

## IMAGE DENOISING BASED ON A NEW ANISOTROPIC MEAN CURVATURE MODEL

YUAN WANG<sup>✉1</sup> AND ZHI-FENG PANG<sup>✉\*,2</sup>

<sup>1</sup>School of Mathematical Sciences, Zhejiang University, Hangzhou, China

<sup>2</sup>School of Mathematics and Statistics, Henan University, Kaifeng, China

(Communicated by Weihong Guo)

**ABSTRACT.** A number of variational models for image denoising have been proposed in the last few years in order to advance the denoising performance. To improve the denoising quality, it is very significant to describe the local structure of image in the proposed models. To this end, this paper proposes a novel denoising model which combines the gradient operator  $\nabla$  with the adaptive weighted matrix  $W$  in the mean curvature regularized term such that the proposed model can describe the local features in image efficiently. Since the proposed model is a high-order nonlinear and nonconvex optimization problem, we need to use the operator splitting method to transform it into a multi-variable optimization problem and then the alternating direction method of multipliers (ADMM) can be applied to solve it. Numerical experiments demonstrate that the proposed model yields good performance compared with other well-known gradient-based models.

**1. Introduction.** Image denoising problem plays an important role in the field of image processing. It aims to remove inconsequential signals like noise while keeping meaningful ones like edges and textures. However, this problem is a classic ill-posed problem due to the lacking of prior information. Therefore, we need to employ some mathematical schemes to overcome it. During the past few decades, numerous methods have been proposed to deal with this problem, such as linear and non-linear filtering [33], partial differential equations [20], statistical methods [9], gradient-based methods [5, 17], sparse models [21], CNN-based models [30, 25, 29, 1, 12]. Among these, the gradient-based methods have become increasingly popular methods. In this paper, we focus on a class of gradient-based methods. Specifically, this class of models can be written as the following optimization problem

$$\min_u \left\{ \frac{\lambda}{2} \int_{\Omega} (f - u)^2 dx dy + \mathcal{R}(\nabla u) \right\}, \quad (1)$$

where the clean image  $u : \Omega \rightarrow \mathbb{R}$  is estimated from the noisy observation  $f : \Omega \rightarrow \mathbb{R}$ . Here the image domain  $\Omega$  denotes a connected bounded open subset of  $\mathbb{R}^2$  with the Lipschitz boundary. The fidelity term  $\int_{\Omega} (f - u)^2 dx dy$  penalizes that the restored image  $u$  is very far away from the noisy observed  $f$ . The regularization term  $\mathcal{R}(\nabla u)$

2020 *Mathematics Subject Classification.* Primary: 58F15, 58F17; Secondary: 53C35.

*Key words and phrases.* Image denoising, anisotropic mean curvature, augmented Lagrangian method, weighted matrix, alternating direction method of multipliers.

\*Corresponding author: Zhi-Feng Pang.

mainly describes some image prior information such as continuity, smoothness or bounded variation. The regularization parameter  $\lambda > 0$  is used to balance the effect of the data fitting term and the regularization term.

In the model (1), choosing an effective regularization scheme is very important for image denoising problem. A classical scheme was developed by Rudin, Osher, and Fatemi [26]. This method mainly wants to find a clean image from a noisy observation by considering the following function:

$$\min_u \left\{ \frac{\lambda}{2} \int_{\Omega} (f - u)^2 dx dy + \int_{\Omega} |\nabla u| dx dy \right\}, \quad (2)$$

where the total variation (TV) term  $\int_{\Omega} |\nabla u| dx dy = \int_{\Omega} \sqrt{u_x^2 + u_y^2} dx dy$ . It has been proved to be very efficient for noise removal and has been widely used in other image processing tasks such as deblurring, inpainting and reconstruction. However, the TV model (2) also suffers some unfavorable properties such as the staircase effect, smears object corners and then leads to the loss of image contrast. To overcome these problems, a popular approach is to add high order derivative into the regularization term [10, 35, 18, 8]. Besides, in [6, 31, 7], authors combined the TV with the Laplacian regularization for improving the quality of the restored image. In [19], authors directly incorporated second order derivative information into the image denoising model. In [2], L. Ambrosio and S. Masnou used the Euler's elastic of the level curves of a smooth function as the regularization. To solve Euler's elastic efficiently, authors proposed a linearized augmented Lagrangian method in [36]. Although the high-order TV-based models can avoid the staircase effect, these models often blur image edges. So other improved models have been proposed such as the total generalized variation (TGV) model[4], hybrid TV-based model [24, 16], nolocal-based model [11], Schatten-norm model [14] and fractional order derivative-based model [3, 37], etc. Since the above TV-based models are isotropic models, they can not efficiently couple with local features in image. Therefore we wish that the proposed model diffuses along the tangent direction of local features. To this end, in [23], authors proposed an anisotropic TV-based (ATV) model to deal with the problem. Inspired by ATV model, Yang et al. proposed an adaptive Euler's elastic-based model to handle the isotropic problem in [34].

Different to the above methods regarding a given image as a function defined on a domain  $\Omega \subset \mathbb{R}^2$ , some methods in [13, 22] considered the image as an embedded surface in  $\mathbb{R}^3$  and used some Riemannian manifold measures as the regularization term. In [39], authors employed the mean curvature (MC) of an image surface of a given image as the regularizer. This model can sweep noise while preserving image contrasts, object corners and edges, and it also ameliorates the staircase effect. However, the associated functions are not easy to be minimized due to high nonlinearity and nonconvexity. In order to efficiently solve the model, Zhu et al. employed the idea of augmented Lagrangian methods to minimize the associated function in [40]. Furthermore, they [38] extended the mean curvature model to be the  $L^p$ -norm of the MC with  $p > 1$  as a new class of regularization term to impose stronger regularity.

Motivated by the above discussions, this paper wants to propose a new image denoising model based on the mean curvature regularization. Actually, since the MC model is an isotropic model, it is difficult to keep the diffusion of the corresponding Euler-Lagrangian equation along with the tangential direction of the edge. In order

to overcome this drawback, inspired by the ATV model [23], we add an adaptive weighted matrix into each component of the gradient operator and then propose the anisotropic mean curvature (AMC) model to increase the ability of the anisotropic diffusion. The contributions can be summarized as follows.

- We propose an anisotropic mean curvature (AMC) model for the image denoising problem by coupling a weighted matrix with the gradient operator.
- Since the proposed model is high nonlinear and non-convex, we apply the classic augmented Lagrangian method to decompose the original problem into several easily solvable subproblems.
- Extensive experiments on classic testing images show the robustness of the proposed method compared with several related denoising models.

The rest of the paper is organized as follows. Section 2 mainly recalls some denoising models related to our proposed model. In section 3, we present the proposed model and the corresponding numerical algorithm. Section 4 gives some numerical comparisons to show the effectiveness of the proposed model. Concluding remarks are given in the last section.

**2. Related models.** In this section, some denoising models related to our proposed model are reviewed, such as the TV model (TV model)[26], the High-order TV-based model (HOTV model)[8], the total generalized variation model (TGV model)[4], the anisotropic TV-based model (ATV model)[23] and the mean curvature model (MC model)[39].

**2.1. TV model.** The TV model proposed by Rudin, Osher and Fatemi [26] has been proven to be an efficient model for image denoising. And it has also been used in other image processing such as image restoration, image deblurring, image inpainting, etc. This model is defined as

$$\min_u \left\{ \frac{\lambda}{2} \int_{\Omega} (f - u)^2 dx dy + \int_{\Omega} |\nabla u| dx dy \right\}. \quad (3)$$

Even though the TV model is very powerful in removing noise while preserving edges and contours of objectives, it also has some unfavorable properties. For instance, it can occur staircase effect and lose image contrasts.

**2.2. High-order TV-based model (HOTV).** In order to solve the staircase effect, quite a few high order variational models have been proposed, such as the Laplacian penalty [27] and the Hessian Schatten-norm regularization [15]. One example of the HOTV models was proposed in [8] as

$$\min_u \left\{ \frac{\lambda}{2} \int_{\Omega} (f - u)^2 dx dy + \int_{\Omega} |\nabla^2 u| dx dy \right\}. \quad (4)$$

This model can remedy staircase effect, but it may lead to blur image edges.

**2.3. Total generalized variation (TGV)-based model.** The TGV model [4] was first defined in the space of symmetric tensors of order  $k$ . However, there are some numerical difficulties due to the tensors. In order to solve the TGV-based model efficiently, one work in [32] employed the Legendre-Fenchel transformation

to consider the dual form and proposed a new TGV model as

$$\min_u \left\{ \frac{\lambda}{2} \int_{\Omega} (f - u)^2 dx dy + \alpha_0 \int_{\Omega} |\nabla u - \mathbf{w}| dx dy + \alpha_1 \int_{\Omega} |\varrho(\mathbf{w})| dx dy \right\}, \quad (5)$$

where  $\alpha_0$  and  $\alpha_1$  are the positive parameters, and  $\varrho(\mathbf{w})$  is defined as

$$\varrho(\mathbf{w}) = \frac{1}{2} \begin{bmatrix} 2\nabla_x w_1 & \nabla_y w_1 + \nabla_x w_2 \\ \nabla_y w_1 + \nabla_x w_2 & 2\nabla_y w_2 \end{bmatrix}$$

for  $\mathbf{w} = (w_1, w_2)^T$ . TGV model is ability to reconstruct piecewise polynomial functions.

**2.4. Anisotropic TV-based (ATV) model.** Above TV-based models are isotropic models because of using the same weight in the gradient operator  $\nabla = (\nabla x, \nabla y)^T$  such that  $\nabla x$  and  $\nabla y$  give the same penalties along the  $x$ -axis and the  $y$ -axis directions. It leads to not coupling with local features in image efficiently. Furthermore, Zhi-Feng Pang et al. proposed the anisotropic TV-based (ATV) model to overcome it and the model is followed as

$$\min_u \left\{ \frac{\lambda}{2} \int_{\Omega} (f - u)^2 dx dy + \int_{\Omega} |\mathbf{T} \nabla u| dx dy \right\}, \quad (6)$$

where  $\mathbf{T} = \text{diag}(t_1, t_2)$ . This model can couple with local features in image when choosing suitable  $t_1$  and  $t_2$  comparing with the classic TV-based models.

**2.5. Mean curvature (MC)-based model.** In [39], authors proposed a variational denoising model that used the mean curvature of the image surface. The minimization function can be written as:

$$\min_u \left\{ \lambda \int_{\Omega} |\kappa_u| dx dy + \frac{1}{2} \int_{\Omega} (f - u)^2 dx dy \right\}, \quad (7)$$

where  $\alpha$  is a positive parameter and  $\kappa_u = \nabla \cdot \left( \frac{\nabla u}{\sqrt{1+|\nabla u|^2}} \right)$  is the mean curvature of the image surface  $\psi(x, y, z) = u(x, y) - z = 0$ . It can remove the noise from the image and preserve image contrasts. However, it is a challenging issue to develop numerical algorithms to minimize its function as it involves high order derivatives and the non-differentiable regularization term. In order to overcome the MC model solving difficulties, the authors in [40] employed the idea of augmented Lagrangian methods to minimize the function (7). Even though this model solved effectively the minimization of the MC model and the advantages of MC model were retained, it was an isotropic model because of using the same weight for gradient operator.

**3. The proposed anisotropic MC (AMC) model.** In this section, we first introduce our proposed anisotropic mean curvature (AMC) model for the denoising problem and then give an efficient numerical method to solve it.

**3.1. The proposed model.** In order to increase the ability of the anisotropy diffusion of the MC model, we introduce different weight to each component of the gradient operator and then propose the following anisotropic mean curvature (AMC) model as

$$\min_u \left\{ \frac{1}{2} \int_{\Omega} (f - u)^2 dx dy + \lambda \int_{\Omega} \left| \nabla \left( \frac{W \nabla u}{\sqrt{1+|W \nabla u|^2}} \right) \right| dx dy \right\}, \quad (8)$$

where the weighted matrix  $W$  is defined by

$$W = \begin{bmatrix} W_x & 0 \\ 0 & W_y \end{bmatrix} = \begin{bmatrix} \frac{1}{1+\beta|G_\sigma * \nabla_x f|} & 0 \\ 0 & \frac{1}{1+\beta|G_\sigma * \nabla_y f|} \end{bmatrix}.$$

Here  $\beta > 0$  and  $\sigma$  is the variance of the Gaussian convolution function  $G_\sigma$ .

**Remark 3.1.** Here we elaborate how to choose the weighted matrix  $W$ . First, we test a group of affine images with different angels such as  $0, \pi/8, \pi/4, 3\pi/8, \pi/2$  and add the white Gaussian noise with variance as  $\sigma = 0.05, 0.1, 0.2$ . Then, we analyse the relationship between  $W_x$  and  $W_y$  for the denoising images, as shown in Figure 1. It is obvious that the ratio of  $\log(W_x/W_y)$  depends on the change of angels. Meanwhile, the ratio is symmetrical. Based on this, we choose  $W := \text{diag}(W_x, W_y)$  as the weighted matrix. In addition, we employ the gradient information in the weighted matrix  $W$ . As is well known that gradient information can relate the structure of objects in an image and identify features of interest for recognition and classification directly. So it is reasonable for weighted matrix  $W$  with gradient information in order to increase the ability of the anisotropy diffusion when proposing the denoising model as done in (8).

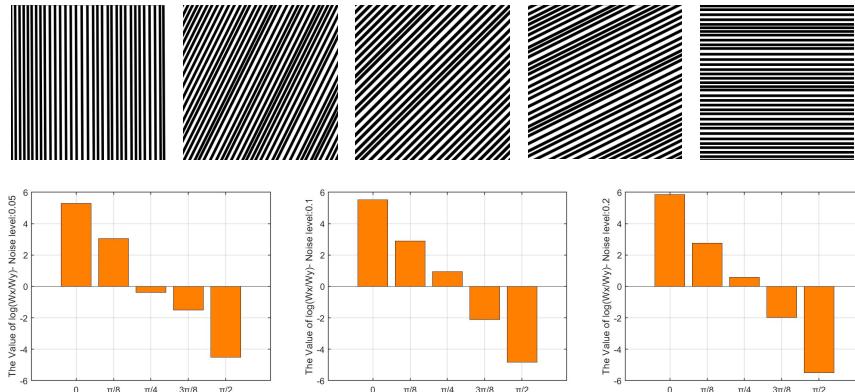


FIGURE 1. 1st: affine images with different angels  $0, \pi/8, \pi/4, 3\pi/8, \pi/2$ . 2nd: the ratio of  $\log(W_x/W_y)$ .

**3.2. The augmented Lagrangian method for the proposed denoising model.** Because the model (8) involves high order derivatives, the related Euler-Lagrange equations are fourth-order equations, which raises a challenge of developing an efficient numerical algorithm to solve them. In addition, the mean curvature  $\nabla \cdot \left( \frac{\nabla u}{\sqrt{1+|\nabla u|^2}} \right)$  is not homogeneous in the variable  $u$ . Therefore, in reference to [40], we introduce new variables and then use augmented Lagrangian method to deal with this problem.

As it makes no difference for the curvature term in the model if we set  $\mathbf{v} = (\nabla u, 1)$ , then the proposed AMC model (8) is transferred to the following constrained optimization problem

$$\begin{cases} \min_{u,q,n,p,v} \frac{1}{2} \int_{\Omega} (f - u)^2 dx dy + \lambda \int_{\Omega} |q| dx dy, \\ \text{s.t. } q = \nabla n, n = \frac{\mathbf{p}}{|\mathbf{p}|}, \mathbf{p} = \bar{\mathbf{W}} \mathbf{v}, \mathbf{v} = (\nabla u, 1), \end{cases} \quad (9)$$

where  $\mathbf{n} := (n_1, n_2, n_3)$ ,  $\mathbf{p} := (p_1, p_2, p_3)$ ,  $\bar{\mathbf{W}} := (W, 1)$  and  $\mathbf{v} := (v_1, v_2, v_3)$ . Then, based on the augmented Lagrange method, we rewrite the problem (9) as

$$\begin{aligned} & \mathcal{L}(u, q, \mathbf{n}, \mathbf{p}, \mathbf{v}, \mathbf{m}; \lambda_1, \boldsymbol{\lambda}_2, \lambda_3, \boldsymbol{\lambda}_4, \boldsymbol{\lambda}_5) \\ &= \frac{1}{2} \int_{\Omega} (f - u)^2 dx dy + \lambda \int_{\Omega} |q| dx dy + r_1 \int_{\Omega} (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}) dx dy \\ &+ \int_{\Omega} \lambda_1 (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}) dx dy + \frac{r_2}{2} \int_{\Omega} (\mathbf{p} - \bar{\mathbf{W}} \mathbf{v})^2 dx dy \\ &+ \int_{\Omega} \boldsymbol{\lambda}_2 (\mathbf{p} - \bar{\mathbf{W}} \mathbf{v}) dx dy + \frac{r_3}{2} \int_{\Omega} (q - \partial_x n_1 - \partial_y n_2)^2 dx dy \\ &+ \int_{\Omega} \lambda_3 (q - \partial_x n_1 - \partial_y n_2) dx dy + \frac{r_4}{2} \int_{\Omega} |\mathbf{v} - (\nabla u, 1)|^2 dx dy \\ &+ \int_{\Omega} \boldsymbol{\lambda}_4 (\mathbf{v} - (\nabla u, 1)) dx dy \\ &+ \frac{r_5}{2} \int_{\Omega} |\mathbf{n} - \mathbf{m}|^2 dx dy + \int_{\Omega} \boldsymbol{\lambda}_5 (\mathbf{n} - \mathbf{m}) dx dy + \delta_{\mathcal{C}}(\mathbf{m}), \end{aligned} \quad (10)$$

where  $r_i, i = 1, \dots, 5$  are the penalization parameters,  $\lambda_1, \lambda_3 \in \mathbb{R}$ ,  $\boldsymbol{\lambda}_2, \boldsymbol{\lambda}_4, \boldsymbol{\lambda}_5 \in \mathbb{R}^3$  are Lagrange multipliers and  $\mathbf{p}, \mathbf{n}, \mathbf{m}, \mathbf{v} \in \mathbb{R}^3$ .  $\mathcal{C} = \{\mathbf{m} \in L^2(\Omega) : |\mathbf{m}| \leq 1\}$ , and  $\delta_{\mathcal{C}}(\cdot)$  is the characteristic function, defined as:

$$\delta_{\mathcal{C}}(\mathbf{m}) = \begin{cases} 0, & \mathbf{m} \in \mathcal{C}, \\ +\infty, & \text{otherwise.} \end{cases}$$

The introduction of new variables adopted here is similar to reference [28]. To get the most out of this paper, we make a few remarks in the following.

**Remark 3.2.** The use of the variable  $\mathbf{m}$  with  $|\mathbf{m}| \leq 1$  can be regarded as a relaxation. Moreover, the constraint  $|\mathbf{m}| \leq 1$  is very important to prevent the unboundedness of  $\mathbf{m}$  when  $\mathbf{p} = 0$ .

It is well-known that the solution of the constrained optimization problem (9) corresponds to one of the saddle points for the augmented Lagrangian function (10). We need to find the saddle point of (10) by using an iterative algorithm. However, the problem (10) is a multi-variable optimization problem, where the variables are coupled together, so it is difficult to solve them directly. In order to decouple with them, one efficient method is ADMM to transform it into several easily solvable subproblems. More specifically, we need to solve one of variables and simultaneously fix others. To this end, we consider the following subproblems and obtain solvers for each of them.

$$\begin{aligned} \xi_1(u) = \min_u \frac{1}{2} \int_{\Omega} (f - u)^2 dx dy + \frac{r_4}{2} \int_{\Omega} |\mathbf{v}^{k-1} - (\nabla u, 1)|^2 dx dy + \\ \int_{\Omega} \boldsymbol{\lambda}_4^{k-1} (\mathbf{v}^{k-1} - (\nabla u, 1)) dx dy, \end{aligned} \quad (11)$$

$$\begin{aligned}\xi_2(q) = \min_q \lambda \int_{\Omega} |q| dx dy + \frac{r_3}{2} \int_{\Omega} (q - \partial_x n_1^{k-1} - \partial_y n_2^{k-1})^2 dx dy + \\ \int_{\Omega} \lambda_3^{k-1} (q - \partial_x n_1^{k-1} - \partial_y n_2^{k-1}) dx dy,\end{aligned}\quad (12)$$

$$\begin{aligned}\xi_3(\mathbf{n}) = \min_{\mathbf{n}} \frac{r_3}{2} \int_{\Omega} (q^k - \partial_x n_1 - \partial_y n_2)^2 dx dy + \int_{\Omega} \lambda_3^{k-1} (q^k - \partial_x n_1 - \partial_y n_2) dx dy + \\ \frac{r_5}{2} \int_{\Omega} |\mathbf{n} - \mathbf{m}^{k-1}|^2 dx dy + \int_{\Omega} \lambda_5^{k-1} (\mathbf{n} - \mathbf{m}^{k-1}) dx dy,\end{aligned}\quad (13)$$

$$\begin{aligned}\xi_4(\mathbf{p}) = \min_{\mathbf{p}} r_1 \int_{\Omega} (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}^{k-1}) dx dy + \int_{\Omega} \lambda_1^{k-1} (|\mathbf{p}| - \mathbf{p} \cdot \mathbf{m}^{k-1}) dx dy + \\ \frac{r_2}{2} \int_{\Omega} (\mathbf{p} - \bar{\mathbf{W}} \mathbf{v}^{k-1})^2 dx dy + \int_{\Omega} \lambda_2^{k-1} (\mathbf{p} - \bar{\mathbf{W}} \mathbf{v}^{k-1}) dx dy,\end{aligned}\quad (14)$$

$$\begin{aligned}\xi_5(\mathbf{v}) = \min_{\mathbf{v}} \frac{r_2}{2} \int_{\Omega} (\mathbf{p}^k - \bar{\mathbf{W}} \mathbf{v})^2 dx dy + \int_{\Omega} \lambda_2^{k-1} (\mathbf{p}^k - \bar{\mathbf{W}} \mathbf{v}) dx dy + \\ \frac{r_4}{2} \int_{\Omega} |\mathbf{v} - (\nabla u^k, 1)|^2 dx dy + \int_{\Omega} \lambda_4^{k-1} (\mathbf{v} - (\nabla u^k, 1)) dx dy,\end{aligned}\quad (15)$$

$$\begin{aligned}\xi_6(\mathbf{m}) = \min_{\mathbf{m}} r_1 \int_{\Omega} (|\mathbf{p}^k| - \mathbf{p}^k \cdot \mathbf{m}) dx dy + \int_{\Omega} \lambda_1^{k-1} (|\mathbf{p}^k| - \mathbf{p}^k \cdot \mathbf{m}) dx dy + \\ \frac{r_5}{2} \int_{\Omega} |\mathbf{n}^k - \mathbf{m}|^2 dx dy + \int_{\Omega} \lambda_5^{k-1} (\mathbf{n}^k - \mathbf{m}) dx dy + \delta_{\mathcal{R}}(\mathbf{m}).\end{aligned}\quad (16)$$

- For the subproblem  $\xi_1(u)$ , it is a least square optimal problem and the corresponding Euler-Lagrange equation is

$$-r_4 \Delta u + u = f - (r_4 v_1^{k-1} + \lambda_{41}^{k-1})_x - (r_4 v_2^{k-1} + \lambda_{42}^{k-1})_y. \quad (17)$$

Choosing different boundary conditions, it is corresponded to different numerical methods. Here we assume that the boundary condition is periodic, then the fast Fourier transform (FFT) can be used to solve it. First, we discretized the problem(17) as

$$-r_4 \operatorname{div}^{-} \nabla^{+} u + u = f - \partial_x^{-} (r_4 v_1^{k-1} + \lambda_{41}^{k-1}) - \partial_y^{-} (r_4 v_2^{k-1} + \lambda_{42}^{k-1}).$$

We then apply the discrete Fourier transform to both sides as

$$\begin{aligned}\mathcal{F}(\partial_x^{\pm} u) &= \pm(e^{\pm\sqrt{-1}z_i} - 1)\mathcal{F}(u), \\ \mathcal{F}(\partial_y^{\pm} u) &= \pm(e^{\pm\sqrt{-1}z_j} - 1)\mathcal{F}(u),\end{aligned}$$

where  $z_i^1 = 2\pi(i-1)/M, i = 1, \dots, M, z_j^2 = 2\pi(j-1)/N, j = 1, \dots, N$ , and we have

$$(-2r_4(\cos z_i^1 + \cos z_j^2 - 2) + 1)\mathcal{F}(u) = \mathcal{F}(g),$$

where  $g = f - \partial_x^{-} (r_4 v_1^{k-1} + \lambda_{41}^{k-1}) - \partial_y^{-} (r_4 v_2^{k-1} + \lambda_{42}^{k-1})$ . So we can calculate the inverse Fourier transform to obtain  $u$ , that is,

$$u = \mathcal{F}^{-1} \left( \frac{\mathcal{F}(g)}{-2r_4(\cos z_i^1 + \cos z_j^2 - 2) + 1} \right). \quad (18)$$

- For the subproblem  $\xi_2(q)$ , we can combine the last two terms in (12) into one term. Then we have

$$\min_q \lambda \int_{\Omega} |q| dx dy + \frac{r_3}{2} \int_{\Omega} \left( q - \left( \partial_x n_1^{k-1} + \partial_y n_2^{k-1} - \frac{\lambda_3^{k-1}}{r_3} \right) \right)^2 dx dy.$$

Based on soft threshold operation, the minimizer of the problem (12) is given as:

$$q = \max \left\{ 0, 1 - \frac{\lambda}{r_3 |\tilde{q}|} \right\} \tilde{q}, \quad (19)$$

$$\text{where } \tilde{q} = \partial_x^- n_1^{k-1} + \partial_y^- n_2^{k-1} - \frac{\lambda_3^{k-1}}{r_3}.$$

- For the subproblem  $\xi_3(n)$ , the solution is similar to that of the problem (11). The associated Euler-Lagrange equation is

$$\begin{aligned} -r_3(\partial_x n_1 + \partial_y n_2)_x + r_5 n_1 &= r_5 m_1^{k-1} - \lambda_{51}^{k-1} - (r_3 q^k + \lambda_3^{k-1})_x, \\ -r_3(\partial_x n_1 + \partial_y n_2)_y + r_5 n_2 &= r_5 m_2^{k-1} - \lambda_{52}^{k-1} - (r_3 q^k + \lambda_3^{k-1})_y, \\ n_3 &= m_3^{k-1} - \frac{\lambda_{53}^{k-1}}{r_5}. \end{aligned} \quad (20)$$

The first two equations of (20) is discretized as:

$$\begin{aligned} -r_3(\partial_x^+ \partial_x^- n_1 + \partial_x^+ \partial_y^- n_2) + r_5 n_1 &= r_5 m_1^{k-1} - \lambda_{51}^{k-1} - \partial_x^+(r_3 q^k + \lambda_3^{k-1}), \\ -r_3(\partial_y^+ \partial_x^- n_1 + \partial_y^+ \partial_y^- n_2) + r_5 n_2 &= r_5 m_2^{k-1} - \lambda_{52}^{k-1} - \partial_y^+(r_3 q^k + \lambda_3^{k-1}). \end{aligned}$$

Applying the discrete Fourier transform, we can get

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \mathcal{F}(n_1) \\ \mathcal{F}(n_2) \end{pmatrix} = \begin{pmatrix} \mathcal{F}(h_1) \\ \mathcal{F}(h_2) \end{pmatrix}, \quad (21)$$

where

$$\begin{aligned} a_{11} &= r_5 - r_3(e^{\sqrt{-1}z_i^1} - 1)(1 - e^{-\sqrt{-1}z_i^1}), \\ a_{12} &= -r_3(e^{\sqrt{-1}z_i^1} - 1)(1 - e^{-\sqrt{-1}z_j^2}), \\ a_{21} &= -r_3(e^{\sqrt{-1}z_j^2} - 1)(1 - e^{-\sqrt{-1}z_i^1}), \\ a_{22} &= r_5 - r_3(e^{\sqrt{-1}z_j^2} - 1)(1 - e^{-\sqrt{-1}z_j^2}), \end{aligned}$$

and

$$\begin{aligned} h_1 &= r_5 m_1^{k-1} - \lambda_{51}^{k-1} - \partial_x^+(r_3 q^k + \lambda_3^{k-1}), \\ h_2 &= r_5 m_2^{k-1} - \lambda_{52}^{k-1} - \partial_y^+(r_3 q^k + \lambda_3^{k-1}). \end{aligned}$$

The determinant of the  $2 \times 2$  matrix equals to  $r_5^2 - 2r_5r_3(\cos z_i^1 + \cos z_j^2 - 2)$ , which is positive when  $r_5 > 0$ , so the solution of the problem (21) is easily solved by the discrete inverse Fourier transform. The third equation of (20) can be computed directly.

- For the subproblem  $\xi_4(p)$ , we can easily solve the variable  $p$  based on the soft threshold operation. The minimizer is

$$\mathbf{p} = \max \left\{ 0, 1 - \frac{r_1 + \lambda_1^{k-1}}{r_2 |\tilde{\mathbf{p}}|} \right\} \tilde{\mathbf{p}}, \quad (22)$$

where  $\tilde{\mathbf{p}} = \bar{\mathbf{W}}\mathbf{v}^{k-1} - \frac{\lambda_2^{k-1}}{r_2} + \frac{(r_1 + \lambda_1^{k-1})}{r_2} \mathbf{m}^{k-1}$ .

• For the subproblem  $\xi_5(\mathbf{v})$ , it is a smoothing optimization problem. With a simple computation, its explicit solution can be obtained by

$$\begin{aligned} v_1 &= \frac{r_2 W_x p_1^k + W_x \lambda_{21}^{k-1} + r_4 \partial_x^+ u^k - \lambda_{41}^{k-1}}{r_2 W_x^2 + r_4}, \\ v_2 &= \frac{r_2 W_y p_2^k + W_y \lambda_{22}^{k-1} + r_4 \partial_y^+ u^k - \lambda_{42}^{k-1}}{r_2 W_y^2 + r_4}, \\ v_3 &= \frac{r_2 p_3^k + \lambda_{23}^{k-1} + r_4 - \lambda_{43}^{k-1}}{r_2 + r_4}. \end{aligned} \quad (23)$$

• For the subproblem  $\xi_5(\mathbf{m})$ , we can rewrite it as

$$\min_{\mathbf{m}} \frac{r_5}{2} \int_{\Omega} \left| \mathbf{m} - \left( \mathbf{n}^k + \frac{r_1 + \lambda_1^{k-1}}{r_5} \mathbf{p}^k + \frac{\lambda_5^{k-1}}{r_5} \right) \right|^2 dx dy + \delta_{\mathcal{R}(\mathbf{m})}. \quad (24)$$

Then, we get the solution from project operation that

$$\mathbf{m} = \begin{cases} \tilde{\mathbf{m}}, & \text{if } |\tilde{\mathbf{m}}| \leq 1, \\ \tilde{\mathbf{m}}/|\tilde{\mathbf{m}}|, & \text{if } |\tilde{\mathbf{m}}| > 1, \end{cases} \quad (25)$$

where  $\tilde{\mathbf{m}} = \mathbf{n}^k - \frac{\lambda_5^{k-1}}{r_5} + \frac{(r_1 + \lambda_1^{k-1})}{r_5} \mathbf{p}^k$ .

Finally, the updates of the Lagrangian multipliers  $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$  are given as follows:

$$\lambda_1^k = \lambda_1^{k-1} + r_1 (|\mathbf{p}^k| - \mathbf{p}^k \cdot \mathbf{m}^k), \quad (26)$$

$$\lambda_2^k = \lambda_2^{k-1} + r_2 (\mathbf{p}^k - \bar{\mathbf{W}}\mathbf{v}^k), \quad (27)$$

$$\lambda_3^k = \lambda_3^{k-1} + r_3 (q^k - \partial_x^- n_1^k - \partial_y^- n_2^k), \quad (28)$$

$$\lambda_4^k = \lambda_4^{k-1} + r_4 (\mathbf{v}^k - \nabla u^k), \quad (29)$$

$$\lambda_5^k = \lambda_5^{k-1} + r_5 (\mathbf{n}^k - \mathbf{m}^k). \quad (30)$$

Specifically, we can summarize the augmented Lagrangian algorithm for the AMC model as Algorithm 1.

---

**Algorithm 1:** Augmented Lagrangian algorithm for the AMC model

---

**Input:** The image  $f$  and parameter  $\lambda$ .

**Initialization:**  $u^0, q^0, \mathbf{n}^0, \mathbf{p}^0, \mathbf{v}^0, \mathbf{m}^0$  and  $\lambda_1^0, \lambda_2^0, \lambda_3^0, \lambda_4^0, \lambda_5^0$ .

**while** (*Stopping conditions are not satisfied*) **do**

Computing  $(u^k, q^k, \mathbf{n}^k, \mathbf{p}^k, \mathbf{v}^k, \mathbf{m}^k)$  by (11-16) with the fixed Lagrangian multipliers  $(\lambda_1^{k-1}, \lambda_2^{k-1}, \lambda_3^{k-1}, \lambda_4^{k-1}, \lambda_5^{k-1})$ .

Updating the Lagrangian multipliers  $(\lambda_1^k, \lambda_2^k, \lambda_3^k, \lambda_4^k, \lambda_5^k)$  by (26-30).

**end**

**Output:** Denoising result  $u := u^k$ .

---

**4. Numerical experiments.** In this section, to demonstrate the performance of our proposed model (8), we present a series of numerical implementations to remove noise for the different images. To this end, we choose some state-of-the-art denoising models such as the TV model [26], the high-order total variation model (HOTV) [8], the total generalized variation model (TGV) [4], the anisotropic TV-based model (ATV) [23] and the mean curvature-based model (MC)[39] compared with our proposed model (8). We unify the stopping criterion for all denoising models, that is, all of numerical methods will be stopped when the relative error satisfies  $\frac{\|u^k - u^{k-1}\|_{\ell^1}}{\|u^{k-1}\|_{\ell^1}} \leq 10^{-5}$  or the iteration reaches 500. All comparison algorithms are implemented and tested in Matlab R2020a under Windows OS on a laptop computer equipped with Intel Core i5 2.11GHz.

As is well known, choosing suitable parameters in the proposed model is essential for obtaining good denoising results. For the numerical algorithm, the suitable parameters can improve the convergence rate and the numerical stability. In the proposed model (8), there is a regularization parameter  $\lambda$ . An appropriate  $\lambda$  can give a good balance, filtering out enough noise without losing too much information in the computed solution. In the augmented Lagrangian function (10), there are five penalization parameters  $r_i, i = 1, \dots, 5$ . All the experiments in our paper, the parameters fall into some intervals, specifically,  $\lambda \in [1 \times 10^2, 4 \times 10^3]$ ,  $r_1, r_2, r_4 \in [1, 80]$  and  $r_3, r_5 \in [1 \times 10^4, 3 \times 10^5]$ .

For numerical comparisons, we choose ten classical testing images as shown in Figure 2, where degraded images are generated by adding the white Gaussian noise with different variance as  $\sigma = 0.05, 0.1, 0.2$ . In addition, in order to evaluate the restored ability, we also use the Matlab functions “*snr*(·)” and “*ssim*(·)” as the evaluation standard. In the numerical implementations, parameters obtaining higher SNR are chose.

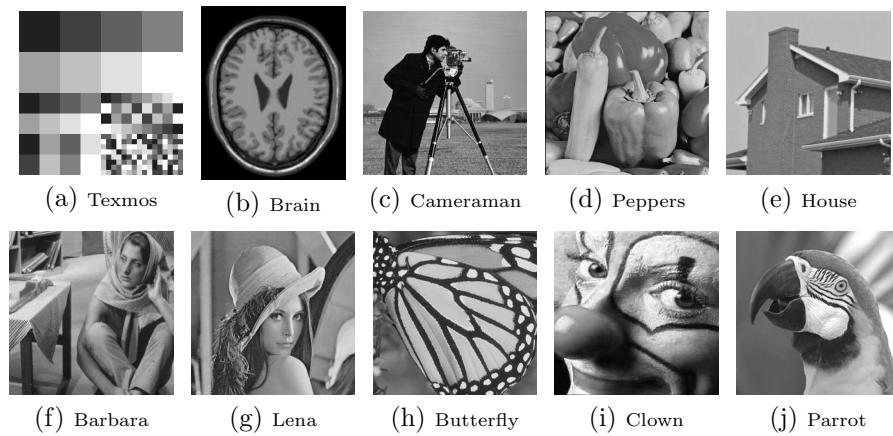


FIGURE 2. Original images.

**4.1. Artificial image.** We start with the piecewise constant image as Texmos degraded by the white Gaussian noise with variance as  $\sigma = 0.05, 0.1, 0.2$ . From the values of the SNR and the SSIM in Table 1, it can be seen that our proposed model (8) performs more robust than other related denoising models. Since humans are involved in evaluating the quality of the denoising result, it becomes absolutely

essential to provide visual quality of the restored images. As shown in Figure 3, we can observe that the images restored by our proposed model have less noise. Especially, some noises can be significantly found in the HOTV model and the TGV model with the noising level of variance  $\sigma = 0.2$ . In order to show denoising results intuitively, we choose the 180th row of the restored images as shown in Figure 4. We can observe from the slice plots that images restored by our proposed model (8) not only contain less noise but also preserve the neat edges and the constant regions efficiently.

		0.05		0.1		0.2	
		SNR	SSIM	SNR	SSIM	SNR	SSIM
Texmos (512 × 512)	TV	34.9389	0.9900	30.7660	0.9864	25.8750	0.9740
	HOTV	31.8945	0.9859	26.3419	0.9657	21.4325	0.9079
	TGV	30.1237	0.9524	27.6509	0.9469	20.2697	0.7842
	ATV	36.5850	0.9959	31.5363	0.9912	26.7421	0.9859
	MC	36.7772	0.9928	32.8587	0.9897	28.1378	0.9859
	AMC	<b>39.5339</b>	<b>0.9974</b>	<b>33.7164</b>	<b>0.9920</b>	<b>28.6771</b>	<b>0.9865</b>

TABLE 1. SNR and SSIM of the Texmos image with different levels of white Gaussian noise.

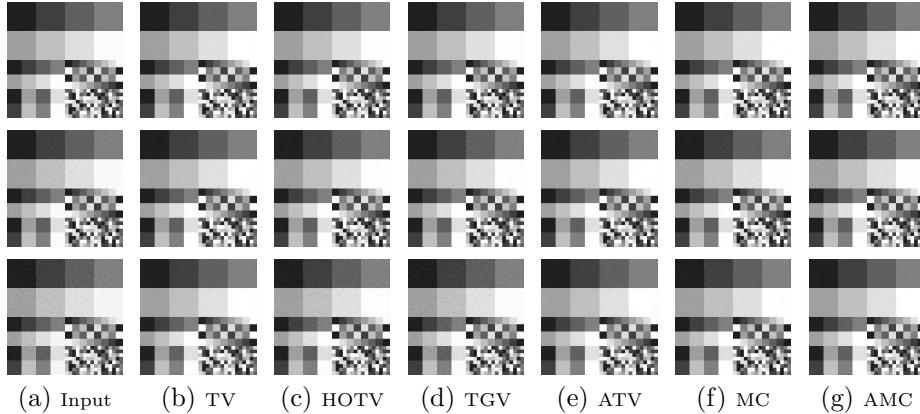


FIGURE 3. Denoising results are shown by different models. Images from left to right are noisy images, denoising images of TV, HOTV, TGV, ATV, MC and the proposed model (AMC). 1st row: the noising level of variance  $\sigma = 0.05$ ; 2nd row: the noising level of variance  $\sigma = 0.1$ ; 3rd row: the noising level of  $\sigma = 0.2$ .

**4.2. Medical image.** This subsection chooses a slice of T1-weighted brain MR image as the testing image, which is contaminated by the white Gaussian noise with different variance as  $\sigma = 0.05, 0.1, 0.2$ . The quantitative comparisons as shown in Figure 5 indicate that our proposed model (8) achieves the best results in terms of the SNR and the SSIM. To show the effectiveness via the visual quality of the brain image, Figure 6 shows the related comparisons between our proposed model and other models. It is easy to observe that all models can provide visual preferable results compared with the noisy images. Specifically, removing the noise by using

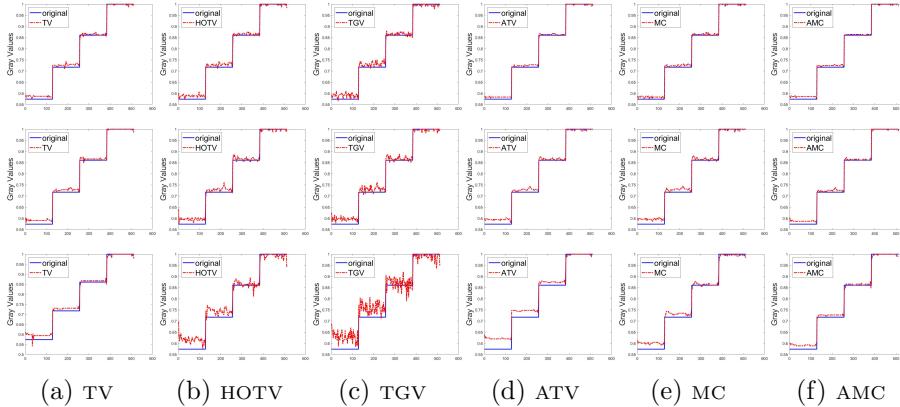


FIGURE 4. The comparisons of the 180th row plots generated from the restored image and the original image. 1st row: restored image with variance  $\sigma = 0.05$ ; 2nd row: restored image with variance  $\sigma = 0.1$ ; 3rd row: restored image with variance  $\sigma = 0.2$ . Images from left to right are restored results of TV, HOTV, TGV, ATV, MC and AMC.

our proposed model and the MC model is more robust than other models. In order to clearly show the restored images, we zoom a part of the brain image degraded by the noise with the standard deviation as 0.2 as an example and then plot them in the bottom column of Figure 6. We can find that restored images by the TV model, HOTV model and TGV model include some noise. Moreover, there is also a little bit of noise in the denoising image by the ATV model. Compared with the MC model, the important difference is that our proposed model present better restored results in edges and give more natural structures.

**4.3. Natural images.** We also select some natural images commonly used in denoising as testing images. For example, Lena contains both sharp edges and smooth regions. Cameraman and peppers include many approximated cartoon regions. So these images are suitable to evaluate our proposed model. To this end, we generate testing images by adding the white Gaussian noise with variance  $\sigma = 0.05, 0.1, 0.2$ . For these degraded images, we evaluate the aforementioned denoising models as shown in Table 2. In general, it concludes that our proposed model outperforms other five models based on these quantitative criterions. To also give a visual impression of these comparisons among different models, we consider the visual assessment of the restored images for cameraman, peppers and house and also include some enlarged regions. Here we only show the noise variance as  $\sigma = 0.2$  to be one example as shown in Figure 7. Obviously, it is easy to observe that our proposed model can efficiently not only remove noise but also keep edges and details, such as, tripod, windows and the edges of peppers.

Figure 8 shows restored images of barbara and lena only for noise density with variance  $\sigma = 0.05$ . It can be seen that the proposed model can successfully remove the most noise and only scattered tiny noise spots left. In order to show the restored abilities, we simultaneously plot the contour curves of the restored image. These

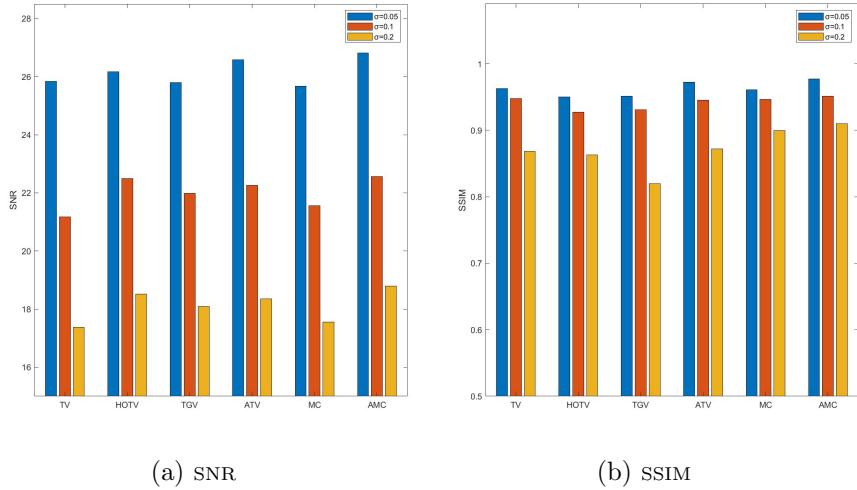


FIGURE 5. Comparisons among six models in terms of the SNR and the SSIM for the brain image.

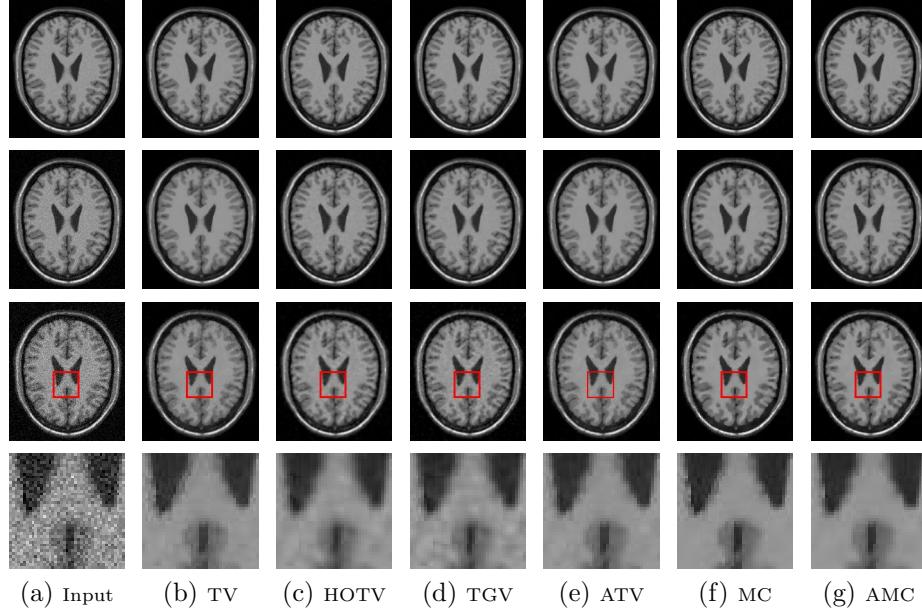


FIGURE 6. Comparisons of six methods on the brain MRI with different levels of noise.

contours clearly reveal that the proposed model, as compared to other models, successfully preserves the details and is closer to the original images.

For testing images of butterfly, clown and parrot, we present the denoising results and plot the corresponding contours in Figure 9 with variance  $\sigma = 0.1$ . We similarly observe that the AMC model outperforms other models. More specifically, the AMC

model has better performance in preserving image structures and reduces the heavy noise effectively. In addition, the different images between the clean images and the restored images are also shown in Figure 9. To show clear difference images, we use image enhancement on the difference images, that is, we adjust the image gray scale and increase the brightness. Especially, we can observe the colorbars of the different images based on our proposed method more dark and less structural information compared with other models, which implies the effectiveness of our proposed method.

In addition, we apply the cameraman image with variance  $\sigma = 0.05, 0.1, 0.2$  as an example to compare the computation time, as shown in Table 3. Compared to the TV-based models, the MC-based models take relatively more computation time. That's mainly because the MC-based models take some time to solve for high order derivatives.

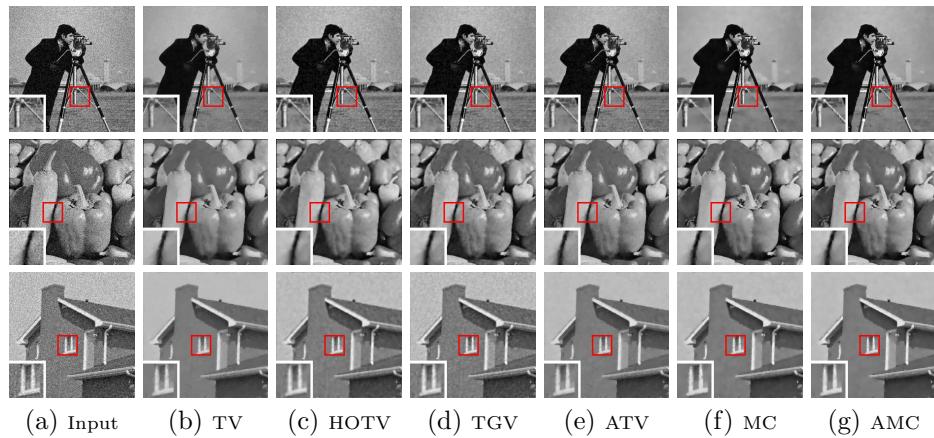


FIGURE 7. Visual examples of restored images with the noise variance as  $\sigma = 0.2$ . The noisy images are displayed in the first column. The second through seventh columns show the restored results of TV, HOTV, TGV, ATV, MC and AMC model.

**4.4. Convergence analysis.** Since our proposed model (8) is non-convex and involves high order derivatives, it is not easy to prove the convergence of the proposed algorithm. In order to show the convergence from the numerical implementation, we need to visually analyze the following relative residuals proposed by [28]:

$$(R_1^k, R_2^k, R_3^k, R_4^k, R_5^k) = \frac{1}{|\Omega|} \left\{ \|\tilde{R}_1^k\|_{\ell^1}, \|\tilde{R}_2^k\|_{\ell^1}, \|\tilde{R}_3^k\|_{\ell^1}, \|\tilde{R}_4^k\|_{\ell^1}, \|\tilde{R}_5^k\|_{\ell^1} \right\}, \quad (31)$$

where

$$\begin{aligned} \tilde{R}_1^k &= |\mathbf{p}^k| - \mathbf{p}^k \cdot \mathbf{m}^k, \tilde{R}_2^k = \mathbf{p}^k - \bar{\mathbf{W}} \mathbf{v}^k, \tilde{R}_3^k = q^k - \partial_x n_1^k - \partial_y n_2^k, \\ \tilde{R}_4^k &= \mathbf{v}^k - (\nabla u^k, 1), \tilde{R}_5^k = \mathbf{n}^k - \mathbf{m}^k. \end{aligned}$$

$|\Omega|$  is the area of domain. In addition, we also compute the relative error as

$$\frac{\|u^k - u^{k-1}\|_{\ell^1}}{\|u^{k-1}\|_{\ell^1}} \quad (32)$$

		0.05		0.1		0.2	
		SNR	SSIM	SNR	SSIM	SNR	SSIM
Cameraman (256×256)	TV	24.4261	0.9403	19.8933	0.8741	15.4665	0.7645
	HOTV	24.7059	0.9389	20.0565	0.8692	15.8390	0.7175
	TGV	24.8285	0.9407	20.2318	0.8660	16.0555	0.7590
	ATV	24.8216	0.9405	19.5423	0.8163	16.1089	0.7329
	MC	24.2884	0.9438	19.7386	0.8844	15.7260	0.8098
	AMC	<b>24.8751</b>	<b>0.9452</b>	<b>20.2513</b>	<b>0.8882</b>	<b>16.8309</b>	<b>0.8301</b>
Peppers (256×256)	TV	23.4557	0.9509	19.0540	0.9031	15.3310	0.8472
	HOTV	23.6027	0.9482	19.4940	0.9123	15.4403	0.8566
	TGV	23.6712	0.9487	19.7737	0.9127	15.4945	0.8449
	ATV	23.5929	0.9508	19.3028	0.8929	15.7272	0.8423
	MC	23.4629	0.9523	19.4053	0.9137	15.9017	0.8659
	AMC	<b>23.8056</b>	<b>0.9535</b>	<b>20.1047</b>	<b>0.9204</b>	<b>16.4739</b>	<b>0.8687</b>
Barbara (256×256)	TV	21.0820	0.9428	17.6311	0.8804	14.2412	0.8008
	HOTV	22.1298	0.9388	17.7057	0.8820	14.6666	0.8107
	TGV	21.7975	0.9455	17.7752	0.8663	14.7037	0.7913
	ATV	22.1881	0.9431	17.7769	0.8856	14.6581	0.8062
	MC	21.7009	0.9425	17.7023	0.8846	14.4138	0.8155
	AMC	<b>22.2248</b>	<b>0.9458</b>	<b>17.7859</b>	<b>0.8878</b>	<b>14.7221</b>	<b>0.8178</b>
House (256×256)	TV	22.7621	0.9302	19.1033	0.8786	16.1751	0.8317
	HOTV	22.8041	0.9263	19.3590	0.8724	15.9549	0.8131
	TGV	22.7787	0.9268	19.4610	0.8772	15.2764	0.7463
	ATV	22.9113	0.9311	19.2093	0.8832	16.4647	0.8369
	MC	22.3283	0.9200	19.4971	0.8845	16.4433	0.8408
	AMC	<b>23.0546</b>	<b>0.9313</b>	<b>19.7522</b>	<b>0.8857</b>	<b>16.9050</b>	<b>0.8483</b>
Lena (256×256)	TV	21.9420	0.9385	17.8192	0.8757	14.2216	0.7932
	HOTV	21.9930	0.9384	17.8427	0.8759	14.1715	0.8014
	TGV	21.7423	0.9405	17.9651	0.8751	14.0864	0.7777
	ATV	21.9480	0.9386	17.8277	0.8760	14.2312	0.8132
	MC	21.9025	0.9408	17.8036	0.8914	14.1209	0.8108
	AMC	<b>22.1449</b>	<b>0.9413</b>	<b>17.9969</b>	<b>0.8924</b>	<b>14.3413</b>	<b>0.8196</b>
Parrot (256×256)	TV	25.1037	0.9472	20.7407	0.9100	17.3382	0.8376
	HOTV	23.6234	0.9489	19.5562	0.9107	15.5233	0.8419
	TGV	25.0737	0.9415	21.3407	0.9063	17.0111	0.7845
	ATV	25.1060	0.9472	20.8167	0.8796	17.5184	0.8586
	MC	25.1114	0.9472	21.0438	0.9125	17.5632	0.8706
	AMC	<b>25.2144</b>	<b>0.9512</b>	<b>21.3952</b>	<b>0.9151</b>	<b>17.9631</b>	<b>0.8707</b>
Clown (256×256)	TV	25.4481	0.9541	21.0520	0.8967	17.2097	0.8068
	HOTV	25.3700	0.9460	21.0189	0.9019	17.2749	0.8228
	TGV	25.1128	0.9572	21.0398	0.8873	17.4512	0.8091
	ATV	25.4768	0.9501	21.0730	0.8969	17.4509	0.8104
	MC	25.2299	0.9557	21.0721	0.9066	16.9260	0.8210
	AMC	<b>25.6417</b>	<b>0.9578</b>	<b>21.3543</b>	<b>0.9085</b>	<b>17.4636</b>	<b>0.8322</b>
Butterfly (256×256)	TV	24.8939	0.9692	20.3859	0.9285	16.0945	0.8755
	HOTV	24.8206	0.9663	20.1207	0.9394	15.5635	0.8768
	TGV	24.7823	0.9656	20.2884	0.9414	15.7252	0.8538
	ATV	24.8984	0.9692	20.3992	0.9288	16.3346	0.8908
	MC	24.5480	0.9713	20.5937	0.9457	16.5272	0.9099
	AMC	<b>25.2768</b>	<b>0.9729</b>	<b>20.9997</b>	<b>0.9495</b>	<b>16.9803</b>	<b>0.9126</b>

TABLE 2. SNR and SSIM of the natural images with different levels of white Gaussian noise.

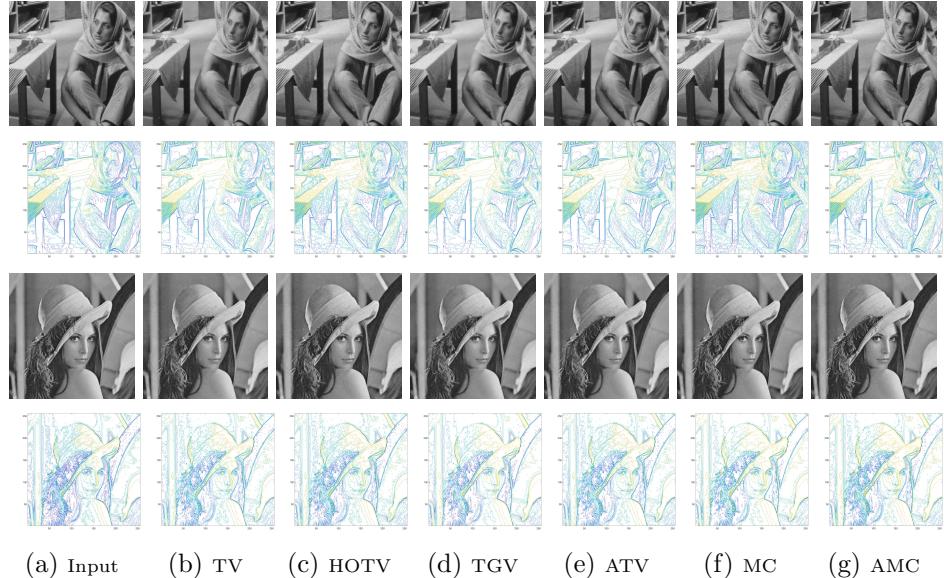


FIGURE 8. Denoising results of six models. Images from left to right are original images, restored results of TV, HOTV, TGV, ATV, MC and the proposed model (AMC). The first and third rows are the restored images in order to efficiently show the restored results. The second and fourth rows are the contour plots. White Gaussian noise with variance as  $\sigma = 0.05$ .

		0.05	0.1	0.2
Time(s)	TV	0.761	0.736	0.781
	HOTV	9.910	9.326	7.181
	TGV	18.405	19.147	18.587
	ATV	1.041	1.378	2.195
	MC	41.277	43.530	44.150
	AMC	42.806	43.994	42.667

TABLE 3. The computation time of different models for the cameraman image with variance as  $\sigma = 0.05, 0.1, 0.2$ .

to show the convergence of the proposed numerical algorithm. Besides all these quantities, we also consider how the energy in (8) is evolving during the iteration by tracking the value of  $E(u^k)$ . For the purpose of presentation, all of the above quantities are shown in log-scale in the Figures.

Here we apply the brain image with  $\sigma = 0.2$  as an example. In Figure 10, we plot the relative residuals (31), relative error in  $u^k$  (32) and the energy versus iteration. As we can see from Figure 10 that the relative residuals, relative error in  $u^k$  and the numerical energy decay when increasing the iteration, which demonstrate that our proposed algorithm converges well numerically.

**5. Conclusion.** In this paper, we proposed a new model for the image denoising problem. To increase the anisotropy for the model, we added a weighted matrix

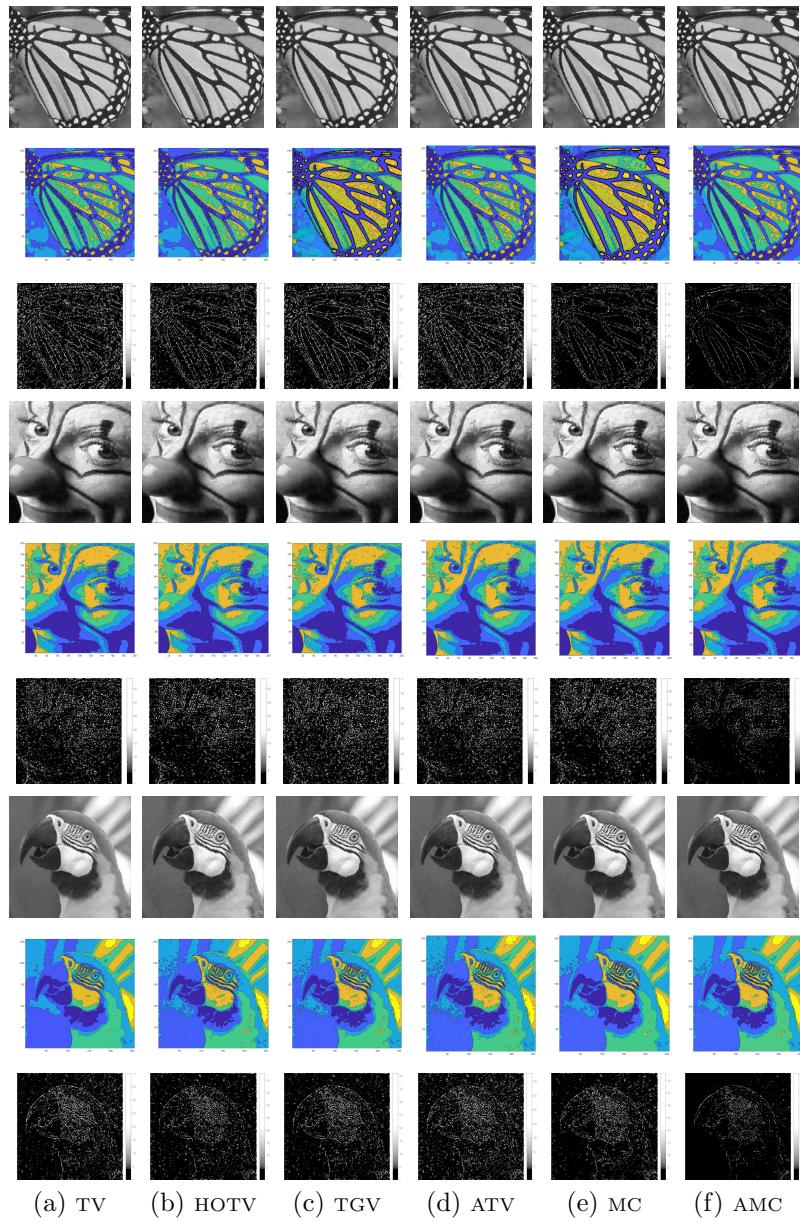


FIGURE 9. Visual examples of restored images with the noise variance as  $\sigma = 0.1$ . The first, fourth and seventh rows show the denoising results by using the different models. The second, fifth and eighth rows show the contours. The third, sixth and ninth rows show the difference images between restored images and the original clean images.

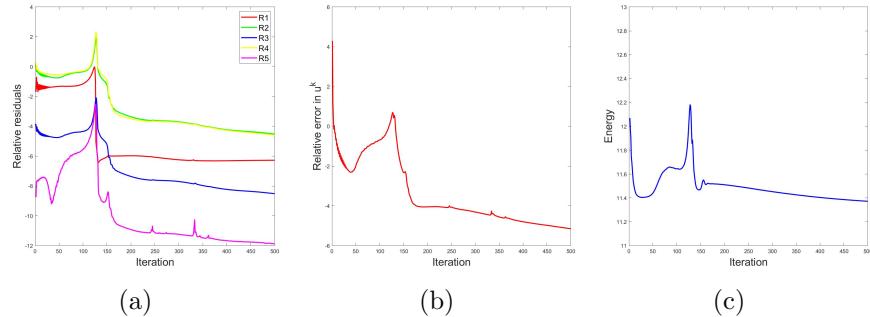


FIGURE 10. The plots of relative residuals, relative error in  $u^k$  and energy for the brain image with variance as  $\sigma = 0.2$ .

to make the gradient direction along with the tangential direction of the edge and then the proposed model can describe the local structures of the image. In addition, we used the alternating direction method of multipliers to decouple the gradient operator and the weighted matrix from the  $L^2$ -norm and then the efficient numerical schemes can be used to solve the related subproblems. Numerical experiments exhibited that the proposed model yields good performance in terms of both quantitative and qualitative criteria.

**Acknowledgments.** This work was partially supported by the Natural Science Foundation of China (No.12071345), the Science and Technology Major Project of Henan Province (No.221100310200), the Programs for Science and Technology Development of Henan Province (No.212102210511, 212102310732) and the SKY research programme (grant: Z-2014-07-2101).

## REFERENCES

- [1] E. Altinsoy, J. Yang and E. Tu, [An improved denoising of G-banding chromosome images using cascaded CNN and binary classification network](#), *The Visual Computer*, **38** (2022), 2139-2152.
- [2] L. Ambrosio and S. Masnou, [A direct variational approach to a problem arising in image reconstruction](#), *Interfaces and Free Boundaries*, **5** (2003), 63-81.
- [3] J. Bai and X.-C. Feng, [Fractional-order anisotropic diffusion for image denoising](#), *IEEE Transactions on Image Processing*, **16** (2007), 2492-2502.
- [4] K. Bredies, K. Kunisch and T. Pock, [Total generalized variation](#), *SIAM Journal on Imaging Sciences*, **3** (2010), 492-526.
- [5] A. Chambolle and P.-L. Lions, [Image recovery via total variation minimization and related problems](#), *Numerische Mathematik*, **76** (1997), 167-188.
- [6] T. F. Chan, S. Esedoglu and F. Park, [A fourth order dual method for staircase reduction in texture extraction and image restoration problems](#), *2010 IEEE International Conference on Image Processing*, 2010, 4137-4140.
- [7] R. H. Chan, H. Liang, S. Wei, M. Nikolova and X.-C. Tai, [High-order total variation regularization approach for axially symmetric object tomography from a single radiograph](#), *Inverse Problems and Imaging*, **9** (2015), 55-77.
- [8] T. Chan, A. Marquina and P. Mulet, [High-order total variation-based image restoration](#), *SIAM Journal on Scientific Computing*, **22** (2000), 503-516.
- [9] C. Cruz, A. Foi, V. Katkovnik and K. Egiazarian, [Nonlocality-reinforced convolutional neural networks for image denoising](#), *IEEE Signal Processing Letters*, **25** (2018), 1216-1220.
- [10] J. Duan, Z. Qiu, W. Lu, G. Wang, Z. Pan and L. Bai, [An edge-weighted second order variational model for image decomposition](#), *Digital Signal Processing*, **49** (2016), 162-181.

- [11] G. Gilboa and S. Osher, **Nonlocal operators with applications to image processing**, *Multiscale Modeling and Simulation*, **7** (2009), 1005-1028.
- [12] S. Goyal, V. Singh, A. Rani and N. Yadav, **Multimodal image fusion and denoising in NSCT domain using CNN and FOTGV**, *Biomedical Signal Processing and Control*, **71** (2022), 103214.
- [13] R. Kimmel, R. Malladi and N. Sochen, **Images as embedded maps and minimal surfaces: Movies, color, texture, and volumetric medical images**, *International Journal of Computer Vision*, **39** (2000), 111-129.
- [14] S. Lefkimiatis and M. Unser, **Poisson image reconstruction with Hessian Schatten-norm regularization**, *IEEE Transactions on Image Processing*, **22** (2013), 4314-4327.
- [15] S. Lefkimiatis, J. P. Ward and M. Unser, **Hessian Schatten-norm regularization for linear inverse problems**, *IEEE Transactions on Image Processing*, **22** (2013), 1873-1888.
- [16] F. Li, C. Shen, J. Fan and C. Shen, **Image restoration combining a total variational filter and a fourth-order filter**, *Journal of Visual Communication and Image Representation*, **18** (2007), 322-330.
- [17] J. Liu, R. W. Liu, J. Sun and T. Zeng, **Rank-one prior: Real-time scene recovery**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
- [18] W. Lu, J. Duan, Z. Qiu, Z. Pan, R. W. Liu and L. Bai, **Implementation of high-order variational models made easy for image processing**, *Mathematical Methods in the Applied Sciences*, **39** (2016), 4208-4233.
- [19] M. Lysaker, A. Lundervold and X.-C. Tai, **Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time**, *IEEE Transactions on Image Processing*, **12** (2003), 1579-1590.
- [20] X. Mao, C. Shen and Y. Yang, **Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections**, *Advances in Neural Information Processing Systems*, **29** (2016), 2802-2810.
- [21] S. Osher, A. Sole and L. Vese, **Image decomposition and restoration using total variation minimization and the  $H^{-1}$  norm**, *Multiscale Modeling and Simulation*, **1** (2003), 349-370.
- [22] Z.-F. Pang, L.-Z. Guo, Y. Duan and J. Lu, **Image restoration based on the minimized surface regularization**, *Computers and Mathematics with Applications*, **76** (2018), 1893-1905.
- [23] Z.-F. Pang, Y.-M. Zhou, T. Wu and D.-J. Li, **Image denoising via a new anisotropic total-variation-based model**, *Signal Processing: Image Communication*, **74** (2019), 140-152.
- [24] K. Papafitsoros and C. B. Schönlieb, **A combined first and second order variational approach for image reconstruction**, *Journal of Mathematical Imaging and Vision*, **48** (2014), 308-338.
- [25] Y. Quan, Y. Chen, Y. Shao, H. Teng, Y. Xu and H. Ji, **Image denoising using complex-valued deep CNN**, *Pattern Recognition*, **111** (2021), 107639.
- [26] L. I. Rudin, S. Osher and E. Fatemi, **Nonlinear total variation based noise removal algorithms**, *Physica D: Nonlinear Phenomena*, **60** (1992), 259-268.
- [27] O. Scherzer, **Denoising with higher order derivatives of bounded variation and an application to parameter estimation**, *Computing*, **60** (1998), 1-27.
- [28] X.-C. Tai, J. Hahn and G. J. Chung, **A fast algorithm for Eulers Elastica model using augmented Lagrangian method**, *SIAM Journal on Imaging Sciences*, **4** (2011), 313-344.
- [29] C. Tian, Y. Xu, W. Zuo, B. Du, C.-W. Lin and D. Zhang, **Designing and training of a dual CNN for image denoising**, *Knowledge-Based Systems*, **226** (2021), 106949.
- [30] L. A. Vese and S. J. Osher, **Modeling textures with total variation minimization and oscillatory patterns in image processing**, *Journal of Scientific Computing*, **19** (2003), 553-579.
- [31] G. Wang, J. Xu, Q. Dong and Z. Pan, **Active contour model coupling with higher order diffusion for medical image segmentation**, *International Journal of Biomedical Imaging*, 2014.
- [32] Y. Wang, J. Yang, W. Yin and Y. Zhang, **A new alternating minimization algorithm for total variation image reconstruction**, *SIAM Journal on Imaging Sciences*, **1** (2008), 248-272.
- [33] D. Yang and J. Sun, **Bm3d-net: A convolutional neural network for transform-domain collaborative filtering**, *IEEE Signal Processing Letters*, **25** (2017), 55-59.
- [34] J. Yang, M. Ma, J. Zhang and C. Wang, **Noise removal using an adaptive Euler's elastica-based model**, *The Visual Computer*, 2022, 1-12.
- [35] Y.-L. You and M. Kaveh, **Fourth-order partial differential equations for noise removal**, *IEEE Transactions on Image Processing*, **9** (2000), 1723-1730.
- [36] J. Zhang, R. Chen, C. Deng and S. Wang, **Fast linearized augmented Lagrangian method for Euler's elastica model**, *Numerical Mathematics: Theory, Methods and Applications*, **10** (2017), 98 -115.

- [37] J. Zhang and Z. Wei, *A class of fractional-order multi-scale variational models and alternating projection algorithm for image denoising*, *Applied Mathematical Modelling*, **35** (2011), 2516-2528.
- [38] W. Zhu, *Image denoising using  $L^p$ -norm of mean curvature of image surface*, *Journal of Scientific Computing*, **83** (2020), No. 32, 26 pp.
- [39] W. Zhu and T. Chan, *Image denoising using mean curvature of image surface*, *SIAM Journal on Imaging Sciences*, **5** (2012), 1-32.
- [40] W. Zhu, X.-C. Tai and T. Chan, *Augmented Lagrangian method for a mean curvature based image denoising model*, *Inverse Problems and Imaging*, **7** (2013), 1409-1432.

Received June 2022; 1st revision December 2022; 2nd revision February 2023;  
early access March 2023.