually approximates the noise image f as time grows. The inverse scale space method is related to the regularization theory, especially Tikhonov regularization [29]. In [28], the inverse scale space method worked effectively on quadratic regularization functions, which can derive a linear evolution equation, but it cannot derive good results for total variational functionals. Then, the author in [30,31] derived a general method of the inverse scale space as a limit of the iterative refining process. With the new method, the nonlinear inverse scale space method is implemented on the total variation model. Later, Shi and Osher extended the nonlinear inverse scale space method to the total variational model of multiplicative noise in [17].

In fact, real images often contain more than one type of noise, but may contain a mixture of additive and multiplicative noise, making the processing of mixed noise more challenging. The above scholars have only denoised specific types of noise using inverse scale space methods. Motivated by these works, we proposed a new variational model to restore the image damaged by combined noise and generalized the nonlinear inverse scale method on our model. This paper is dedicated to recovering images corrupted by a mixture of additive noise and multiplicative noise. The method provides a unified framework for mixed additive and multiplicative noise.

This paper is organized as follows. In Section 2, we introduce some basic mathematical knowledge and review some classic variational models for removing additive noise or multiplicative noise. In Section 3, we propose our new restoration model and prove that the minimum value exists. In Section 4, the relaxed inverse scale space method is introduced for the solution of the proposed model. Some numeric results of the experiments are presented in Section 5 and the paper concludes with Section 6.

2. Preliminary knowledge

2.1. Basic knowledge of bounded variation function space

Image denoising methods based on variational methods and partial differential equations (PDEs) have been widely studied because there is a complete set of mathematical theoretical systems and calculation methods.

The basic theory of bounded variation (BV) function space is the theoretical basis of image processing based on the variational method and PDEs. The total variation (TV) model is a classic model of BV function space theory applied to image denoising. Next, we introduce the theory of bounded variation function space.

Definition 2.1. ***(Total variation).*** *Let $\Omega \subset R^n$ be a bounded open subset and $u \in L^1_{loc}(\Omega)$ be a locally integrable function. The total variation of u is $TV(u) = \left\{sup\int_\Omega udiv\varphi dx; \varphi = (\varphi_1, \varphi_2, \cdots, \varphi_n) \in C_0^1(\Omega)^n, |\varphi|_{L^\infty(\Omega)} \le 1\right\}$. If $TV(u) < \infty$ is satisfied, then u is called a bounded variation function, denoted as $|u|_{BV(\Omega)}$. The space composed of all bounded variation functions in Ω is denoted as $BV(\Omega)$.*

Definition 2.2. *With a grayscale image u as a two-dimensional image, let it be the Sobolev space $W^{(1,1)}(\Omega)$, and then $\nabla u \in L^1(\Omega)$ and*

$$-\int_\Omega udiv\varphi dx = \int_\Omega \nabla u \cdot \varphi dx.$$

The total variation of the two-dimensional image u can be defined as

$$TV(u) = \int_\Omega |\nabla u| dx$$

where $\nabla u = (\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}) = (u_x, u_y)$, denoting the derivative of a two-dimensional gray image in regard to variables x and y, is called the gradient of the image, and $|\nabla u| = \sqrt{u_x^2 + u_y^2}$ denotes the modulus of the image gradient.

Definition 2.3. ***(Bounded variation function space).*** *The space of the bounded variation function can be defined as*

$$BV(\Omega) = \left\{u | u \in L^1(\Omega), TV(u) < \infty\right\}.$$

Under the definition of natural norm $\| u \|_{BV} = \| u \|_{L^1} + TV(u)$, $BV(\Omega)$ forms a Banach space.

Remark 2.1. *The function of bounded variation $\| u \|_{BV}$ is convex in the $BV(\Omega)$ space, but not strictly convex.*

2.2. Review of some related models

2.2.1. Total variation-based additive mode

- ROF model

In 1992, Rudin, Osher, and Fatemi used the space of a function of bounded variation establishing a total variation regularized image denoising model (ROF) in [6]. The authors constructed the following energy functional:

$$E(u) = \int_\Omega | \nabla u | dxdy, \tag{2.1}$$

which meets the following two constraints:

$$\begin{cases} \int_\Omega udxdy = \int_\Omega fdxdy, \\ \frac{1}{|\Omega|}\int_\Omega (u-f)^2 dxdy = \sigma^2. \end{cases} \tag{2.2}$$

The first formula represents that the mean value of noise is 0, and the second formula represents that the variance of noise is σ^2. By introducing the Lagrange multiplier, (2.1) and (2.2) are transformed into a new energy functional without restriction.

$$E(u) = \int_\Omega | \nabla u | dxdy + \frac{\lambda}{2}\int_\Omega \| u - f \|^2 dxdy$$

where λ is a positive parameter, which plays an important balance between smoothing and denoising, u denotes the image to be restored, and f denotes the noisy image.

2.2.2. Total variation-based multiplicative model

Among the many denoising methods for multiplicative noise, variational models have attracted the attention of many scholars. In this section, we review several classic total variation models for removing multiplicative noise.

- RLO model

In 2003, the first variational model was proposed by Rudin, Lions, and Osher in [15]. Assume that the noise in (1.2), the mean noise is 1, with variance σ^2. In other words, the noise satisfies (2.3).

$$\begin{cases} \int_\Omega \frac{f}{u} dxdy = 1, \\ \frac{1}{|\Omega|}\int_\Omega \left(\frac{f}{u} - 1\right)^2 dxdy = \sigma^2. \end{cases} \tag{2.3}$$

Use total variation $\int_\Omega | \nabla u | dxdy$ as the regular term. The model (RLO) is expressed as follows:

$$E(u) = \int_\Omega | \nabla u | dxdy + \lambda\int_\Omega \frac{f}{u} dxdy + \mu\int_\Omega \left(\frac{f}{u} - 1\right)^2 dxdy$$

where λ and μ are two Lagrange multipliers.

- AA model

In 2008, based on the maximum a posterior (MAP) estimator, Aubert and Aujol in [16] derived a non-convex variational model (AA) (2.4) and transformed the solution of the AA model into solving its corresponding Euler-Lagrange equation. This model is suitable for the multiplicative noise of the Gamma distribution with a mean value of 1.

$$E(u) = \int_\Omega \mid \nabla u \mid dxdy + \lambda \int_\Omega \left(\log u + \frac{f}{u}\right) dxdy \qquad (2.4)$$

where λ is parameter.

Although the AA model is non-convex, it still proves the existence of its minimum and gives a sufficient condition to ensure the uniqueness of the solution. However, since the AA model is non-convex, its solution may not be the optimal solution, and the solution depends on the initial value selection.

- I-divergence model

In 2009, Steidl and Teuber in [18] proposed a convex variational model using I-divergence as the fidelity term to remove multiplicative noise.

Definition 2.4. *Let $u \in R_+^n$ and $f \in R_+^n$. The I-divergence of f from u is defined in [32] by*

$$I(f \parallel u) := \int_\Omega \left(f \log \frac{f}{u} - f + u\right) dx.$$

With the property $I(f \parallel u) \geq 0$, if and only if $u = f$, the equal sign holds. Ignore the constant terms, the data fidelity term is simplified to $\int_\Omega (u - f \log u) dx$, and the denoising model is as follows:

$$E(u) = \int_\Omega \mid \nabla u \mid dxdy + \lambda \int_\Omega (u - f \log u) dxdy$$

where λ is parameter.

Although I-divergence is generally used as a fidelity term to remove Poisson noise, the author applied Douglas-Rachford splitting techniques and alternating direction methods of multipliers, and proved that it can also effectively remove Gamma noise.

3. The proposed model and analysis

In this section, we propose a novel variational model combining additive noise (AN) and multiplicative noise (MN). This model is more common in real life, and its solving process is also more challenging. We begin with an observation image contaminated by AN and MN, with the assumption that AN and MN are independent or unrelated. In [24, 33], the following model is proposed for formulating the image:

$$f = u + k_0\eta + k_1 u\eta$$

where k_0 and k_1 are constants and η represents Gaussian noise with variance 1.

We propose the following convex model:

$$\begin{aligned} E(u) = &\int_\Omega \mid \nabla u \mid dxdy + \alpha_1 \int_\Omega (u - f \log u) dxdy \\ &+ \frac{\alpha_2}{2} \int_\Omega \parallel u - f \parallel_2^2 dxdy \end{aligned} \qquad (3.1)$$

where α_1 and α_2 are parameters that can be adjusted to achieve a balance between the additive and multiplicative parts of the fidelity.

- When $\alpha_1 = 0$, this reduces to the model in [6].
- When $\alpha_2 = 0$, this reduces to the model in [18].

For the proof of the existence of the model solution, the boundary condition of the solution must be proved [34].

Lemma 3.1. ***(Boundedness).*** *Let $f \in L^2(\Omega)$ and $\inf_\Omega f > 0$, where $\Omega \subset R^2$. If the optimization problem Eq (3.1) has a solution $\hat{u}$, then $\inf_\Omega f \leq \hat{u} \leq \sup_\Omega f$.*

Proof. Define $\alpha = \inf_\Omega f$, $\beta = \sup_\Omega f$. Notice that $u > f$, $u - f \log u$, and $(u - f)^2$ strictly monotonically increase. Then

$$\int_\Omega (\inf(u,\beta) - f \log(\inf(u,\beta))) dxdy \leq \int_\Omega (u - f \log u) dxdy, \qquad (3.2)$$

$$\int_\Omega (\inf(u,\beta) - f)^2 dxdy \leq \int_\Omega (u - f)^2 dxdy. \qquad (3.3)$$

Besides, Lemma 1 in [35] states that there is

$$TV(\inf(u,\beta)) \leq TV(u). \qquad (3.4)$$

Combining (3.2)–(3.4),

$$E(\inf(u,\beta)) \le E(u). \tag{3.5}$$

If $u \le \beta$, the equation is valid.

Since $\hat{u}$ is the minimum solution of Eq (3.1), the equation holds when $u = \hat{u}$, thus $\hat{u} \le \beta$.

It is also the case that

$$E(u) \le E(\sup(u,\alpha)). \tag{3.6}$$

Thus $\hat{u} \ge \alpha$.

So there is $\alpha \le \hat{u} \le \beta$, that is, $\inf_{\Omega} f \le \hat{u} \le \sup_{\Omega} f$. □

Theorem 3.1. ***(Existence).*** *Let $f \in L^2(\Omega)$ and $\inf_{\Omega} f > 0$, where $\Omega \subset R^2$, and then there exists at least one solution in the solution space $BV(\Omega)$ by Eq (3.1).*

Proof. Define $\alpha = \inf_{\Omega} f, \beta = \sup_{\Omega} f$. Because there is $u \equiv \beta \in BV(\Omega)$, the solution space is not empty [10]. Consider the optimization problem's minimization sequence $u_n \in BV(\Omega)$. According to Lemma 3.1, we have $\alpha \le u_n \le \beta$.

Notice that $\{u_n\}$ is a bounded sequence in $BV(\Omega)$, Ω is also bounded, and then

$$\| u_n \|_{L^1(\Omega)} < +\infty. \tag{3.7}$$

Moreover, by the definition of a minimization sequence, we know there is a constant $C > 0$ and $E(u_n) \le C$. When $u_n = f$, $\int_\Omega (u_n - f \log u_n) dxdy$ reaches a positive minimum $\int_\Omega (f - f \log f) dxdy$ and $\int_\Omega (u_n - f)^2 dxdy$ reaches 0. Since $TV(u_n) < \infty$, we have

$$TV(u_n) \le C^{'}. \tag{3.8}$$

From (3.7) and (3.8), there will be evidence that $\{u_n\}$ is consistently bounded in $BV(\Omega)$ space with respect to n. Because of the compactness of $BV(\Omega)$ space, there are subcolumns in $\{u_n\}$ (this remains as $\{u_n\}$ for simplicity) and function u in $BV(\Omega)$ space such that $\{u_n\}$ converges strongly to u in $L^1(\Omega)$ space. More, assume that $u_n(x,y) \to u(x,y)$, a.e. $\forall x, y \in \Omega$.

The Lebesgue control convergence theorem allows us to deduce

$$\int_\Omega (u - f \log u) dxdy = \lim_{n\to\infty} \int_\Omega (u_n - f \log u_n) dxdy. \tag{3.9}$$

According to the lower semi-continuity of the function, we have the following formula:

$$E(u) \le \lim_{n\to\infty} \inf E(u_n). \tag{3.10}$$

Since $\{u_n\}$ is a minimization sequence, it is the smallest solution to the optimization problem Eq (3.1), and then Eq (3.1) has at least one solution in its solution space.

The objective function $E(u)$ in the optimization problem Eq (3.1) is convex, and the uniqueness of the solution thus follows directly from the convexity of the objective function. □

4. Numerical methods

In this section, we will introduce the basic principles and solution process of the inverse scale space method and apply this method to our proposed model. Some scholars have used the inverse scale space method to remove additive noise, and some scholars have used this method to remove multiplicative noise. However, no one has used this method to remove combined noise on our model. Encouraged by this, we have made some improvements on the inverse scale space method to remove combined noise.

4.1. Inverse scale space method

In the following we generalize nonlinear inverse scale space flow [30, 31] to the proposed model (3.1), a relaxation method using two evolution equations to approximate the real evolution equations.

The method relies on a convex variance problem:

$$\min_u \{J(u) + \lambda G(u,f)\} \tag{4.1}$$

where $J(u)$ is the regularization term, and $G(u,f)$ is the data fidelity term. We assume that f has a zero mean: $\int_\Omega f dxdy = 0$.

Definition 4.1. ***(Subdifferential).*** *Let f be a proper lower-semi-continuous convex function on R^m, where R^m is $m-$dimensional Euclidean space. The subdifferential of f is defined for $y \in R^m$ by*

$$\partial f := \{y \in R^m and f(z) \ge f(x) + \langle y, z - x\rangle, \forall z \in R^m\}.$$

Definition 4.2. ***(Bregman distance).*** *Assuming $J : X \to R$ is a convex function, where $X \in R^m$, $u, v \in X$, $p \in \partial J(v)$, then the Bregman distance between u and v is defined as $D_J^p(u, v)$:*

$$D_J^p(u, v) = J(u) - J(v) - \langle u - v, p\rangle .$$

From the definition of the subdifferential Bregman distance $D_J^p(u, v) \geq 0$, the Bregman distance does not satisfy symmetry, $D_J^p(u, v) \neq D_J^p(v, u)$. Assume $p \in \partial J(u) \cap \partial J(v)$, and then

$$\begin{aligned} D_J^p(u, v) &= J(u) - J(v) - \langle u - v, p\rangle \\ &= J(u) - J(v) + \langle v - u, p\rangle \\ &= -(J(v) - J(u) - \langle v - u, p\rangle) \\ &= -D_J^p(v, u). \end{aligned}$$

Applying Bregman's regularization method, (4.1) can be rewritten as

$$u_k = \arg\min_u \{D(u, u_{k-1}) + \lambda G(u, f)\} . \tag{4.2}$$

According to the generalized form of Bregman distance, $D(u, u_{k-1}) = J(u) - J(u_{k-1}) - \langle u - u_{k-1}, p_{k-1}\rangle$, where $p_{k-1} \in \partial J(u_{k-1})$, so (4.2) can be rewritten as

$$u_k = \arg\min_u \{J(u) - J(u_{k-1}) - \langle u - u_{k-1}, p_{k-1}\rangle + \lambda G(u, f)\} .$$

Omitting the constant term (the term about u_{k-1}) irrelevant to minimization yields

$$u_k = \arg\min_u \{J(u) - \langle u, p_{k-1}\rangle + \lambda G(u, f)\} . \tag{4.3}$$

The Euler-Lagrange equation of (4.3) is

$$p(u_k) - p(u_{k-1}) + \lambda \partial_u G(u_k, f) = 0 \tag{4.4}$$

where $\partial_u G$ denotes the variation of G with respect to u.

We are led to the relation:

$$\frac{p(u_k) - p(u_{k-1})}{\lambda} = -\partial_u G(u_k, f). \tag{4.5}$$

Let $\lambda = \triangle t$, $p(u_k) = p(u_k(t))$, and

$$\frac{p(u_k) - p(u_{k-1})}{\triangle t} = -\partial_u G(u_k(t), f).$$

When $\triangle t \to 0$, we arrive at a continuous process

$$\frac{\partial p}{\partial t} = -\partial_u G(u(t), f), p \in \partial J(u). \tag{4.6}$$

The initial conditions for (4.6) are $u\ |_{t=0} = c_0$, $p\ |_{t=0} = 0$, $\int_\Omega \partial_u G(c_0, f) = 0$.

The flow begins with $u(0) = c_0$ and it converges back to the image f as $t \to \infty$, i.e., $\lim_{t\to\infty} u(t) = f$. By (4.6) the image $u(t)$ flows from the smoothest possible image ($u(0) = c_0$) to the noisy image f. We are trying to use the flow to denoise the image.

Since the relationship between p and u is difficult to calculate in the case of non-linearity, (4.6) cannot be directly calculated, so we propose a relaxation method to approximate the direct flow.

The first-order optimality condition and $v_k = p_k/\lambda$ yield the following formula from (4.5):

$$v_k - v_{k-1} = -\partial_u G(u_k, f),$$

$$v_k = v_{k-1} - \partial_u G(u_k, f), k \geq 1, v_0 = 0. \tag{4.7}$$

Therefore,

$$v_k = -\sum_{j=1}^{k} -\partial_u G(u_j, f), k \geq 1$$

and a system of equations about u_k(4.3) and v_k(4.7) can be expressed as

$$\begin{cases} u_k = \arg\min_u \{J(u) + \lambda(G(u, f) - \langle u, v_{k-1}\rangle)\} , \\ v_k = v_{k-1} - \partial_u G(u_k, f), \end{cases} \tag{4.8}$$

where $u_0 = c_0$, $v_0 = 0$, $k = 1, 2, 3, \cdots$, $\int_\Omega \partial_u G(c_0, f) = 0$.

It can also be rewritten as

$$\begin{aligned} u_k &= \arg\min_u \{J(u) - \langle u, \lambda v_{k-1}\rangle + \lambda G(u, f)\} \\ &= \arg\min_u \left\{ J(u) - \left\langle u, -\lambda \sum_{j=1}^{k-1} \partial_u G(u_j, f)\right\rangle + \lambda G(u, f)\right\} \\ &= \arg\min_u \left\{ J(u) + \lambda \sum_{j=1}^{k-1} \left\langle u, \partial_u G(u_j, f)\right\rangle + \lambda G(u, f)\right\}. \end{aligned}$$

After converging to a minimizer u_k in the first formula of (4.8), it is easy to compute $\partial_u G(u_k, f)$ and update v_k. However, the amount of calculation in this process is quite high.

The standard way of solving u_k and v_k is first to evolve a steepest descent flow for u_k, having a fixed v_k, based on the Euler-Lagrange equations of (4.8):

$$p + \lambda(\partial_u G(u, f) - v_{k-1}) = 0, p \in \partial J(u).$$

The corresponding diffusion equation is

$$\frac{\partial u}{\partial t} = -p \ + \ \lambda(-\partial_u G(u, f) \ + \ v_k), p \in \partial J(u), u \mid_{t=0} = \ u_{k-1}. \tag{4.9}$$

The update for v_k in the second formula of (4.8) can be viewed as an iterative descent in v_{k-1} for minimizing $G(u_k, f)$. This is an indirect minimization, which affects u_k by its coupling with v_k. We write the solution for v_k in the following manner:

$$\frac{\partial v}{\partial \tau} = -\partial_u G(u_k, f), v \mid_{\tau=0} = \ v_{k-1}.$$

τ is a time variable similar to t. This extends the definition of the sequence v_k to a continuous formulation.

We propose two continuous flows $u(t)$ and $v(t)$ to approximate the sequences u_k and v_k. Let $\tau = mt$, hence we obtain the following:

$$\frac{\partial v}{\partial t} = m\frac{\partial v}{\partial \tau} = -m\partial_u G(u_k, f). \tag{4.10}$$

Replacing v_k in (4.9) by $v(t)$ and u_k in (4.10) by $u(t)$ yield the relaxed inverse scale space flow:

$$\begin{cases} \frac{\partial u}{\partial t} = -p + \lambda(-\partial_u G(u, f) + v) \\ \frac{\partial v}{\partial t} = -m\partial_u G(u_k, f) \end{cases} \tag{4.11}$$

with $p \in \partial J(u)$ and initial conditions $u \mid_{t=0} = \ c_0, v \mid_{t=0} = \ 0$, $\int_\Omega \partial_u G(c_0, f) = \ 0$.

4.2. Analysis of our model

4.2.1. Inverse scale space flow of our model

Consider our model

$$E(u) = \int_\Omega \mid \nabla u \mid dxdy + \alpha_1 \int_\Omega (u - f \log u)dxdy + \frac{\alpha_2}{2} \int_\Omega \parallel u - f \parallel_2^2 dxdy.$$

In [17, 20], the authors used the inverse scale space method to solve the denoising model containing only one fidelity term. For the two fidelity terms in our paper, we do the following processing to apply the inverse scale space method theory.

The greatest common divisor is the greatest common divisor shared by two or more integers, and the parameters of the fidelity term of the model proposed in this paper may be small, thus we generalize the concept of the greatest common divisor as follows. We extend the concept of the greatest common divisor from integers to decimals. Define $\lambda = d(\alpha_1, \alpha_2)$,

$$d(\alpha_1, \alpha_2) = \frac{\alpha_1}{\mu_1} = \frac{\alpha_2}{\mu_2}$$

where λ denotes the greatest common divisor of α_1 and α_2, μ_1 and μ_2 are integers, and α_1 and α_2 can be integers or decimals. Below are two examples of λ.

When $\alpha_1 = 1$, $\alpha_2 = 2$, then $\lambda = d(1, 2) = 1$. λ is an integer.

When $\alpha_1 = 0.36$, $\alpha_2 = 0.45$, then $\lambda = d(0.36, 0.45) = 0.09$. λ is a decimal.

In the model of this paper, the function $G(u, f)$ is composed of two terms and the coefficients of the fidelity term are two different decimals, whereas the traditional inverse scale space method only applies to one fidelity term, so we promote the concept of the greatest common divisor to combine $d(\alpha_1, \alpha_2)$ with the traditional inverse scale space method.

We will mark the last two terms of $E(u)$ as $g(u, f)$. Let $g(u, f) = \lambda G(u, f)$, and therefore,

$$G(u, f) = \mu_1 \int_\Omega (u - f \log u)dxdy + \frac{\mu_2}{2} \int_\Omega \parallel u - f \parallel_2^2 dxdy.$$

Thus, it is possible to obtain:

$$\partial_u G(u(t), f) = \mu_1(1 - \frac{f}{u}) + \mu_2(u - f).$$

Then, we get the following inverse scale space model from (4.6):

$$\begin{gathered} \frac{\partial p}{\partial t} = \mu_1 \left(\frac{f}{u} - 1\right) + \mu_2(f - u), \\ p \in \partial E(u). \end{gathered}$$

We also have

$$\frac{\partial v}{\partial t} = -m\partial_u G(u(t), f) = -m[\mu_1(1 - \frac{f}{u}) + \mu_2(u - f)]$$

from (4.10).

We can also obtain the corresponding relaxed inverse scale space flow of the model proposed in this paper:

$$\begin{cases} \frac{\partial u}{\partial t} = -p(u) + \lambda \left[\mu_1 \left(\frac{f}{u} - 1\right) + \mu_2(f - u) + v\right], \\ \frac{\partial v}{\partial t} = m \left[\mu_1 \left(\frac{f}{u} - 1\right) + \mu_2(f - u)\right], \\ p(u) = -\nabla \cdot \frac{\nabla u}{|\nabla u|}. \end{cases} \tag{4.12}$$

Its initial conditions are $u_0 = c_0$, $p_0 = 0$, $\int_\Omega \partial_u G(c_0, f) = 0$.

The procedure of solving (4.12) is summarized as follows.

Algorithm: Inverse scale space method (ISSM).

Input: $\triangle t, \mu_1, \mu_2, \lambda, \beta$ are small fixed parameters to avoid singularity.

1. Initialization: $u_0 = c_0, v_0 = 0$.

2. Iteration. For every $k = 1, 2, 3, \cdots$, compute u_{k+1} and v_{k+1}.

$u_{k+1} = u_k$

$+\triangle t \left\{ div\left(\frac{\nabla u_k}{\sqrt{|\nabla u_k|^2+\beta^2}} \right) + \lambda \left[\mu_1 \left(\frac{f}{u_k} - 1 \right) + \mu_2 (f - u_k) + v_k \right] \right\}$,

$v_{k+1} = v_k + m\triangle t \left[\mu_1 \left(\frac{f}{u_k} - 1 \right) + \mu_2 (f - u_k) \right]$.

3. $k = k + 1$, and go to Step 2. If the stop condition is met, the recovered image u is used for quitting.

4.2.2. The convergence behavior

It is obvious that these equations ($\frac{\partial u}{\partial t} = 0$, $\frac{\partial v}{\partial t} = 0$) are the equilibrium: $u = f, v = \frac{p(f)}{\lambda}$. We still try to show that for any $f \in BV$, the solutions converge to this steady state.

The following energy function is defined to analyze the convergence:

$$e(t) = \frac{1}{2\lambda} \parallel u - f \parallel_2^2 + \frac{1}{2\alpha} \parallel v - \frac{q}{\lambda} \parallel_2^2$$

where $q \in \partial J(f), u = u(t), v = v(t)$.

Proposition 4.1. *Let $u(0), v(0)$ be an initial value such that $e(0) < \infty$. Let $u(t), v(t)$ be the solution of (4.12) with $\lambda > 0, \alpha > 0$. Then the energy $e(t)$ decreases monotonically. Moreover, when $t \to \infty$, then $e(t) \to 0$.*

Proof. We calculate the time derivative of the energy functional to obtain

$$\begin{aligned}
\frac{de(t)}{dt} &= \frac{1}{\lambda} \langle u - f, u_t \rangle + \frac{1}{\alpha} \left\langle v - \frac{q}{\lambda}, v_t \right\rangle \\
&= \frac{1}{\lambda} \left\langle u - f, -p(u) + \lambda \left[\mu_1 \left(\frac{f}{u} - 1 \right) + \mu_2 (f - u) + v \right] \right\rangle \\
&\quad + \frac{1}{\alpha} \left\langle v - \frac{q}{\lambda}, \alpha \left[\mu_1 \left(\frac{f}{u} - 1 \right) + \mu_2 (f - u) \right] \right\rangle \\
&= \left\langle u - f, -\frac{p(u)}{\lambda} + \mu_1 \left(\frac{f}{u} - 1 \right) + \mu_2 (f - u) + v \right\rangle \\
&\quad + \left\langle v - \frac{q}{\lambda}, \mu_1 \left(\frac{f}{u} - 1 \right) + \mu_2 (f - u) \right\rangle.
\end{aligned}$$

In steady state, $u - f = 0$, this is equivalent to $\mu_1 \left(\frac{f}{u} - 1 \right) + \mu_2 (f - u) = 0$. We replace $\mu_1 \left(\frac{f}{u} - 1 \right) + \mu_2 (f - u)$ by $f - u$.

$$\begin{aligned}
\frac{de(t)}{dt} &= \left\langle u - f, -\frac{p(u)}{\lambda} + f - u + v \right\rangle + \left\langle v - \frac{q}{\lambda}, f - u \right\rangle \\
&= \langle u - f, f - u \rangle + \left\langle u - f, -\frac{p(u)}{\lambda} + v \right\rangle + \left\langle v - \frac{q}{\lambda}, f - u \right\rangle \\
&= - \parallel u - f \parallel^2 - \left\langle u - f, \frac{p(u)}{\lambda} - \frac{q}{\lambda} \right\rangle \\
&= - \parallel u - f \parallel^2 - \frac{1}{\lambda} \langle u - f, p(u) - q \rangle .
\end{aligned}$$

According to the definition of the Bregman distance, we know $D(u, f) = J(u) - J(f) - \langle u - f, \partial J(f) \rangle$ and $D(f, u) = J(f) - J(u) - \langle f - u, \partial J(u) \rangle$. Therefore, $D(u, f) + D(f, u) = \langle u - f, \partial J(u) - \partial J(f) \rangle$.

Finally, we get

$$\frac{de(t)}{dt} = - \parallel u - f \parallel^2 - \frac{1}{\lambda} (D(u, f) + D(f, u)) \leq 0$$

which implies that the energy $e(t)$ decreases monotonically. Because of $e(t) \geq 0$ and $\frac{de(t)}{dt} \leq 0$, we get $\lim_{t \to \infty} e(t) = 0$. □

The dual function of u is $E^*(p) := \sup_u \{\langle u, p \rangle - E(u)\}$, and then $p \in \partial_u E(u)$ and $u \in \partial_p E^*(p)$ are equivalent. If we can compute the dual function E^*, we can get an explicit relationship for $u(p)$. Based on the above conditions, we can get the following estimator.

First, the derivatives of the fitted function and $u(t)$ in the time direction are calculated separately:

$$\frac{1}{2} \frac{d}{dt} \parallel u(t) - f \parallel_{L^2}^2 = \langle u(t) - f, \partial_t u(t) \rangle , \tag{4.13}$$

$$\begin{aligned}
\partial_t u(t) &= \frac{d}{dt} \partial_p E^*(p) = G^*(p(t)) \partial_t p(t) \\
&= -G^*(p(t)) \left(\frac{\mu_1}{u} + \mu_2 \right) (u(t) - f).
\end{aligned} \tag{4.14}$$

Because $u > 0$, u is bounded, $\mu_1 > 0$, and $\mu_2 > 0$, we might as well set $\frac{\mu_1}{u} + \mu_2 = a > 0$, and then

$$\partial_t u(t) = -aG^*(p(t))(u(t) - f) \tag{4.15}$$

where $G^* = \partial_{pp} E^*$ is the Hessian matrix of the dual function.

If E^* is strictly convex, then there is a constant $b > 0$ that makes

$$\langle \varphi, G^*(\omega) \varphi \rangle \geq b \parallel \varphi \parallel^2 \tag{4.16}$$

where $\forall \varphi, \omega \in BV^*$. Therefore combine (4.13), (4.15), and (4.16), and we get

$$\frac{1}{2}\frac{d}{dt} \parallel u(t) - f \parallel_{L^2}^2 = \langle u(t) - f, -aG^*(p(t))(u(t) - f)\rangle \leq -ab \parallel u(t) - f \parallel_{L^2}^2 .$$

Using Gronwall's inequality, there is

$$\parallel u(t) - f \parallel_{L^2} \leq e^{-ab(t-s)} \parallel u(s) - f \parallel_{L^2} \leq e^{-abt} \parallel f \parallel_{L^2}$$

if $t > s$. Thus, $t \to \infty$, $u \xrightarrow{L^2} f$.

In the estimate of L^2 above, f can be either a clean image or an image containing noise. Assume f is a clean image and $E(f) < \infty$, and we can get the Bregman distance error estimator

$$\begin{aligned}\frac{d}{dt}D(f, u(t)) &= \frac{d}{dt}\left[E(f) - E(u(t)) - \langle f - u(t), p(t)\rangle\right] \\ &= -\langle f - u(t), \partial_t p(t)\rangle \\ &= -\left\langle f - u(t), \left(\frac{\mu_1}{u} + \mu_2\right)(f - u(t))\right\rangle \\ &= -a \parallel u(t) - f \parallel_{L^2}^2 .\end{aligned}$$

If $q(t) \in L^2 \cap \partial E(f)$, we can get estimators for $p(t)$ and $q(t)$,

$$\begin{aligned}\frac{1}{2}\frac{d}{dt} \parallel p(t) - q(t) \parallel_{L^2}^2 &= \langle \partial_t p(t), p(t) - q(t)\rangle \\ &= \left\langle \left(\frac{\mu_1}{u} + \mu_2\right)(f - u(t)), p(t) - q(t)\right\rangle \\ &= a\langle (f - u(t)), p(t) - q(t)\rangle \\ &= a(-D(f, u) - D(u, f)).\end{aligned}$$

Integrate the above formula:

$$\begin{aligned}&\frac{1}{2}\int_0^t \frac{d}{dt} \parallel p(t) - q(t) \parallel_{L^2}^2 dt \\ &= \int_0^t a(-D(f, u) - D(u, f))dt, \\ &\frac{1}{2}(\parallel p(t) - q(t) \parallel_{L^2}^2 - \parallel p(0) - q(t) \parallel_{L^2}^2) \\ &+ \int_0^t a(D(f, u(s)) + D(u(s), f))ds \\ &= 0.\end{aligned}$$

Since $p(0) = 0, \parallel p(t) - q(t) \parallel_{L^2}^2$, $D(u(s), f)$ are all non-negative and $D(f, u(s)) \geq D(f, u(t))$ $(s \leq t)$, $D(f, u(t)) \leq \frac{\|q(t)\|_{L^2}^2}{2at}$. Thus, $t \to \infty$, $u \xrightarrow{D} f$.

When f is a noisy image, w is a clean image, and $E(w) < \infty$, there may be $E(f) = \infty$. In this case, there are the following propositions:

Proposition 4.2. *Based on the above conditions, if* $\parallel f - u(t) \parallel_{L^2} > \sigma$ *and* $\parallel f - g \parallel_{L^2} \leq \sigma$, *and* σ *is the standard deviation of the noise, then* $D(g, u(t))$ *is monotonically decreasing in time.*

Proof.

$$\begin{aligned}\frac{d}{dt}D(g, u(t)) &= \langle -\partial_t p(t), g - u(t)\rangle \\ &= \left\langle \left(\frac{\mu_1}{u} + \mu_2\right)(f - u(t)), g - u(t)\right\rangle \\ &= -\langle a(f - u(t)), g - f + f - u(t)\rangle \\ &= -a \parallel f - u(t) \parallel_{L^2}^2 - a\langle f - u(t), g - f\rangle \\ &\leq -\frac{a \parallel f - u(t) \parallel_{L^2}^2}{2} + \frac{a \parallel f - g \parallel_{L^2}^2}{2}.\end{aligned}$$

When $\parallel f - u(t) \parallel_{L^2} > \parallel f - g \parallel_{L^2}$, the right-hand side of the inequality is negative, and $D(g, u(t))$ is monotonically decreasing in time. □

4.2.3. Initial conditions

For the initialization of the inverse scale space, $u(0) = c_0$, we use the fact that

$$\int_\Omega \partial_t p = 0$$

which means that $\int_\Omega \partial_u G(u(t), f) = 0$ is time invariant for all time.

As $t \to 0$, the expression vanishes, leading to

$$\int_\Omega \partial_u G(c_0, f) = 0.$$

Combining $\partial_u G(u, f) = \mu_1\left(1 - \frac{f}{u}\right) + \mu_2(u - f)$, we obtain

$$\begin{aligned}\int_\Omega \left[\mu_1\left(1 - \frac{f}{c_0}\right) + \mu_2(c_0 - f)\right] &= 0, \\ \int_\Omega \left(\frac{\mu_1(c_0 - f) + \mu_2(c_0^2 - fc_0)}{c_0}\right) &= 0, \\ \int_\Omega \left(\frac{\mu_2 c_0^2 + (\mu_1 - \mu_2 f)c_0 - \mu_1 f}{c_0}\right) &= 0,\end{aligned}$$

which means that

$$\frac{\mu_2 c_0^2 + (\mu_1 - \mu_2 f)c_0 - \mu_1 f}{c_0} = 0, c_0 \neq 0,$$

so

$$\mu_2 c_0^2 + (\mu_1 - \mu_2 f)c_0 - \mu_1 f = 0. \tag{4.17}$$

The solution of this Eq (4.17) for c_0 is: $c_{01} = f, c_{02} = -\frac{\mu_1}{\mu_2} < 0$.

Thus the initial condition of ISS flow is given: $c_0 = f$.

4.2.4. Second order in time formulation

Our relaxed model can be written in a general form as

$$\begin{cases} \frac{\partial u}{\partial t} = -p(u) + \lambda(-\partial_u G(u,f) + v), \\ \frac{\partial v}{\partial t} = -m\partial_u G(u,f). \end{cases} \tag{4.18}$$

The first equation in (4.18) can be written in u with a second-order derivative in the time domain.

$$\partial^2_{tt} u = -\partial_t p(u) + \lambda(-\partial_t(\partial_u G(u,f)) + \partial_t v).$$

Replacing $\partial_t v$ in the above formula with the second equation in (4.18), we can obtain

$$\partial^2_{tt} u = -\partial_u p(u)\partial_t u + \lambda(-(\partial^2_{uu} G(u,f)\partial_t u) - m\partial_u G(u,f)).$$

The above formula can be rewritten as

$$\partial^2_{tt} u + \partial_u p(u)\partial_t u + \lambda(\partial^2_{uu} G(u,f)\partial_t u) + \lambda m\partial_u G(u,f) = 0$$

with the initial conditions $u\mid_{t=0} = c_0, u_t\mid_{t=0} = -\lambda\partial_u G(0,f)$.

As in [30, 31] we approximate $p(u) = -\Delta u$, which leads to

$$\partial^2_{tt} u + (-\Delta + \lambda\partial^2_{uu} G(u,f))\partial_t u + \lambda m\partial_u G(u,f) = 0. \tag{4.19}$$

We fix u as a constant u_0 and f as a constant f_0. Rewriting (4.19) in the frequency domain ξ by taking the Fourier transform, we obtain the characteristic equation

$$r^2 + (\mid \xi \mid^2 + \lambda\partial^2_{uu} G(u_0,f_0))r + \lambda m\partial_u G(u_0,f_0) = 0$$

with two solutions

$$r_\pm = \frac{-(\mid \xi \mid^2 + \lambda\partial^2_{uu} G(u_0,f_0))}{2} \pm \frac{\sqrt{(\mid \xi \mid^2 + \lambda\partial^2_{uu} G(u_0,f_0))^2 - 4\lambda m\partial_u G(u_0,f_0)}}{2}.$$

Both roots exist for all frequencies if

$$(\partial^2_{uu} G(u_0,f_0))^2 \geq \frac{4m}{\lambda}\partial_u G(u_0,f_0)$$

and we know

$$\partial_u G(u_0,f_0) = \mu_1\left(1 - \frac{f_0}{u_0}\right) + \mu_2(u_0 - f_0),$$

$$\partial^2_{uu} G(u_0,f_0) = \frac{\mu_1 f}{u_0^2} + \mu_2.$$

Thus, we should require the parameters α and λ meet

$$\frac{\alpha}{\lambda} \leq \frac{(\frac{\mu_1 f}{u_0^2} + \mu_2)^2}{4\left(\mu_1\left(1 - \frac{f_0}{u_0}\right) + \mu_2(u_0 - f_0)\right)}$$

to ensure monotone evolution of u toward f.

5. Numerical results

In this section, some numerical experiments on medical images show the performance of our proposed model, expanding the application field of the model. The peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) are currently the most widely used methods for evaluating image quality. The larger the value of PSNR is, the closer the denoised image is to the original image, which means that the effect of image restoration is better.

First, we introduce the mean square error (MSE).

$$MSE(f,u) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(f_{ij} - u_{ij})^2$$

where f_{ij} and u_{ij} respectively represent the pixel value of the restored image and the original image at (i, j), and M and N are the number of rows and columns of the image.

The specific calculation formula of the PSNR is as follows:

$$PSNR = 10 \times \log_{10}\left(\frac{L^2}{MSE}\right)$$

where L denotes the maximum gray level of the image. For 8-bit grayscale images, L is equal to 255.

The expression for SSIM is:

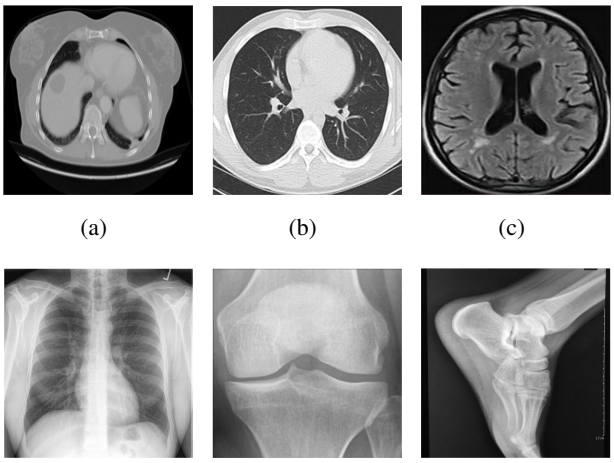
$$SSIM = \frac{(2\mu_f\mu_u + c_1)(2\sigma_{fu} + c_2)}{(\mu_f^2 + \mu_u^2 + c_1)(\sigma_f^2\sigma_u^2 + c_2)}$$

where μ_f and μ_u are the means of f and u, respectively; σ_f^2, σ_u^2 are the variances of f and u, respectively, σ_{fu} is the covariance of f and u, and c_1 and c_2 are constants used to maintain stability.

5.1. Analysis of medical image denoising performance with ISSM

5.1.1. Different figure and noise level test

As shown in Figure 1, our numerical experiments use six real medical images, which are called “Adenocarcinoma”, “Chest”, “Brain”, “CXR”, “Knee”, and “Foot”.



(a) (b) (c)

(d) (e) (f)

Figure 1. (a) The original "Adenocarcinoma" image; (b) the original "Chest" image; (c) the original "Brain" image; (d) the original "CXR" image; (e) the original "Knee" image; (f) the original "Foot" image.

First, a more thorough analysis is given. We have compared the denoising effect of different sizes of the same image at the same noise level $(25, 0.1)$ using the ISSM denoising method of this paper and the results are shown in Table 1. According to Table 1, we can find that as the image size increases, the PSNR value after denoising also increases. As the image size is larger, the number of pixels contained in the image is more and the observed image is also clearer. The result is that image processing takes longer. Based on the above considerations, we chose an image of size 512*512 for the experiment in this paper.

In the first experiment, three images with a size of $512 * 512$ were destroyed by different degrees of noise: $(k_0, k_1) = (10, 0), (25, 0), (25, 0.1), (25, 0.2), (50, 0)$. The PSNR values of the image contaminated by noise and the PSNR values after denoising are shown in Table 2. After a lot of experiments, in order to make the noise reduction effect better, we chose the parameters $dt = 0.03$, $\beta = 0.01, \alpha_1 = 0.01, \alpha_2 = 0.005$, and the number of iterations equal to 800. It can be seen that the PSNR values of the three images have been improved after denoising by our model.

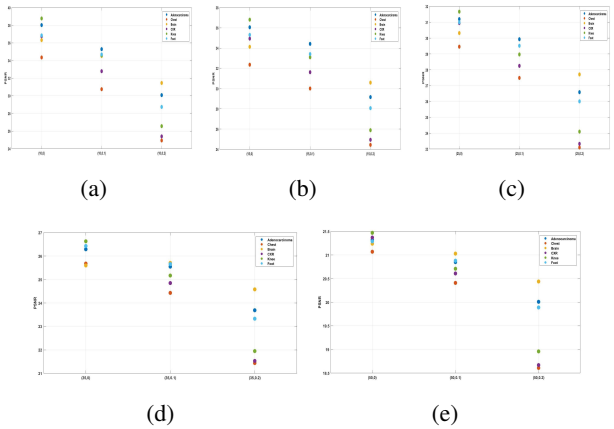
Figure 2 shows the PSNR values of the six medical CT images tested in Figure 1 at different noise levels, from which we can see that the denoising effect of these six images varies at the same noise level. Among them, under the noise level of $(10, 0)$ and $(25, 0)$, image "Knee" has the highest PSNR value after denoising and the best recovery effect; under the noise levels of $(25, 0.1)$, the PSNR of images "Adenocarcinoma", "Brain", "Knee", "Foot", are basically about the same, but image "Brain" is slightly higher; under the noise level of $(25, 0.2)$, the PSNR of image "Brain" is the highest and the denoising effect is the best; under the noise level of $(50, 0)$, image "Adenocarcinoma" has the highest PSNR value, which is not distinctly high compared with other tested images.

Table 1. Comparison results of PSNR values for different image sizes.

Image / Size	Adenocarcinoma		Chest	
	Noisy	Denoised	Noisy	Denoised
128*128	19.30	27.05	18.63	25.08
256*256	19.30	28.76	18.63	26.24
512*512	19.30	29.52	18.63	27.75
1024*1024	19.30	30.53	18.63	28.66
Image / Size	**Brain**		**CXR**	
	Noisy	Denoised	Noisy	Denoised
128*128	19.68	25.18	18.48	27.01
256*256	19.68	27.68	18.48	27.91
512*512	19.68	29.86	18.48	28.62
1024*1024	19.68	30.42	18.48	28.93
Image / Size	**Knee**		**Foot**	
	Noisy	Denoised	Noisy	Denoised
128*128	18.64	28.11	19.32	28.10
256*256	18.64	28.67	19.32	28.81
512*512	18.64	29.33	19.32	29.79
1024*1024	18.64	29.44	19.32	30.31

Table 2. The PSNR and SSIM values of noisy and denoised images.

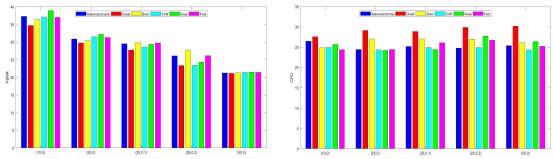
Image	(k_0, k_1)	noisy PSNR/SSIM	denoised PSNR/SSIM
Adenocarcinoma	(10,0)	28.13/0.2650	37.29/0.3986
	(25,0)	20.16/0.1383	30.92/0.3237
	(25,0.1)	19.30/0.1271	29.52/0.2962
	(25,0.2)	17.25/0.1080	26.13/0.2358
	(50,0)	14.14/0.0727	21.96/0.1496
Chest	(10,0)	28.13/0.4677	34.72/0.6835
	(25,0)	20.16/0.2190	29.73/0.4887
	(25,0.1)	18.63/0.1872	27.75/0.4370
	(25,0.2)	15.88/0.1454	23.36/0.3434
	(50,0)	14.15/0.1056	21.16/0.2386
Brain	(10,0)	28.13/0.4351	36.42/0.7259
	(25,0)	20.19/0.2071	30.40/0.5453
	(25,0.1)	19.68/0.1936	29.86/0.5209
	(25,0.2)	18.51/0.1696	27.79/0.4599
	(50,0)	14.16/0.0897	21.32/0.2404
CXR	(10,0)	28.12/0.2974	37.14/0.6335
	(25,0)	20.19/0.0881	31.50/0.4123
	(25,0.1)	18.48/0.0694	28.62/0.3212
	(25,0.2)	15.54/0.0458	23.44/0.2053
	(50,0)	14.16/0.0272	21.44/0.1096
Knee	(10,0)	28.13/0.1884	38.94/0.5220
	(25,0)	20.18/0.0533	32.18/0.3122
	(25,0.1)	18.64/0.0408	29.33/0.2286
	(25,0.2)	15.91/0.0259	24.31/0.1261
	(50,0)	14.16/0.0167	21.51/0.0685
Foot	(10,0)	28.12/0.2627	37.01/0.3787
	(25,0)	20.17/0.0903	31.30/0.2920
	(25,0.1)	19.32/0.0764	29.79/0.2527
	(25,0.2)	17.42/0.0586	26.17/0.1837
	(50,0)	14.16/0.0338	21.42/0.1003



(a) (b) (c) (d) (e)

Figure 2. The PSNR of the tested images at different noise levels (k_0, k_1).

We plot the PSNR and running time (CPU) of the tested images after denoising at different noise levels as bar charts (as shown in Figure 3), where 3(a) is the PSNR after denoising and 3(b) is the running time (CPU). The dark blue, red, yellow, light blue, green, and purple bars in the figure represent images "Adenocarcinoma", "Chest", "Brain", "CXR", "Knee", and "Foot", respectively. From 3(a), it can be found that the denoising ability of different images is different for different noise levels. From 3(b), it can be seen that the denoising runtime for the "Chest" image is longer for all the different noise levels, and the runtime for the other five images is basically the same. We can also see from 3(b) that there is not much difference in the running time when the noise level is relatively small; however, when the denoising level is large, the difference in the running time is more obvious, and the running time of the X-ray (1(d)–1(f)) is more similar compared to that of the CT image (1(a)–1(c)).



(a) (b)

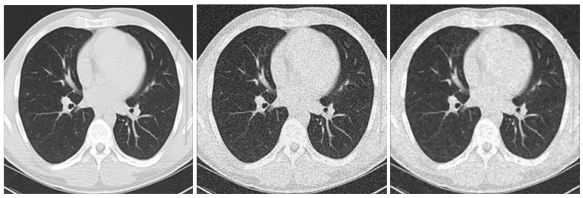
Figure 3. (a) PSNR of the tested images at different noise levels (k_0, k_1); (b) CPU of the tested images at different noise levels (k_0, k_1).

5.1.2. Different perspectives

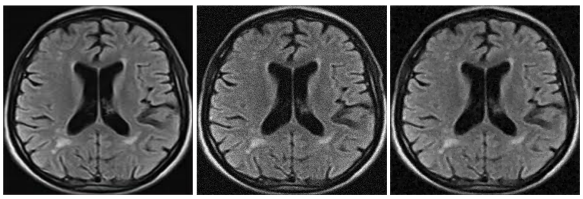
In the second experiment, we illustrate from all aspects that our model has good performance. We take $(k_0, k_1) = (25, 0.1)$ to denoise the image in Figure 1. The choice of parameters remains $dt = 0.03, \beta = 0.01, \alpha_1 = 0.01, \alpha_2 = 0.005$.

Figure 4 shows that our proposed model can remove the mixture of additive and multiplicative noise to some extent.

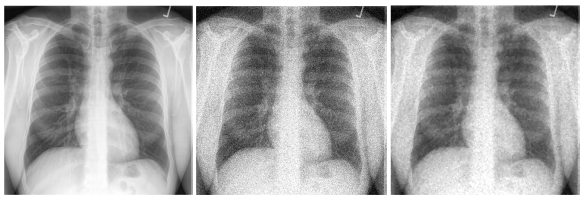
(a) (b) (c)



(d) (e) (f)



(g) (h) (i)



(j) (k) (l)

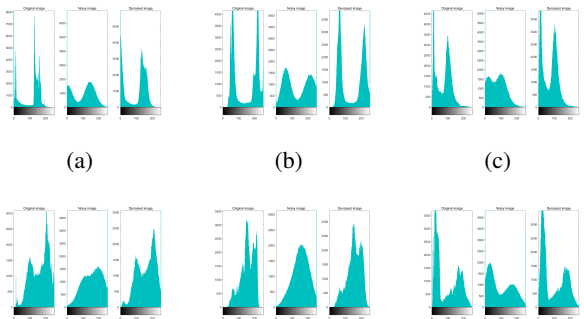
(m) (n) (o)

(p) (q) (r)

Figure 4. (a)–(c) The original image, the noisy image, and the denoised image of "Adenocarcinoma", respectively. (d)–(f) The original image, the noisy image, and the denoised image of "Chest", respectively. (h)–(g) The original image, the noisy image, and the denoised image of "Brain", respectively. (k)–(m) The original image, the noisy image, and the denoised image of "CXR", respectively. (n)–(p) The original image, the noisy image, and the denoised image of "Knee", respectively. (q)–(s) The original image, the noisy image, and the denoised image of "Foot", respectively.

In Figure 5, the histogram of the image shows the distribution of different gray levels in the total pixels. The horizontal axis represents the value of the pixel, and the vertical axis represents the number of times that the corresponding pixel value appears. It can be seen that the histogram of the image destroyed by noises is completely different from the original image, but the histogram of the denoised image is very close to the original image.

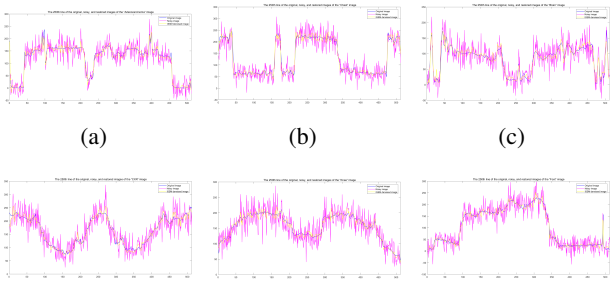
In Figure 6, we show the 250th line of the tested images. The original image, noisy image, and denoised image are represented by blue, purple, and yellow solid lines, respectively. From these figures, we can see that when the original image is contaminated by noise, the one-dimensional image oscillates greatly. The yellow solid line is highly consistent with the blue, which represents that the denoising effect of our proposed model is very effective.



(a) (b) (c)

(d) (e) (f)

Figure 5. (a) Histogram of the "Adenocarcinoma" image; (b) histogram of the "Chest" image; (c) histogram of the "Brain" image; (d) histogram of the "CXR" image; (e) histogram of the "Knee" image; (f) histogram of the "Foot" image, where the left is the original image, the center is the noisy image, and the right is the denoised image.

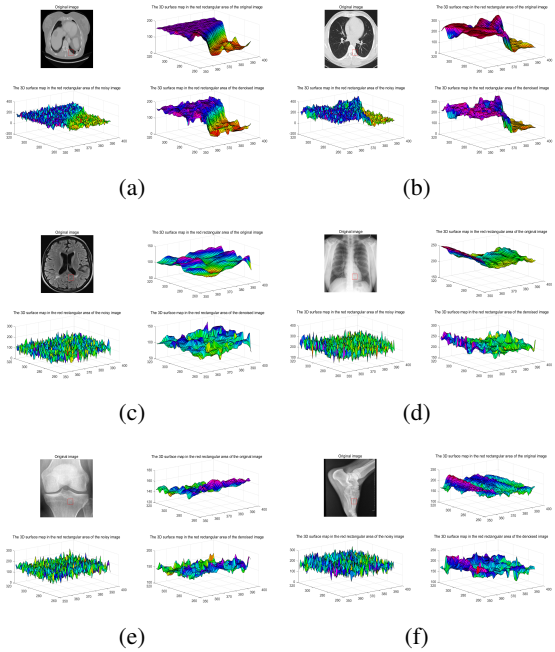


(a) (b) (c)

(d) (e) (f)

Figure 6. (a) The 250th line of the "Adenocarcinoma" image; (b) the 250th line of the "Chest" image; (c) the 250th line of the "Brain" image; (d) the 250th line of the "CXR" image; (e) the 250th line of the "Knee" image; (f) the 250th line of the "Foot" image.

In Figure 7, we look at the same areas of the tested image and mark them with a red rectangle. The three-dimensional surface map of the original image is relatively flat, and that of the noisy image is rugged and accompanied by many sharp protrusions. It can be clearly seen that the three-dimensional surface map is highly consistent with the original image, which concludes that our proposed model has a good denoising effect.



(a) (b)

(c) (d)

(e) (f)

Figure 7. (a) The 3D surface map of the "Adenocarcinoma" image; (b) the 3D surface map of the "Chest" image; (c) the 3D surface map of the "Brain" image; (d) the 3D surface map of the "CXR" image; (e) the 3D surface map of the "Knee" image; (f) the 3D surface map of the "Foot" image. In each figure, the top left corner is the original image with the red rectangular box, the top right corner is its corresponding 3D surface map, the bottom left corner is the 3D surface map with noise, and the bottom right corner is the 3D surface map after denoising.

5.1.3. α_1 and α_2-dependent test

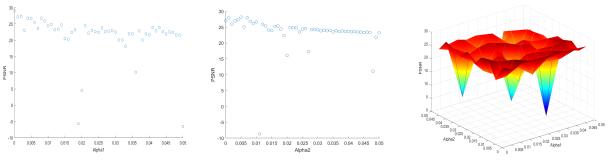
In the third experiment, we first selected several different sets of parameter values and conducted experiments, and the results are shown in the following Table 3. According to Table 3, we can find that different parameter values affect the denoising effect (PSNR value and SSIM value), and therefore, to further find the suitable parameter values, we utilize a grid search for them.

We examine the impact of changing the values of $\alpha_1 \in [0.001, 0.05]$ and $\alpha_2 \in [0.001, 0.05]$ on the proposed variational method. Our method is executed with the parameters $dt = 0.03, \beta = 0.01$, the number of iterations $N = 700$, and a noise level equal to $(25, 0.1)$. As clearly seen

in Figure 8(a), we set $\alpha_2 = 0.005$, according to the PSNR value, and it can be seen that when $\alpha_1 \in [0.001, 0.015]$, the effect of removing this type of noise is excellent. From 8(b), we set $\alpha_1 = 0.003$, and it can be seen that when $\alpha_2 \in [0.001, 0.01]$, the effect of removing this type of noise is excellent. In Figure 8(c), we show a three-dimensional plot of the changes in the PSNR values corresponding to changes in the regularization parameters α_1 and α_2. The darker the red, the higher the PSNR value, and the red areas are those with higher PSNR values, which we find are consistent with the conclusions reflected in 8(a) and 8(b). For more details, see Figure 8.

Table 3. Comparison results of PSNR and SSIM values with different parameters α_1 and α_2.

α_1	α_2	Noisy		Denoised	
		PSNR	SSIM	PSNR	SSIM
0.001	0.005	19.36	0.1156	29.88	0.2851
0.005	0.005	19.38	0.1160	29.94	0.2842
0.01	0.005	19.35	0.1155	29.90	0.2840
0.05	0.005	19.34	0.1157	24.95	0.2830
α_1	α_2	Noisy		Denoised	
		PSNR	SSIM	PSNR	SSIM
0.01	0.001	19.36	0.1157	28.75	0.2850
0.01	0.005	19.36	0.1156	29.97	0.2831
0.01	0.01	19.34	0.1152	29.86	0.2818
0.01	0.05	19.36	0.1157	29.82	0.2829



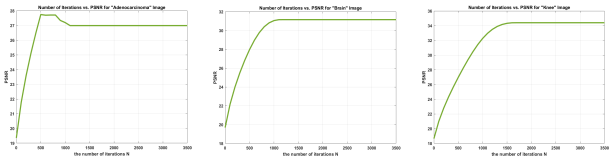
(a) (b) (c)

Figure 8. (a) The PSNR value of the restored image with different values of α_1; (b) the PSNR value of the restored image with different values of α_2; (c) the PSNR value of the restored image with different values of α_1 and α_2.

From Figure 8, we can further see that although the size of parameters α_1 and α_2 do not have a significant impact on the PSNR value after image denoising as a whole, there are individual parameter values that will make the PSNR value drop dramatically, or even become negative. In order to make the model's denoising effect more robust and accurate, we need to discuss the parameters.

5.1.4. N-dependence test

In the fourth experiment, in order to decide the value of the N, we study the relation between the PSNR and the number of the iterations N. The Figure 9 shows the relation between the PSNR and the number of iterations N on "Adenocarcinoma" , "Brain", and "Knee" images corrupted by the combined additive and multiplicative noise with noise intensity $(k_0, k_1) = (25, 0.1)$. Figure 9 shows that the curves of the relationship between the PSNR and number of iterations are slightly different from other images. In 9(a), we can see that the curve first shows an upward trend and then decreases, but finally reaches a steady state; in 9(b) and 9(c), we can see that the curve shows an upward trend and finally reaches a steady state. Therefore, for images "Adenocarcinoma" and "Brain", the optimal number of iterations N for removing the mixed noise is in the interval $(1000, 1250)$, and for image "Knee", the optimal number of iterations N for removing the mixed noise is around 1500.



(a) (b) (c)

Figure 9. The relationship between PSNR and the number of iterations N. (a) The "Adenocarcinoma" image; (b) the "Brain" image; (c) the "Knee" image.

5.2. ISSM and GDE denoising performance analysis

5.2.1. Different algorithms

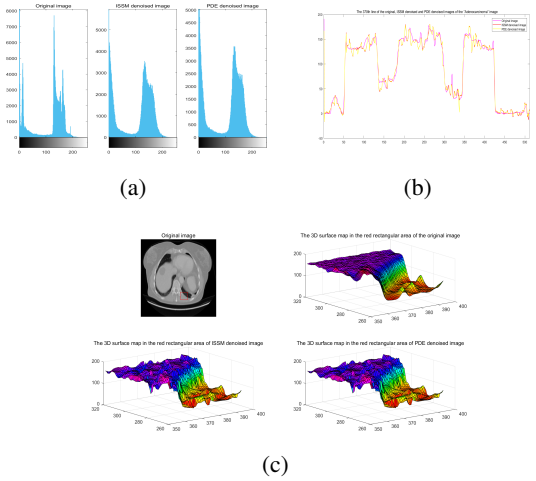
In the following, we still test the six images of Figure 1. A noise of $(k_0, k_1) = (25, 0.1)$ is added to the images and denoised with gradient descent (GDE) and the inverse scale space algorithm (ISSM), which shows the PSNR and

running time (CPU) after denoising as shown in Table 4 where higher PSNR values and smaller CPUs are in bold. In this experiment, the parameter values are tuned as $dt = 0.03, \beta = 0.01, (\alpha_1, \alpha_2) = (0.01, 0.005)$, and the number of iterations is equal to 500. In Table 4, we can find that the PSNR and SSIM values of the ISSM method are slightly higher than those of the GDE, and for images 1(a),1(c),1(f), the running time of the GDE is slightly longer than that of the ISSM, whereas for images 1(b),1(d),1(e), the running time of the GDE is about 2 seconds longer than that of the ISSM. In general the ISSM is a little bit shorter in terms of running time than the GDE. Therefore the method in this paper is better than the GDE in terms of PSNR, SSIM, and runtime.

Table 4. Comparison results of PSNR, SSIM values, and runtime (CPU) of the two methods.

Image	**PSNR**		
	Noisy	GDE	ISSM
Adenocarcinoma	19.30	28.34	**29.52**
Chest	18.63	27.51	**27.75**
Brain	19.68	29.38	**29.86**
CXR	18.48	28.29	**28.62**
Knee	18.64	28.96	**29.33**
Foot	19.32	29.24	**29.79**
Image	**SSIM**		
	Noisy	GDE	ISSM
Adenocarcinoma	0.1271	0.2923	**0.2962**
Chest	0.1872	0.4317	**0.4370**
Brain	0.1936	0.5095	**0.5209**
CXR	0.0694	0.3103	**0.3212**
Knee	0.0408	0.2174	**0.2286**
Foot	0.0764	0.2479	**0.2527**
Image	**Time**		
		GDE	ISSM
Adenocarcinoma		25.7696	**25.1879**
Chest		30.0489	**28.8607**
Brain		27.0294	**27.0153**
CXR		27.5674	**24.9295**
Knee		27.6340	**24.4423**
Foot		26.8773	**26.0720**

We used the "Adenocarcinoma" image to compare the ISSM method with the GDE method in detail. As shown in Figure 10, 10(a) shows that the GDE denoising method yields a greater number of low-pixel values than the ISSM method. 10(b) shows that the ISSM method and the GDE method are essentially identical at line 370th of the image after denoising. However, in certain instances, the ISSM method outperforms the GDE method. As shown in 10(c), the 3D surface maps of the images denoised by the ISSM method and the GDE method are basically the same. Thus, Figure 10 as a whole reflects that the inverse scale space method (ISSM) is superior to the gradient descent method (GDE).



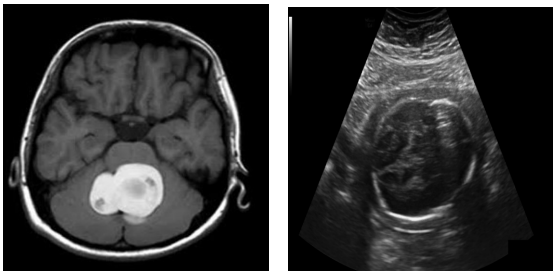
(a) (b)

(c)

Figure 10. (a) The histogram of the "Adenocarcinoma" image; (b) the 370th line of the "Adenocarcinoma" image; (c) the 3D surface map of the "Adenocarcinoma" image.

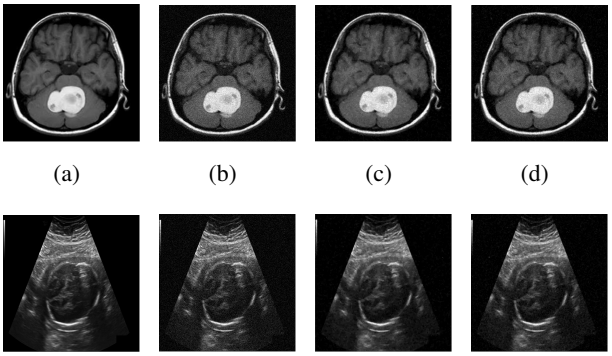
5.2.2. Testing of other types of medical images

In this subsection of the experiment, we tested brain tumour MRI and fetal head ultrasound images, as shown in Figure 11. The original image, the noisy image, the ISSM denoised image, and the GDE denoised image are shown from left to right in Figure 12, where the noise level of the noisy image is $(25, 0.1)$. As can be seen from the figure, visually we can see very little difference with the naked eye, so we analyzed and compared them again from a quantitative point of view, as shown in Table 5.



(a) (b)

Figure 11. (a) Brain tumour MRI image; (b) fetal head ultrasound image.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 12. Original image, noisy image, and image after denoising by different methods.

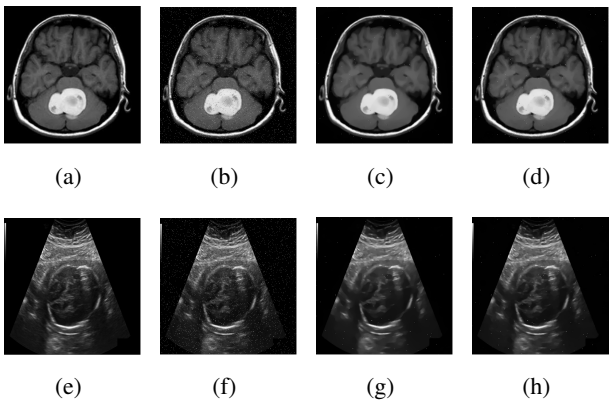
Table 5. Comparison results of PSNR, SSIM values, and runtime (CPU) of the two methods.

Image	**PSNR**		
	Noisy	GDE	ISSM
MRI	19.76	26.65	**30.30**
Ultrasound	19.98	25.76	**30.40**
Image	**SSIM**		
	Noisy	GDE	ISSM
MRI	0.1271	0.2601	**0.3390**
Ultrasound	0.1117	0.2330	**0.2836**
Image	**Time**		
		GDE	ISSM
MRI		25.9930	**24.7379**
Ultrasound		27.5113	**27.2909**

In Table 5, we can clearly see that the PSNR and SSIM values obtained by our method in this paper are higher than those of GDE, which indicates that our method can effectively denoise another two types of images, MRI and ultrasound, besides X-ray and CT; and according to the comparison of the last two columns of the running time, it can be seen that our method has a slight improvement in comparison with GDE, which further illustrates the advantages of our method for the adaptability of medical image denoising.

5.2.3. Other noise type denoising

In this section of the experiment, we add salt and pepper noise with a noise level of 0.02 to the image in Figure 11 and denoise it with the method of this paper and GDE, and the results are shown in Figure 13 and Table 6. From the figure, we can find that both this paper's method and GDE can effectively remove part of the salt and pepper noise, and ISSM is better than the GDE denoising effect, but both ISSM and the GDE method cannot effectively retain the texture details. From Table 6, it can be seen that the PSNR value after denoising is not high, and there is not much difference between the two methods.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 13. Original image, noisy image, and image after denoising by different methods.

Table 6. Comparison results of PSNR values of the two methods.

Image	PSNR		
	Noisy	GDE	ISSM
MRI	9.86	10.23	**10.26**
Ultrasound	13.06	14.02	**14.09**

6. Conclusions

In this paper, we combine the data fidelity terms from the ROF model and the I-divergence model to restore medical images with combined noises. We have proved the existence and uniqueness of the minimum solution of the model. The inverse scale space method has been used to solve the non-linear PDE resulting from minimizing the proposed functional. Experimental results demonstrate that the proposed model can remove combined noise on medical images from different perspectives. The inverse scale space method on this model improves PSNR better and has faster runtime than traditional methods on medical images. Parameter sensitivity is also discussed.

Medical imaging is often affected by noise, resulting in image quality degradation, while the model and method proposed in this paper can effectively denoise medical images, thus enabling doctors to more accurately identify lesions and reduce the possibility of misdiagnosis and omission of diagnosis, thus improving diagnostic accuracy and enhancing lesion detection. It can also help patients to reduce the number of repeated examinations and enhance the patient experience. Learning-based methods are currently more popular research. Combining the modeling approach of this paper with deep learning as well as further exploring real medical images with noise are directions for future work.

Author contributions

Conceptualization, Chenwei Li and Donghong Zhao; methodology, Chenwei Li and Donghong Zhao; software, Chenwei Li; formal analysis, Chenwei Li and Donghong Zhao; writing—original draft preparation, Chenwei Li; writing—review and editing, Chenwei Li and Donghong Zhao; visualization, Chenwei Li. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors thank the reviewers for their helpful comments and valuable suggestions. This research is funded by the National Natural Science Foundation of China (grant number 12371481) and Youth Teaching Talents Training Program of USTB (2016JXGGRC-002).

Conflict of interest

All authors declare no conflicts of interest.

References

1. G. Aubert, P. Kornprobst, *Mathematical problems in image processing—partial differential equations and the calculus of variations*, 2 Eds., Applied Mathematical Sciences, 2006. http://dx.doi.org/10.1007/978-0-387-44588-5
2. D. L. Donoho, I. M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, *J. Am. Stat. Assoc.*, **90** (1995), 1200–1224. http://dx.doi.org/10.1080/01621459.1995.10476626
3. S. Osher, M. Burger, D. Goldfarb, J. Xu, W. Yin, An iterative regularization method for total variation-based image restoration, *Multiscale Model. Simul.*, **4** (2005), 460–489. http://dx.doi.org/10.1137/040605412
4. X. Yu, D. Zhao, A weberized total variance regularization-based image multiplicative noise model, *Image Anal. Stereol.*, **42** (2023), 65–76. http://dx.doi.org/10.5566/ias.2837
5. P. Kornprobst, R. Deriche, G. Aubert, Image sequence analysis via partial differential equations, *J. Math. Imaging Vis.*, **11** (1999), 5–26. http://dx.doi.org/10.1023/A:1008318126505

6. S. M. A. Sharif, R. A. Naqvi, Z. Mehmood, J. Hussain, A. Ali, S. W. Lee, Meddeblur: medical image deblurring with residual dense spatial-asymmetric attention, *Mathematics*, **11** (2023), 115. http://dx.doi.org/10.3390/math11010115

7. R. A. Naqvi, A. Haider, H. S. Kim, D. Jeong, S. W. Lee, Transformative noise reduction: leveraging a transformer-based deep network for medical image denoising, *Mathematics*, **12** (2024), 2313. http://dx.doi.org/10.3390/math12152313

8. S. Umirzakova, S. Mardieva, S. Muksimova, S. Ahmad, T. Whangbo, Enhancing the super-resolution of medical images: introducing the deep residual feature distillation channel attention network for optimized performance and efficiency, *Bioengineering*, **10** (2023), 1332. http://dx.doi.org/10.3390/bioengineering10111332

9. M. Shakhnoza, S. Umirzakova, M. Sevara, Y. I. Cho, Enhancing medical image denoising with innovative teacher student model-based approaches for precision diagnostics, *Sensors*, **23** (2023), 9502. http://dx.doi.org/10.3390/s23239502

10. L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Phys. D*, **60** (1992), 259–268. http://dx.doi.org/10.1016/0167-2789(92)90242-F

11. A. Boukdir, M. Nachaoui, A. Laghrib, Hybrid variable exponent model for image denoising: a nonstandard high-order pde approach with local and nonlocal coupling, *J. Math. Anal. Appl.*, **536** (2024), 128245. http://dx.doi.org/10.1016/j.jmaa.2024.128245

12. W. Lian, X. Liu, Non-convex fractional-order TV model for impulse noise removal, *J. Comput. Appl. Math.*, **417** (2023), 114615. http://dx.doi.org/10.1016/j.cam.2022.114615

13. J. Shen, On the foundations of vision modeling: I. Webers law and Weberized TV restoration, *Phys. D*, **175** (2003), 241–251. http://dx.doi.org/10.1016/S0167-2789(02)00734-0

14. C. Liu, X. Qian, C. Li, *A texture image denoising model based on image frequency and energy minimization*, Springer Berlin Heidelberg, 2012, 939–949. http://dx.doi.org/10.1007/978-3-642-34531-9_101

15. C. Li, D. Zhao, A non-convex fractional-order differential equation for medical image restoration, *Symmetry*, **16** (2024), 258. http://dx.doi.org/10.3390/sym16030258

16. J. S. Lee, Digital image enhancement and noise filtering by use of local statistics, *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-2** (1980), 165–168. http://dx.doi.org/10.1109/TPAMI.1980.4766994

17. A. Achim, A. Bezerianos, P. Tsakalides, Novel bayesian multiscale method for speckle removal in medical ultrasound images, *IEEE Trans. Med. Imaging*, **20** (2001), 772–783. http://dx.doi.org/10.1109/42.938245

18. L. Rudin, P. L. Lions, S. Osher, Multiplicative denoising and deblurring: theory and algorithms, In: *Geometric level set methods in imaging, vision, and graphics*, New York: Springer, 2003, 103–119. https://dx.doi.org/10.1007/0-387-21810-6_6

19. G. Aubert, J. F. Aujol, A variational approach to removing multiplicative noise, *SIAM J. Appl. Math.*, **68** (2008), 925–946. http://dx.doi.org/10.1137/060671814

20. J. Shi, S. Osher, A nonlinear inverse scale space method for a convex multiplicative noise model, *SIAM J. Imaging Sci.*, **1** (2008), 294–321. http://dx.doi.org/10.1137/070689954

21. G. Steidl, T. Teuber, Removing multiplicative noise by douglas-rachford splitting methods, *J. Math. Imaging Vis.*, **36** (2010), 168–184. http://dx.doi.org/10.1007/s10851-009-0179-5

22. L. Xiao, L. L. Huang, Z. H. Wei, A weberized total variation regularization-based image multiplicative noise removal algorithm, *EURASIP J. Adv. Signal Process.*, **2010** (2010), 1–15. http://dx.doi.org/10.1155/2010/490384

23. L. Huang, L. Xiao, Z. Wei, A nonlinear inverse scale space method for multiplicative noise removal based on weberized total variation, *2009 Fifth International Conference on Image and Graphics*, Xi'an, China, 2009, 119–123. http://dx.doi.org/10.1109/ICIG.2009.19

24. X. Liu, T. Sun, Hybrid non-convex regularizers model for removing multiplicative noise, *Comput. Math. Appl.*, **126** (2022), 182–195. http://dx.doi.org/10.1016/j.camwa.2022.09.012

25. C. Li, C. He, Fractional-order diffusion coupled with integer-order diffusion for multiplicative noise removal, *Comput. Math. Appl.*, **136** (2023), 34–43. http://dx.doi.org/10.1016/j.camwa.2023.01.036

26. K. Hirakawa, T. W. Parks, Image denoising using total least squares, *IEEE Trans. Image Process.*, **15** (2006), 2730–2742. http://dx.doi.org/10.1109/TIP.2006.877352

27. N. Chumchob, K. Chen, C. Brito-Loeza, A new variational model for removal of combined additive and multiplicative noise and a fast algorithm for its numerical approximation, *Int. J. Comput. Math.*, **90** (2013), 140–161. http://dx.doi.org/10.1080/00207160.2012.709625

28. A. Ullah, W. Chen, M. A. Khan, H. G. Sun, An efficient variational method for restoring images with combined additive and multiplicative noise, *Int. J. Appl. Comput. Math.*, **3** (2017), 1999–2019. http://dx.doi.org/10.1007/s40819-016-0219-y

29. Y. Chen, W. Feng, R. Ranftl, H. Qiao, T. Pock, A higher-order mrf based variational model for multiplicative noise reduction, *IEEE Signal Process Lett.*, **21** (2014), 1370–1374. http://dx.doi.org/10.1109/LSP.2014.2337274

30. P. Ochs, Y. Chen, T. Brox, T. Pock, ipiano: inertial proximal algorithm for nonconvex optimization, *SIAM J. Imaging Sci.*, **7** (2014), 1388–1419. http://dx.doi.org/10.1137/130942954

31. O. Scherzer, C. Groetsch, Inverse scale space theory for inverse problems, In: *Scale-space and morphology in computer vision*, Springer, Berlin, Heidelberg, **2106** (2001), 317–325. http://dx.doi.org/10.1007/3-540-47778-0_29

32. C. W. Groetsch, O. Scherzer, Nonstationary iterated tikhonovmorozov method and third order differential equations for the evaluation of unbounded operators, *Math. Methods Appl. Sci.*, **23** (2000), 1287–1300.

33. M. Burger, S. Osher, J. Xu, G. Gilboa, Nonlinear inverse scale space methods for image restoration, In: *Variational, geometric, and level set methods in computer vision*, Springer, Berlin, Heidelberg, **3752** (2005), 25–36. http://dx.doi.org/10.1007/11567646_3

34. M. Burger, G. Gilboa, S. Osher, J. Xu, Nonlinear inverse scale space methods, *Commun. Math. Sci.*, **4** (2006), 179–212. http://dx.doi.org/10.4310/CMS.2006.v4.n1.a7

35. I. Csiszar, Why least squares and maximum entropy? an axiomatic approach to inference for linear inverse problems, *Ann. Statist.*, **19** (1991), 2032–2066. http://dx.doi.org/10.1214/AOS/1176348385

AIMS Press