Image Anal Stereol 2023;42:65-76 doi: 105566/ias.2837
Original Research Paper

A WEBERIZED TOTAL VARIANCE REGULARIZATION-BASED IMAGE MULTIPLICATIVE NOISE MODEL

XINYAO YU AND DONGHONG ZHAO✉
University of Science and Technology Beijing, 30 Xueyuan Road, Haidian District, Beijing, 100083
e-mail: yuxinyao1998@qq.com; zdh751111@ustb.edu.cn.
(Received November 24, 2022, revised April 13, 2023; accepted June 11, 2023)

ABSTRACT

This paper considers Weber's law and proposes a new non-convex model for images contaminated by Gaussian noise and Rayleigh noise. The alternating direction method of multipliers (abbreviated as ADMM) is a recent popular method that can handle convex and non-convex problems well. This paper compares denoising effect between ADMM and the Euler-Lagrange equation method applied to the non-convex model. The numerical experimental results show that ADMM performs better and has a higher Peak Signal to Noise Ratio.

Keywords: ADMM, Euler-Lagrange equation, image denoising, multiplicative noise, partial differential equation, Weberized total variation.

INTRODUCTION

Image processing is a rapidly growing discipline that has a wide range of applications in many fields. The image has become an effective tool for studying visual perception in psychology, physiology, computer science and other aspects. In addition, image processing plays a pivotal role in scientific research, industrial production, medical and health care, education, management and other large-scale applications, especially with the rapid development of multimedia technology, images and human life are more inseparable. The fast progress of computer science, the popularity of image digitization and image display equipment also provides good conditions for the development of image processing and become the main driving force for its development.

Image denoising is one of important problems in image restoration, where the key is to remove noise while preserving image features such as image edges. The total variation image model was originally introduced to solve the image denoising problem. The success of the total variation denoising model lies in the utilization of the inherent regularity of the natural image, which can easily reflect the geometric regularity of the real image, such as the smoothness of the boundary

Compared to additive Gaussian white noise, multiplicative noise still conforms to Rayleigh or Gamma distribution functions. Multiplicative noise is a serious contamination to the image, and it is difficult to process multiplicative noise images effectively because of the sharp fluctuations and low uniformity of multiplicative noise. Therefore, multiplicative noise is chosen to be removed. (Wu *et al.*, 2008)

Let u be the original clear image, f be the observed image contaminated by noise and v be noise. The noise model can be expressed as

$$f = u \cdot v \tag{1}$$

Most traditional image processing methods rarely take into account the influence of human visual psychology. The Weber's law suggests that human reaction and perception to visual signal intensity fluctuations are obtained by weighting the background stimulus rather than being apparently uniform. In this paper, an attempt is made to develop an image recovery model aiming to use one of the most famous psychological results, namely Weber's law. Its main mathematical properties, such as existence and uniqueness, are investigated, as well as the computational method of the associated nonlinear partial differential equations. (Shen, 2003)

Weber's law was first proposed (Weber, 1834) in 1834. It was later rationally elaborated by Fechner . This law reveals the general effect of human sensitivity to intensity increment δu under background stimulus u , i.e., JND (just-noticeable-difference). The Weber fraction is a constant: $\frac{\delta u}{u} = const$.

Many experiments have shown that for a large range of stimulus u , Weber's law does provide a good approximation. Applying Weber's law in visual perception, u

representing background light intensity, δu represents intensity fluctuation.

There are many methods to process image models, such as the Euler-Lagrange equation method, nonlinear inverse scale space methods, and Bregman iterative regularization algorithms. In this paper, we will use ADMM to process the model.

The structure of this paper is as follows. Some preliminaries about model and alternating are summarized, in Section Introduction. In Section Material and Methods, we propose a new non-convex model, and prove the uniqueness of the existence of the minimum value. Then, the calculation method of Weberized total variation minimization is discussed. In Section Results, numerical experimental results are given. Finally, we give some conclusions in Section Discussion.

Model review

In this subsection, we briefly review the classical multiplicative noise cancellation problem and the Weberized total variation restoration

Classical multiplicative noise model

In 2003, Rudin, Lions, and Osher (Rudin *et al.*, 2003) firstly proposed the total variation model with multiplicative noise removed, which is a multiplicative version of the ROF model. We call it the RLO model. In this model, the noise image can be expressed as $f = u \cdot v$. Given two constraints that the mean of noise v is 1 and the variance is 0. Minimizing $\int_\Omega |\nabla u|$ under these two constraints, the RLO model is obtained as follows:

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

The RLO model is a classical total variation model for Gaussian multiplicative noise. The model can remove the Gaussian multiplicative noise to some extent. However, because the model is proposed by combining the statistical properties of Gaussian multiplicative noise, it is not universal to other multiplicative noise.

As the research progressed, researchers gradually found that most multiplicative noise obeys the gamma distribution rather than the Gaussian distribution. In 2008, Aubert and Aujol (Aubert and Aujol, 2008) assumed that the noise is gamma noise with mean 1, derived a data fidelity term based on the maximum a posteriori estimate. They proposed a non-convex model for removing gamma noise, called as AA model,

$$E(u) = \int_\Omega |\nabla u| dx + \lambda \int_\Omega \left(\log u + \frac{f}{u} \right) dx \qquad (3)$$

And a sufficient condition for the existence of uniqueness of this model is given, $u \in (0, 2f)$.

The AA model can effectively remove gamma multiplicative noise. However, the fidelity term of the AA model is non-convex, which is prone to block effect in the denoising process and the denoising effect is poor.

Shi and Osher (Shi and Osher, 2008) used the corresponding relaxed inverse scale spatial flow as a denoising technique and used the data fidelity term of the AA model. The regular term part is $|\log u|_{BV}$ instand $|u|_{BV}$. They proposed a convex variational model called as SO model:

$$E(w) = \int_\Omega |w| dx + \lambda \int_\Omega \left[afe^{-w} + \frac{b}{2} f^2 e^{-2w} + (a+b)w \right] dx, \qquad (4)$$

where $w = \log(u)$.

Inspired by the AA model, in 2010, Steidl and Teuber evolved the I-divergence model (Steidl and Teuber, 2010). I-divergence is called the generalized Kullback-Leibler divergence. The I-divergence of f and u is $I(f,u) = \int_\Omega \left(f \log u \frac{f}{u} - f + u \right) dx$. Since f is the observed image, neglecting the constant term reduces to $\int_\Omega (u - f \log u) dx$. The model takes the generalized I-divergence, as the fidelity term and the full variance as the regular term, and the mathematical expression is as follows:

$$E(u) = \int_\Omega |\nabla u| dx + \int_\Omega (u - f \log u) dx \qquad (5)$$

In this paper, we will propose a new model based on RLO model, I-divergence model and Weber's law.

Weberized total variation restoration

Most traditional image restoration models do not consider that the visual sensitivity of local fluctuations depends on the intensity of the ambient level. If the local variation is 0.01 and the background intensity is 0.1, the fluctuation level is more significant than when the background intensity is 0.8.

In 2003, Jianhong Shen (Shen, 2003) combined Weber's law with the ROF model and proposed the Weberized classical total variation model with removal of additive noise:

$$E(u) = \int_\Omega \frac{|\nabla u|}{u} dx + \frac{\lambda}{2} \int_\Omega (u - f)^2 dx \qquad (6)$$

In 2010, Lili Huang, Liang Xiao and Zhihui Wei (Xiao *et al.*, 2010) proposed a multiplicative noise removal model based on Weber's total variation regularization:

$$E(u)=\int_{\Omega}\frac{|\nabla u|}{u}dx+\lambda\int_{\Omega}\left(\log u+\frac{f}{u}\right)dx \tag{7}$$

this model combines Weberized total variation restoration with the AA model.

Distributed ADMM

ADMM is an effective method to solve decomposable convex or non-convex optimization problems. First, ADMM can decompose the objective function of the original problem into some resolvable sub-problems. Each subproblem is then solved in parallel. Finally, the solution of the coordinated subproblem yields a global solution to the original problem. This decomposition of dividing the original problem into sub-problems is particularly important and productive for solving large-scale problems (He, 2018). Glowinski and Marrocco proposed ADMM in 1975, firstly. It was revisited and proved to be applicable to large-scale distributed optimization problems by Boyd *et al.* in 2011. This had led to the widespread use of ADMM in engineering applications (Glowinski and Marrocco,1975; Gabay and Mercier, 1976; Boyd *et al.*,2011). In China, Prof. Bingsheng He has made outstanding contributions to the research and application of ADMM (Shi *et al.*, 2014). Currently, ADMM is one of the most frequently used methods for solving total variational regularization models in the image processing field. (Li and Li, 2021; Wei, 2020; Wu *et al.*, 2017; Xiang and Wei, 2020; Chang *et al.*, 2016)

Consider the following optimization problem (He, 2018),

$$\begin{aligned}&\min_{x,z} f(x)+g(z)\\ &s.t.\ Ax+Bx=c\end{aligned}, \tag{8}$$

where $x,z\in X;c\in Y;A,B:X\to Y$, X,Y is Banach space.

The augmented Lagrange function corresponding to the constrained optimization problem equation (8) is

$$L(x,z;w)=f(x)+g(z)+\frac{\mu}{2}\left\|Ax+Bz-c+w\right\|_2^2-\frac{\mu}{2}\left\|w\right\|_2^2 \tag{9}$$

where $w\in Y$ is the Lagrange multiplier and $\mu>0$ is a sufficiently large penalty parameter.

Minimization of the augmented Lagrange function using the alternating iteration method, and then the following ADMM iteration format can be obtained (Nishihara *et al.*, 2015; Wang ei al., 2015; Zhang *et al.*, 2010),

$$\begin{aligned}x^{k+1}&=\arg\min_{x}L(x,z^k;w^k)\\ z^{k+1}&=\arg\min_{z}L(x^{k+1},z;w^k)\\ w^{k+1}&=w^k+(Ax^{k+1}+Bz^{k+1}-c)\end{aligned}. \tag{10}$$

MATERIAL AND METHODS

Proposed model for image restoration

As mentioned in Introduction, the regular term of the existing Weberized total variation restoration model uses Weber's law and the fidelity term is the AA model. The AA model is prone to block effect in the denoising process, and the effect is not good, so we combine the Weberized total variation restoration model with the I-divergence model. The new model is:

$$E(u)=\int_{\Omega}\frac{|\nabla u|}{u}dx+\lambda\int_{\Omega}(u-f\log u)dx \tag{11}$$

Due to the ordinary least squares energy control in $E(u)$, we assume that $f\in L^2(\Omega)$. As a result, $u\in L^2(\Omega)$. Throughout the paper, we also assume that Ω is a Lipschitz open domain with a finite Lebesgue measure $|\Omega|<\infty$. The following natural admissible space for the Weberized TV restoration (Shen, 2003)

$$\Pi(\Omega)=\left\{u>0\middle|u\in L^2(\Omega),\mathrm{TV}(\ln u)<\infty\right\}$$

The model consists of two parts. The first term at the right end of the equation is the regular term, which is composed of the Weberized total variation of the image in the model proposed in this paper, and it can provide a priori information about the original image. The second term is the fidelity term, which is derived from the constraint or derivation principle and mainly guarantees the closeness of the denoised image to the original image. In this paper, we hope that while noise removal is performed on the image, the denoised image and the observed image can be as consistent as possible in terms of information other than noise, and the difference between the two outside of noise signal interference is minimized. We hope that information loss is minimized. Based on this, I-divergence is a good measure that can reflect the information loss between the recovered image and the received image very precisely. Therefore, it is possible to minimize the information loss by minimizing the I- divergence, which is also in line with the idea of model construction of the minimized energy generalization in the variational method.

λ is the regularization parameter. The function of the regularization model is to suppress the amplification of noise and estimate the image stably, while the optimi-

zation model often contains data fidelity terms and a priori regularization terms, and the function of the regularization parameter is to balance the data fidelity terms and a priori regular terms. Mathematically, the solution u of the model is the canonical solution and is related to the parameter λ. As the parameter λ gradually increases, the regularization penalty becomes smaller and the smoothness (regularity) of the obtained solution gradually decreases. However, as the parameters gradually increase, the effect of the data fidelity term increases and the regularized solution will be close to the original observed image.

At the right side of the equation, the first term is the regular term which is non-convex and $\mathrm{TV}(\log u)=\int_\Omega \frac{|\nabla u|}{u}dx$ (Shen, 2003). The second term, which is the fidelity term is convex. Therefore, Eq. (11) is a non-convex model. Convexity not only has good geometric properties, such as support and separation planes, but also has good global analytical properties. In convex optimization problems, the local optimal solution is the global optimal solution, while non-convex optimization problems do not have this property. Therefore, the non-convex model will have the case that there is no global optimal solution when solving for the extremum.

The existence and uniqueness of the solution of Eq. (11) is proved below.

Existence of Weberized total variation restoration

In order to prove the existence of the solution of the optimization Eq. (11), the proof of boundedness of the solution is given first.

Lemma 1(Boundedness) Let $f\in L^\infty(\Omega)$ and $\inf_\Omega f>0$, then the optimization problem Eq.(11) has a solution $\hat{u}$, and $\inf_\Omega f\le\hat{u}\le\sup_\Omega f$.

Proof. Define $\alpha=\inf_\Omega f$, $\beta=\sup_\Omega f$. If $u>f$, $u-f\log u$ strongly increasing. Then

$$\int_\Omega(\inf(u,\beta)-f\log(\inf(u,\beta)))dx\le\int_\Omega(u-f\log u)dx. \quad (12)$$

In addition, according to Lemma 1 in (Kornprobst, 1999), there are

$$\begin{aligned}&\mathrm{TV}(\inf(u,\beta))\le\mathrm{TV}(u)\ \text{and}\\&\mathrm{TV}(\log(\inf(u,\beta)))\le\mathrm{TV}(\log u).\end{aligned}\quad (13)$$

Combine (12) and (13),

$$E(\inf(u,\beta))\le E(u). \quad (14)$$

If and only if, $u\le\beta$ the equation holds.

Since $\hat{u}$ is the minimum solution of problem Eq.(11), the equation holds when $u=\hat{u}$, thus $\hat{u}\le\beta$.

Similarly, it is obtained that

$$E(u)\le E(\sup(u,\alpha)), \quad (15)$$

thus $\hat{u}\ge\alpha$. This completes the proof.

Theorem 1 (Existence) Assume that $f\in L^\infty(\Omega)$ and $\inf_\Omega f>0$, then the optimization problem Eq.(11) has at least one solution in its solution space $\prod(\Omega)$.

Proof. Define $\alpha=\inf_\Omega f$, $\beta=\sup_\Omega f$. Notice that $u\equiv\beta\in\prod(\Omega)$, thus the solution space is non-empty (Shen, 2003). Consider the sequence of minimizations of the optimization problem $\{u_n\}\subset\prod(\Omega)$. According to **Lemma 1** (boundedness), it has $\alpha<u_n<\beta$. Then

$$\int_\Omega\frac{|\nabla u_n|}{u_n}dx\ge\frac{1}{\beta}\int_\Omega|\nabla u_n|dx$$

Because $\{u_n\}$ is a bounded sequence in the Banach space $\mathrm{BV}(\Omega)=\{f:f\in L^1(\Omega)\,and\,\mathrm{TV}(f)<\infty\}$ endowed with the BV norm $\|f\|_{\mathrm{BV}}=\|f\|_{L^1}+\mathrm{TV}(f)$.

And since Ω is bounded, then

$$\|u_n\|_{L^1(\Omega)}<+\infty \quad (16)$$

Moreover, by the definition of minimization sequence $\{u_n\}$, we know that there exists a constant $C>0$, let $E(u_n)\le C$. Since $\mathrm{TV}(\log u_n)<\infty$, $\int_\Omega(u_n-f\log u_n)dx$ reaches a positive minimum $\int_\Omega(f-f\log f)dx$ when $u_n=f$. We have

$$\mathrm{TV}(u_n)\le C. \quad (17)$$

From Eq. (16) and (17), it can be shown that u_n is consistently bounded in $\mathrm{BV}(\Omega)$ space with respect to n. By the tightness of $\mathrm{BV}(\Omega)$ space, there exist subcolumns in $\{u_n\}$ (still denoted by $\{u_n\}$ for convenience) and functions u in $\mathrm{BV}(\Omega)$ space such that $\{u_n\}$ converges strongly to u in $L^1(\Omega)$ space. Further, assume that

$$u_n(x)\to u(x),\quad a.e.\ x\in\Omega.$$

Because of Lebesgue's control convergence theorem, we have

$$\int_\Omega(u-f\log u)dx=\lim_{n\to\infty}\int_\Omega(u_n-f\log u_n)dx. \quad (18)$$

Let $v_n=\log u_n$, $v=\log u$, then

$$v_n(x)\to v(x),\quad a.e.\ x\in\Omega.$$

Assume that $\vec{p}=(p_1,p_2)\in C_0^1(\Omega,R^2)$ is a Vector-valued function and $\|\vec{p}\|=\sqrt{p_1^2+p_2^2}\le 1$, then

$$v_n\nabla\cdot\vec{p}\to v\nabla\cdot\vec{p},\quad |v_n\nabla\cdot\vec{p}|\le|\log\beta\|\nabla\cdot\vec{p}|\ a.e.\ x\in\Omega. \quad (19)$$

Since $\vec{p}$ has a compact support, the left end function of inequality Eq.(19) belongs to the $L^1(\Omega)$ space. Because of Lebesgue's control convergence theorem,

$$\int_\Omega v\nabla\cdot\vec{p}dx=\lim_{n\to\infty}\int_\Omega v_n\nabla\cdot\vec{p}dx. \quad (20)$$

Eq.(20) takes the upper exact bound on $\vec{p}$,

$$\int_\Omega v\nabla\cdot\vec{p}dx\le\liminf_{n\to\infty}\int_\Omega|\nabla v_n|dx. \quad (21)$$

We have (Enrico, 1983)

$$\int_\Omega|\nabla v|dx\le\liminf_{n\to\infty}\int_\Omega|\nabla v_n|dx. \quad (22)$$

Combine Eq.(18) and (22),

$$E(u)\le\liminf_{n\to\infty}E(u_n).$$

It has $u\in\prod(\Omega)$. Since u_n is the minimization sequence, u is the minimum solution of the optimization problem Eq.(11). This completes the proof.

Uniqueness of Weberized total variation restoration

According to optimization theory, when the objective function is strictly convex and forcing, there has a unique minimum solution for the fixed solution problem (Aubert and Kornprobst, 2006). The objective function $E(u)$ in the optimization problem Eq. (11) contains a non-convex negative log likelihood function and a Weberized total variation regular term, so $E(u)$ is non-convex. And thus the uniqueness of the solution cannot be derived directly from the convexity of the objective function. The Euler-Lagrange equation of the optimization problem Eq. (11) is used below to prove the uniqueness of its solution.

Theorem 2 (Uniqueness) Assume that $f\in L^\infty(\Omega)$, $f>0$ and $\lambda>0$. The solution to the optimization problem Eq. (11) is u .Then u is unique.

Proof. Due to the Lemm2 in (Shen, 2003), the solution u of the optimization problem Eq. (11) satisfies the following Euler-Lagrange equation,

$$\begin{aligned}&\lambda\left(1-\frac{f}{u}\right)-\frac{1}{u}\nabla\left[\frac{\nabla u}{|\nabla u|}\right]=0\\&\frac{\partial u}{\partial\vec{n}}=0,x\in\partial\Omega\end{aligned} \quad (23)$$

$\vec{n}=\dfrac{\nabla u}{|\nabla u|}$ is normal vector of $\partial\Omega$. Since $u>0$, Eq.(23) can be equivalently rewritten as

$$\begin{aligned}&-\nabla(\frac{\nabla u}{|\nabla u|})+\lambda(u-f)=0\\&\frac{\partial u}{\partial\vec{n}}=0,x\in\partial\Omega\end{aligned} \quad (24)$$

At each fixed pixel $x\in\Omega$, and define a cubic potential of u by $F'=\lambda(u-f)$

Then

$$F=\lambda(\frac{u^2}{2}-fu)$$
$$F''=\lambda$$

It is easy to deduce that when $\lambda>0$, $F''>0$, and thus F is a strictly convex function.

A new reference objective function can be defined from $E(u)$ as follows

$$E_r(u)=\int_\Omega|\nabla u|+\lambda(\frac{u^2}{2}-fu)dx.$$

It is easy to prove that Eq.(24) is exactly the Euler-Lagrange equation for $E_r(u)$. Since the TV Radon measure is semi-convex, the objective function $E_r(u)$ has global strict convexity, and thus there has a unique minimum solution. The original objective function $E(u)$ also has a unique minimum solution. This completes the proof.

The computational approach

In this section, we will solve the model proposed in this paper with the numerical solution of Euler-Lagrange equations and ADMM.

Numerical solution of Euler-Lagrange equation method

This method is the classical method for image denoising.

We first obtain the Euler-Lagrange equation of the form Eq.(11) as follows:

$$\lambda\left(1-\frac{f}{u}\right)-\frac{1}{u}\nabla\left[\frac{\nabla u}{|\nabla u|}\right]=0 \quad (25)$$

or equivalently:

$$-\nabla\left(\frac{\nabla u}{|\nabla u|}\right)+\lambda(u-f)=0\,. \tag{26}$$

The corresponding diffusion equation is

$$\frac{\partial u}{\partial t}=-\lambda\left(1-\frac{f}{u}\right)+\frac{1}{u}\left[\frac{\nabla u}{|\nabla u|}\right]. \tag{27}$$

Define a new reference energy for $E_r(u)$as follows:

$$E_r(u)=\int_\Omega|\nabla u|+\lambda\left(\frac{u^2}{2}-fu\right)dx \tag{28}$$

It is easy to derive that Eq. (25) is exactly the Euler–Lagrange equation for Eq. (28). Eq. (11) and Eq. (28) have the same Euler-Lagrange equation. The solution of the model Eq. (11) is also the solution of the optimization model Eq. (28). The optimization model Eq. (28) has a unique solution that satisfies Eq. (25). Therefore, the models Eq. (11) and Eq. (28) are equivalent. Eq. (28) can be viewed as an alternative function of the Weberized total variation regularization minimization problem Eq. (11). Eq. (28) is strictly convex and therefore has a unique minimizer. The problem of finding the minimizer of Eq. (11) can be converted to finding the minimizer of Eq. (28).

The Euler-Lagrange equation is solved as follows.

First, the image is discredited by the central difference in half-point format to an eight-neighborhood system. h denotes the step size of the grid, and h is 1 in the iteration. Δt denotes the time step of the iteration. The difference form of the derivative of the pixel point at (i,j) with respect to the auxiliary variable time t can be obtained as shown in Eq. (29).

$$\frac{\partial u}{\partial t}=\frac{u^{n+1}(i,j)-u^n(i,j)}{\Delta t} \tag{29}$$

Combine Eq. (27) and (29), it can be concluded that

$$\frac{u^{n+1}(i,j)-u^n(i,j)}{\Delta t}=\tilde{\lambda}\left(-\lambda\left(1-\frac{f}{u}\right)+\frac{1}{u}\nabla\left(\frac{\nabla u}{|\nabla u|}\right)\right) \tag{30}$$

where $\tilde{\lambda}$ is the proportionality factor.

According to Eq. (30), the iterative form of the Weberized total variation noise removal model when solved using the time stepping method can be obtained as shown

$$u^{n+1}(i,j)=u^n(i,j)+\Delta t\left[-\lambda_1\left(1-\frac{f}{u^n(i,j)}\right)+\lambda_2\frac{1}{u^n(i,j)}\frac{u^n_{\zeta\zeta}(i,j)}{|\nabla u(i,j)|^n+1}\right] \tag{31}$$

where $\lambda_1=\tilde{\lambda}\cdot\lambda$ and $\lambda_2=\tilde{\lambda}$.

There are also subtle differences between the regular terms of the noise removal model in the theoretical expression and in the iterative form. In the theoretical expression, the denominator of the regular term is $|\nabla u|$. While in the iterative form, the denominator of the regular term is $|\nabla u|+1$. This is due to the fact that during the experiment, it was found that the gradient mode of the image would have a small value, when the computer would process it directly to 0 and a pathological situation would occur. Therefore, for the sake of improving the regularity of the model, the denominator is added by one to avoid the above mentioned problems in the iterative process, and this improvement has been found to be effective and reasonable through experiments.

The steps to solve the model equation Eq. (11) with Euler-Lagrange equation are as follows.

Euler-Lagrange equation method for the model
Initialize: $u^0=f$, Parameter: λ_1 , λ_2 For $k=1:n$, n is the number of iterations update u^{k+1} by Eq.(31), if $k=n$, break End

Using the ADMM to the model

This part uses the ADMM for solving the model proposed in Material and Methods.

Based on the format of the optimization problem, Eq. (8) is solved by the ADMM method. The total variation image recovery regularization model Eq. (11) can be equivalently rewritten by introducing the variable d as

$$\min_{u,d}\int_\Omega d\,dx+\frac{\lambda}{2}\|u-f\log u\|_2^2 \quad s.t.\ \frac{|\nabla u|}{u}=d \tag{32}$$

The corresponding augmented Lagrange function is

$$L(x,d;z)=\int_\Omega d\,dx+\frac{\lambda}{2}\|u-f\log u\|_2^2+\frac{\mu}{2}\left\|\frac{|\nabla u|}{u}-d+z\right\|_2^2-\frac{\mu}{2}\|z\|_2^2 \tag{33}$$

where $z=(z_x,z_y)\in X\times X$ is Lagrange multiplier, and $\mu>0$ is a penalty parameter.

From this, ADMM iteration format can be constructed as

$$u-subproblem: u^{k+1}=\arg\min_u \frac{\lambda}{2}\|u-f\log u\|_2^2+\frac{\mu}{2}\left\|\frac{|\nabla u|}{u}-d^k+z^k\right\|_2^2$$
$$d-subproblem: d^{k+1}=\arg\min_d \int_\Omega d\,dx+\frac{\mu}{2}\left\|\frac{|\nabla u^{k+1}|}{u^{k+1}}-d+z^k\right\|_2^2$$
$$z-subproblem: z^{k+1}=z^k+\left(\frac{|\nabla u^{k+1}|}{u^{k+1}}-d^{k+1}\right) \tag{34}$$

For the solution of the u-subproblem, the discrete case is equivalent to solving the following equation (Li and Li, 2021)

$$(\lambda I+\mu\nabla^*\cdot\nabla)u^{k+1}=\lambda f+\mu\nabla_x^*(d_x^k-z_x^k)+\mu\nabla_y^*(d_y^k-z_y^k) \tag{35}$$

Fourier transform on both sides of Eq.(35) simultaneously

$$(\lambda I+\mu(F(\nabla_x)^*\circ F(\nabla_x)+F(\nabla_y)^*\circ F(\nabla_y)))\circ F(u^{k+1})$$
$$=\lambda F(f)+\mu(F(\nabla_x)^*\circ F(d_x^k-z_x^k)+F(\nabla_y)^*\circ F(d_y^k-z_y^k)) \tag{36}$$

where F is discrete Fourier transform, $\circ$ is dot product operations between matrices and I is 1-pattern matrix.

Can be solved using Fourier inversion transform Eq.(36) as

$$u^{k+1}=F^{-1}\left(\frac{\lambda F(f)+\mu(F(\nabla_x)^*\circ F(d_x^k-z_x^k)+F(\nabla_y)^*\circ F(d_y^k-z_y^k))}{\lambda I+\mu(F(\nabla_x)^*\circ F(\nabla_x)+F(\nabla_y)^*\circ F(\nabla_y))}\right) \tag{37}$$

The d-subproblem can be solved analytically using a soft threshold operator,

$$d^{k+1}=S_{\frac{1}{\mu}}\left(\frac{|\nabla u^{k+1}|}{u^{k+1}}+z^k\right) \tag{38}$$

where $S_\tau(x)=\max(|x|-\tau,0)\cdot\frac{x}{|x|}$, when $|x|=0$, let $\frac{x}{|x|}=0$.

Let $\tau=\frac{1}{\mu}$ and $x=\frac{|\nabla u^{k+1}|}{u^{k+1}}+z^k$, Eq.(38) can be written as

$$d^{k+1}=S_{\frac{1}{\mu}}\left(\frac{|\nabla u^{k+1}|}{u^{k+1}}+z^k\right)=\max\left(\left|\frac{|\nabla u^{k+1}|}{u^{k+1}}+z^k\right|-\frac{1}{\mu},0\right)\cdot\frac{\frac{|\nabla u^{k+1}|}{u^{k+1}}+z^k}{\left|\frac{|\nabla u^{k+1}|}{u^{k+1}}+z^k\right|} \tag{39}$$

The steps to solve the model Eq.(11) with ADMM are as follows

ADMM for the model
Initialize: $u^0=f$, $d^0=(d_x^0,d_y^0)=(0,0)$, $z^0=(z_x^0,z_y^0)=(0,0)$
Parameter: $\lambda=1$, μ, tol
For $k=1:MaxIter$
update u^{k+1} by Eq.(37)
d^{k+1} by Eq.(39)
$z^{k+1}=z^k+\left(\frac{\lvert\nabla u^{k+1}\rvert}{u^{k+1}}-d^{k+1}\right)$
if $\frac{\lVert u^{k+1}-u^k\rVert_2}{\lVert u^k\rVert_2}\le tol$
break
End

RESULTS

Evaluation methods and noise image

There are two methods for noise removal. One method is the subjective evaluation method, and the other method is the objective evaluation. As the name implies, the subjective evaluation method refers to the evaluation of noise removal through direct observation by the observer. This type of method is intuitive and convenient, but it is often difficult to identify the advantages and disadvantages of noise removal when higher accuracy is required for analysis and comparison. At this point, the objective evaluation method is necessary. The objective evaluation method refers to the evaluation of image noise removal with quantifiable indexes. There are many such quantitative indicators, and this paper mainly uses the peak signal-to-noise ratio to evaluate the image noise removal situation. Let the size of the image be $M\times N$ size, then the Peak Signal to Noise Ratio (PSNR) is defined as follows:

$$PSNR=10\cdot\log_{10}\frac{255^2}{\frac{1}{M\times N}\sum_{m=1}^{M}\sum_{n=1}^{N}(\hat{u}-u)^2} \tag{40}$$

where $\hat{u}$ refers to the restored image and u refers to the original image. The PSNR can reflect the difference between the two images. The larger the PSNR, the smaller the difference between the two images.

For the problem of similarity between the denoised image and the original image, the aspects of image detail retention and color difference will be evaluated relatively subjectively in combination with PSNR.

In next subsection, experiments are performed with a "Lena" image of size 369×369. Gaussian multiplicative noise is added, as in Fig.1. The probability density function of Gaussian noise is

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

In this paper, the noise parameter is $\mu=1$, $\sigma=0.1$. After adding noise, the gray value of the point (40,40) is 182.7083 and the PSNR of image is 25.6485.

(a) (b)

Fig. 1. *Lena (a)the original (b) add Gaussian noise*

Numerical experiments on the solution of Euler-Lagrange equation method

First, let the number of iterations be 200 and experiment for different parameters, and compare the grayscale value of the point (40, 40) and the PSNR of image after denoised. The results are shown in Table 1 and Fig.2 is more intuitive.

Table 1. *Denoising effect of different parameters when the number of iterations is 200*

λ_1	λ_2	Gray value	PSNR
10	10	179.9137	26.6139
10	100	166.8813	30.8723
10	200	162.4177	32.1859
100	10	182.3758	25.7947
100	100	179.6093	26.8804
100	200	176.9864	27.8194
200	10	182.5414	25.7225
200	100	181.0966	26.3188
200	200	179.6093	26.8807

It can be seen that the best result is obtained with $\lambda_1=10,\lambda_2=200$. Next, fixing the parameters as $\lambda_1=10,\lambda_2=200$, and experiment for different iterations.

As shown in Fig.3, with the increase of the number of iterations, the denoising effect gradually increases and the degree of increase decreases. It can be seen that the PSNR will show a stable state after many iterations during the iterative process. Therefore, it is concluded that the denoising effect will become better and better as the number of iterations increases, while it tends to be stable.

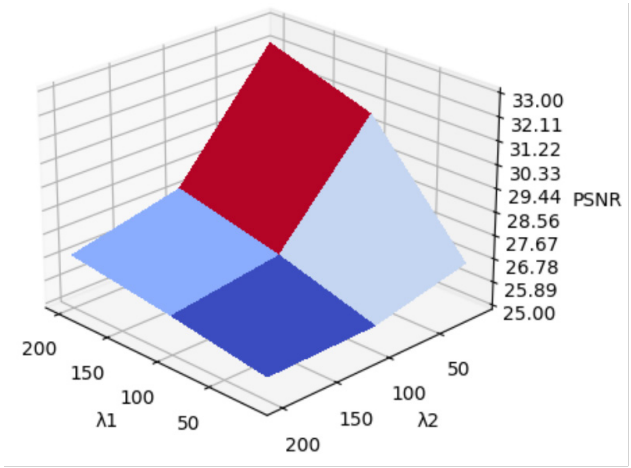


Fig. 2 *Denoising effect of different parameters when the number of iterations is 200*

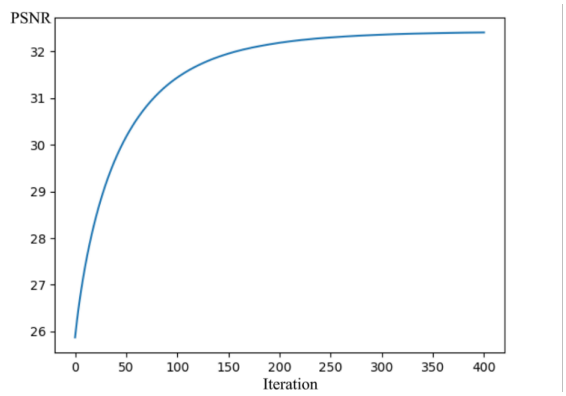


Fig. 3. *Denoising effect with different iterations when* $\lambda_1=10,\lambda_2=200$.

Next, the denoising effect is examined by visual observation. Fig.4 shows the image after denoising. It can be seen that there are better effects, but the image details need to be further improved. It can be intuitively seen that this method has a better denoising effect, and the denoised image is significantly smoother and the edges and details are better preserved compared with the noisy image. After increasing the number of iterations to 400, the continued increase does not significantly improve the denoising effect, but the image details need to be further improved.

(a) (b) (c)

Fig. 4. Denoising image when $\lambda_1 = 10$, $\lambda_2 = 200$ (a) iteration is 100 (b) iteration is 200 (c) iteration is 400.

Numerical Experiments on the solution of ADMM

Continue the experiment with the noise picture of Fig.1. When $tol = 10^{-3}$, different values of the parameter μ are taken for denoising, and compare PSNR. The results are as follows.

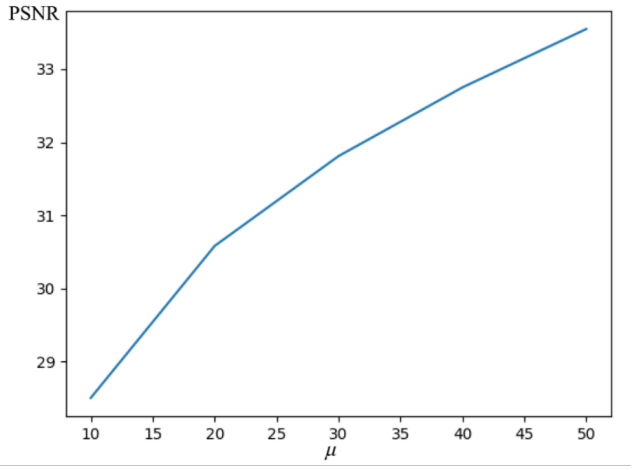


Fig. 5 *Different values of the parameter μ are taken for denoising.*

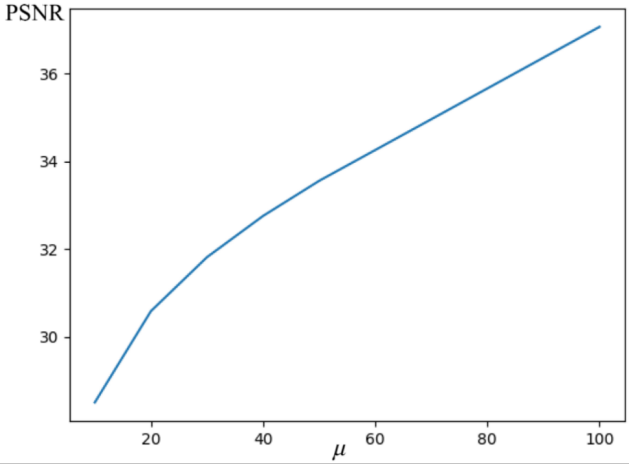


Fig. 6 *Different values of the parameter μ are taken for denoising.*

Table 2. *Runtime and PSNR of different tol*

tol	Time(s)	PSNR
10^{-1}	0.4848	45.95
10^{-2}	0.6708	33.57
10^{-3}	1.1747	33.55
10^{-4}	1.8518	33.58
10^{-5}	2.2178	33.58

According to Fig.5, it can be seen that the the larger the μ, the larger the PSNR. Keep trying $\mu = 100$. As shown in Fig.6, the PSNR increases substantially. It is preliminarily speculated that the larger μ is, the better the denoising effect will be. With fixed μ, experiments were carried out under different tol conditions, as shown in Table 2. According to Table 2, the larger the tol the longer the run time, and PSNR tends to stabilize. When $tol = 10^{-1}$, PSNR reaches its maximum. It is further examined by visual observation whether there exists a better denoising effect when μ is larger or tol is larger.

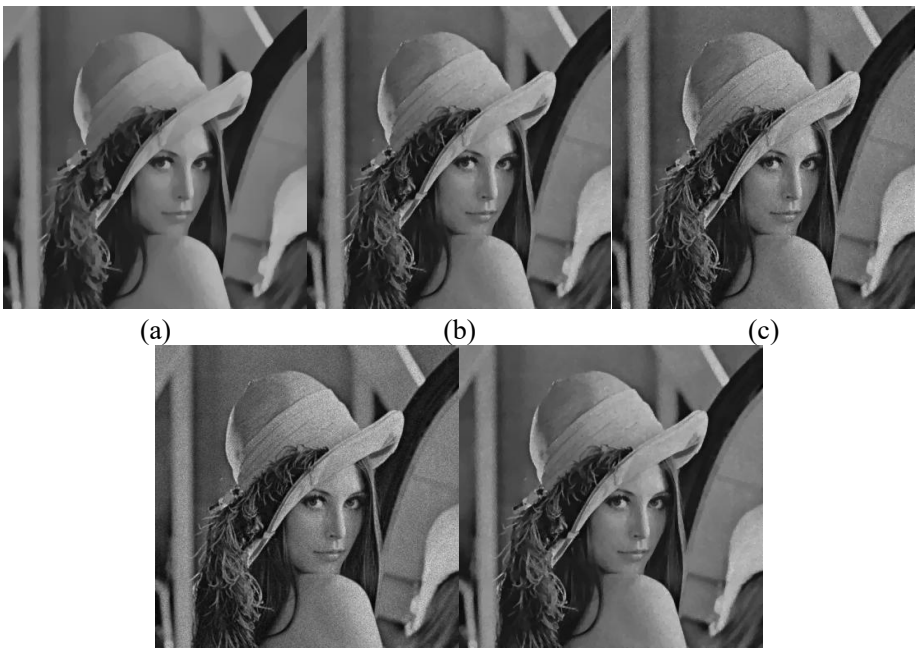
Compare (a), (b), and (c) in Fig.7. When $\mu = 20$, image details are not preserved and many textures disappear. When $\mu = 100$, noise is still present in the image and denoising effect is not good. The best result is obtained at $\mu = 50$. Compare (b) and (d) in Fig.7. It can be seen that when tol is too large, the PSNR is large but the denoising effect is not good. Compare (b) and (e). The difference of denoising effect is not obvious. Considering the denoising effect and the running time, tol is selected as 10^{-3}.

ADMM is better for removing Gaussian multiplicative noise. Noise can be removed, image details are well preserved, and step effects are avoided. The image after ADMM denoising is intuitively more detailed.

Comparison of the two algorithms

In this subsection, we compare the Euler-Lagrange equation method with the ADMM.

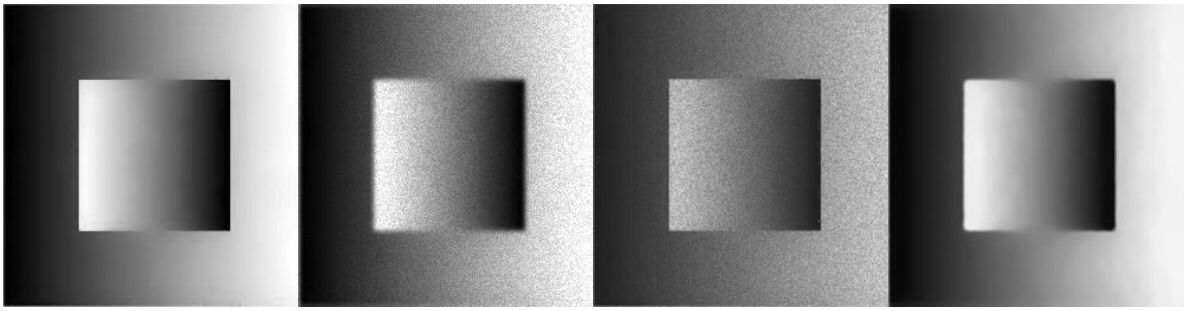
For a more direct and objective comparison, this part of the experiment was compared on synthetic images. Fig. 8 shows the noise removal for the synthetic image. Fig.8 (a) shows the original figure. (c) is the



(a) (b) (c)

(d) (e)

Fig. 7. *(a)* $\mu = 20, tol = 10^{-3}$, *(b)* $\mu = 50, tol = 10^{-3}$, *(c)* $\mu = 100, tol = 10^{-3}$, *(d)* $\mu = 50, tol = 10^{-1}$, *(e)* $\mu = 50, tol = 10^{-1}$.



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

Fig. 8. *(a) the original, (b) add Gaussian noise, (c) Euler-Lagrange equation, (d) ADMM.*

denoising by the partial differential equation method with the parameters $\lambda_1 = 1$, $\lambda_2 = 10$, and the iteration is 10, at which time the best results are obtained under the partial differential equation method. (d) is the denoising by the ADMM method with parameter $\mu = 50, tol = 10^{-3}$. ADMM has better denoising effect than the Euler-Lagrange equation method, while the brightness is closer to the original image.

The composite image is further studied. To show the restored abilities, we choose the 80th line of the original image of composite image and the restored image as shown in Fig.9. In order to observe the denoising effect in more details, we take a part of the original image and make 3D images for noise image and restored images. It is shown in Fig.10.

It can be seen that the denoising effect of ADMM is better. The denoised image is smoother and closer to the original image.

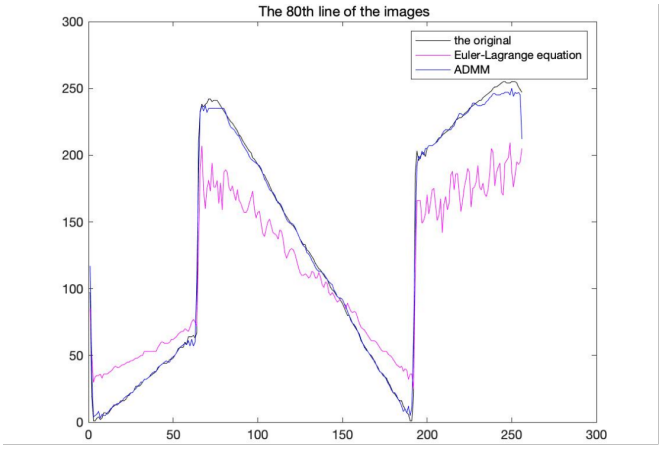


Fig. 9. *The 80th line of the restored of composite image*

Next, change the noise type. In Fig.11, add Rayleigh noise, which is common in medical noise images, to the image and observe the denoising effect. Its probability density is given by the following equation,

$$p(x)=\begin{cases}\dfrac{2}{b}(x-a)\exp\left[-\dfrac{(x-a)^2}{b}\right], x\geq a\\ 0, x<a\end{cases}$$

The noise parameters are $a=0$, $b=0.03$. In Fig.11, (c) with parameter $\lambda_1=10$, $\lambda_2=200$ and 200 iterations. (d) is the ADMM method for denoising with parameter $\mu=50, tol=10^{-3}$.

It can be seen that ADMM is effective for Rayleigh noise, but the picture brightness is not recovered. The Euler-Lagrange equation method works better than ADMM. Therefore, our model used in the field of medical imaging needs to be further investigated.

Discussion

In this paper, we study a new non-convex variational model in the Weberized total variation regularization framework for the multiplicative noise removal problem. The new nonconvex model is obtained by combining the Weberized total variation regularization term and the I- divergence fidelity term. The existence of the model solution is proved and there is a unique solution. Euler-Lagrange equation method and ADMM are used to denoise the images respectively. ADMM outperforms.

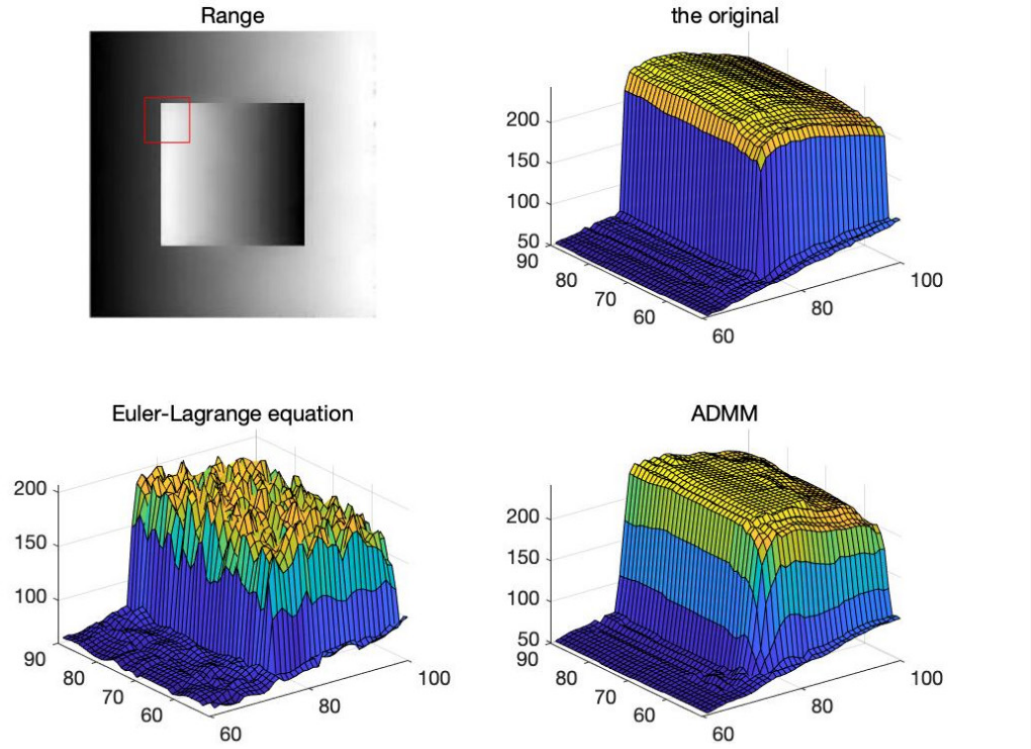
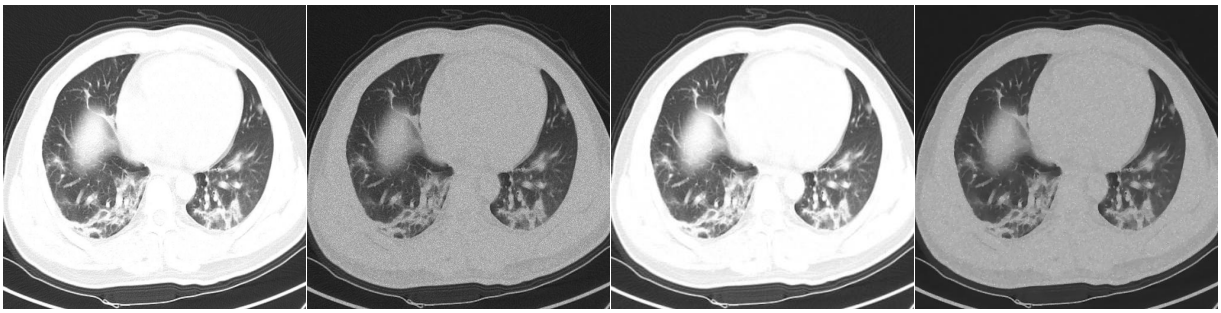


Fig. 10. *3D images for noise image and restored images.*



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

Fig. 11. *(a) the original, (b) add Rayleigh noise, (c) Euler-Lagrange equation, (d) ADMM.*

REFERENCES

Aubert G, Aujol JF (2008). A variational approach to removing multiplicative noise. SIAM J Appl Math 68: 925–46.

Aubert G, Kornprobst P (2006). Mathematical problems in image processing: Partial differential equations and the calculus of variations (applied mathematical sciences). Appl Intell 40:291-304.

Bingsheng H (2018). My 20 years research on alternating directions method of multipliers. Operations Research Transactions 22:1-31.

Boyd S, Parikh N, Chu E, *et al* (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. Found Trends Mach Learn 3:1-122.

Chang TH, Hong M, Liao WC, *et al.* (2016). Asynchronous Distributed ADMM for Large-Scale Optimization—Part I: Algorithm and Convergence Analysis. IEEE T Signal Proces 64: 3118-30.

Enrico G (1983). Minimal surfaces and functions of bounded variation. Monographs in Mathematics 45:7-26.

Gabay D, Mercier B(1976). A dual algorithm for the solution of nonlinear variational problems via finite element approximaton . Comput Math Appl 2:17-40.

Glowinski R, Marrocco A(1975). Sur 1' approximation, pare´le´ments find d' order 1, et la re´solution, par pe´nelisation-dualite´, d' une classe de problèmes de dirichlet non linéaires. J Equine Vet Sci 2:41-76.

Kornprobst P, Deriche R, Aubert G(1999). Image sequence analysis via partial differential equations [J]. J Math Imaging Vis 11:5-26.

Li M, Li B (2021). A novel weighted total variation model for image denoising. IET Image Process 15:2749-60.

Nishihara R, Lessard L, Recht B, *et al.* (2015). A general analysis of the convergence of ADMM. ICML JMLR org 1502:02009.

Rudin LI, Lions PL, Osher S(2003). Multiplicative denoising and deblurring: Theory and algorithms. Springer New York, 103–20.

Shen J (2003). On the foundations of vision modeling I. Weber's law and Weberized TV restoration. Physica D 175.

Shi J, Osher S (2008). A non-linear inverse scale space method for a convex multiplicative noise model. SIAMJ Imaging Sci 01: 294–321.

Shi W, Ling Q, Yuan K, *et al*. (2014). On the linear convergence of the admm in decentralized consensus optimization. IEEE T Signal Proces 62:1750-61.

Steidl G, Teuber T (2010). Removing Multiplicative Noise by Douglas-Rachford Splitting Methods. J Math Imaging Vis 36: 168-84.

Wang Y, Yin W, Zeng J (2015). Global Convergence of ADMM in Nonconvex Nonsmooth Optimization. CoRR,2015,abs/1511.06324.

Weber EH (1834). De pulsu, resorptione, audita et tactu, Annotationes anatomicae et physiologicae. Koehler, Leipzig.

Wei J (2020). Research on Image Denoising Algorithm Based on ADMM. Harbin Engineering University.

Wu B, Wu Y, Zhang H (2008). Image recovery technique based on variational partial differential equations. Beijing University Press.

Wu Z, Li M, Wang DZW, *et al.* (2017). A Symmetric Alternating Direction Method of Multipliers for Separable Nonconvex Minimization Problems. Asia Pac J Oper Res 34:1750030.

Xiang J, Wei J (2020). An improved image denoising method based on ADMM. Applied Science and Technology 47: 14-19.

Xiao L, Huang L, Wei Z (2010). A Weberized Total Variation Regularization-Based Image Multiplicative Noise Removal Algorithm. Eurasip J Adv Sig Pr 2010, 15p.

Zhang X, Burger M, Bresson X, *et al*. (2010). Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. SIAM J Imaging ScI 3:253-76.