EURASIP Journal on
Advances in Signal Processing
a SpringerOpen Journal

RESEARCH **Open Access**

CrossMark

Edge and contrast preserving in total variation image denoising

Liming Tang[*] and Zhuang Fang

Abstract

Total variation (TV) regularization can very well remove noise and simultaneously preserve the sharp edges. But it has the drawback of the contrast loss in the restoration. In this paper, we first theoretically analyze the loss of contrast in the original TV regularization model, and then propose a forward-backward diffusion model in the framework of total variation, which can effectively preserve the edges and contrast in TV image denoising. A backward diffusion term based on a nonconvex and monotony decrease potential function is introduced in the TV energy, resulting in a forward-backward diffusion. In order to finely control the strength of the forward and backward diffusion, and separately design the efficient algorithm to numerically implement the forward and backward diffusion, we propose a two-step splitting method to iteratively solve the proposed model. We adopt the efficient projection algorithm in the dual framework to solve the forward diffusion in the first step, and then use the simple finite differences scheme to solve the backward diffusion to compensate the loss of contrast occurred in the previous step. At last, we test the models on both synthetic and real images. Compared with the classical TV, forward and backward diffusion (FBD), two-step methods (TSM), and TV-FF models, our model has the better performance in terms of peak signal-to-noise ratio (PSNR) and mean structural similarity (MSSIM) indexes.

Keywords: Total variation, Forward diffusion, Backward diffusion, Image denoising, Contrast

1 Introduction

Image denoising plays an important role in various applied areas, such as pattern recognition, medical imaging, remote sensing, video processing, and so on. So far, image denoising has seen many experienced vigorous developments, and emerged in a number of important theories and research results [1], e.g., spatial filtering method [2–4]; transform domain filtering method [5–7]; and PDE-based method [8, 9].

In this paper, we focus on variational method that achieves image denoising by functional regularization and minimization [10, 11]. Since the total variation (TV) minimization model was proposed by Rudin, Osher, and Fatemi [12] in 1992, variational method has attracted more and more research attention due to its sound theoretical basis and good experimental results, where the TV of u TV (u) is defined as

Definition 1 Let $\Omega \in \mathrm{R}^2$ be an open subset with Lipschitz boundary, and $u(\boldsymbol{x}) : \Omega \to \mathrm{R}$. Then TV$(u)$ is defined as

$$
\begin{aligned}
\mathrm{TV}(u) &= \int_{\Omega} |\nabla u| d\boldsymbol{x} \\
&= \sup \left\{ \int_{\Omega} u \mathrm{div}(\boldsymbol{\phi}) d\boldsymbol{x}, \boldsymbol{\phi} \in C_C^1\left(\Omega, \mathrm{R}^2\right), \|\boldsymbol{\phi}\|_{\infty} \le 1 \right\}
\end{aligned}
$$

where div represents divergence operator. Further, TV$(u) + \|u\|_{L^1(\Omega)}$ is known as the bounded variation (*BV*) of u, and TV(u) is so called *BV*-seminorm.

Using TV(u) to measure the restoration, Rudin et al. [12] proposed the *f*ollowing constrained minimization problem called TV model,

$$
\min_u \int_{\Omega} |\nabla u| d\boldsymbol{x}, \quad \text{subject to } \|u - u_0\|_{L^2(\Omega)}^2 = \sigma^2 \qquad (1)
$$

where σ is the standard deviation of noise which is assumed to be known. In order to numerically solve the constrained minimization problem (1), Rudin et al. [12] transformed it to an unconstrained problem:

* Correspondence: tlmcs78@foxmail.com
School of Science, Hubei University for Nationalities, Enshi 445000, People's Republic of China

$$\min_{u}\left\{\alpha\int_{\Omega}|\nabla u|d\boldsymbol{x}+\|u-u_0\|^2_{L^2(\Omega)}\right\} \qquad (2)$$

and solved it by a simple finite difference scheme. Chambolle and Lions [13] showed that problem (1) is equivalent to problem (2) when $\alpha = 1/\lambda$, where λ is the Lagrange multiplier found in solving (1). TV model (2) is convex and easy to solve in practice. In addition, the solution u allows for discontinuities along curves; therefore, edges and contours can be preserved in the restored image.

Later on, many scholars researched into the TV model (2), and proposed lots of improved TV-based models (e.g., [14–17]). We note that many of the current TV-based models have good performance in noise removing and edge preserving. However, these models often cause "loss of contrast" effects in the restoration u [18, 19]. Two examples of loss of contrast in restorations from TV model (2) can be observed in Fig. 1 for the 1-d and 2-d cases.

There are some improved methods to preserve the contrast in TV regularization, which can be categorized into three major classes: two-step methods, partial differential equation (PDE)-based methods and variation methods.

Two-step methods (TSM) remove the noise and enhance the contrast on two separate steps [20]. Some two-step methods firstly remove the noise by TV regularization, and then enhance the contrast in the restoration obtained from the previous regularization step by some classical enhancement methods, such as histogram equalization or gray-scale transformation algorithms. These methods have the drawbacks of losing weak edges. While, there are also some methods that implement the restoration and enhancement in the reverse order, i.e., first enhancement and then regularization. These methods have the drawbacks of poor denoising ability.

PDE-based methods enhance the contrast by evolving a PDE or PDEs. Here, we would like to list some classical PDE-based models in contrast enhancement which are the motivation of our present works. Osher and Rudin [21] proposed a shock filter to enhance the contrast.

$$\begin{cases} u_t = -\text{sign}(u_{\eta\eta})|\nabla u| \\ u(x, t = 0) = u_0 \\ u_{\eta\eta} = \dfrac{u_{xx}u_x^2 + 2u_xu_yu_{xy} + u_{yy}u_y^2}{|\nabla u|^2} \end{cases}$$

Alvarez and Mazorra [22] proposed a regularization shock filter to enhance the contrast for noisy image.

$$\begin{cases} u_t = -\text{sign}(G_\sigma * u_{\eta\eta})|\nabla u| \\ u(x, t = 0) = u_0 \\ u_{\eta\eta} = \dfrac{u_{xx}u_x^2 + 2u_xu_yu_{xy} + u_{yy}u_y^2}{|\nabla u|^2} \end{cases}$$

Weickert [9] proposed a tensor diffusion (i.e., anisotropic diffusion) model,

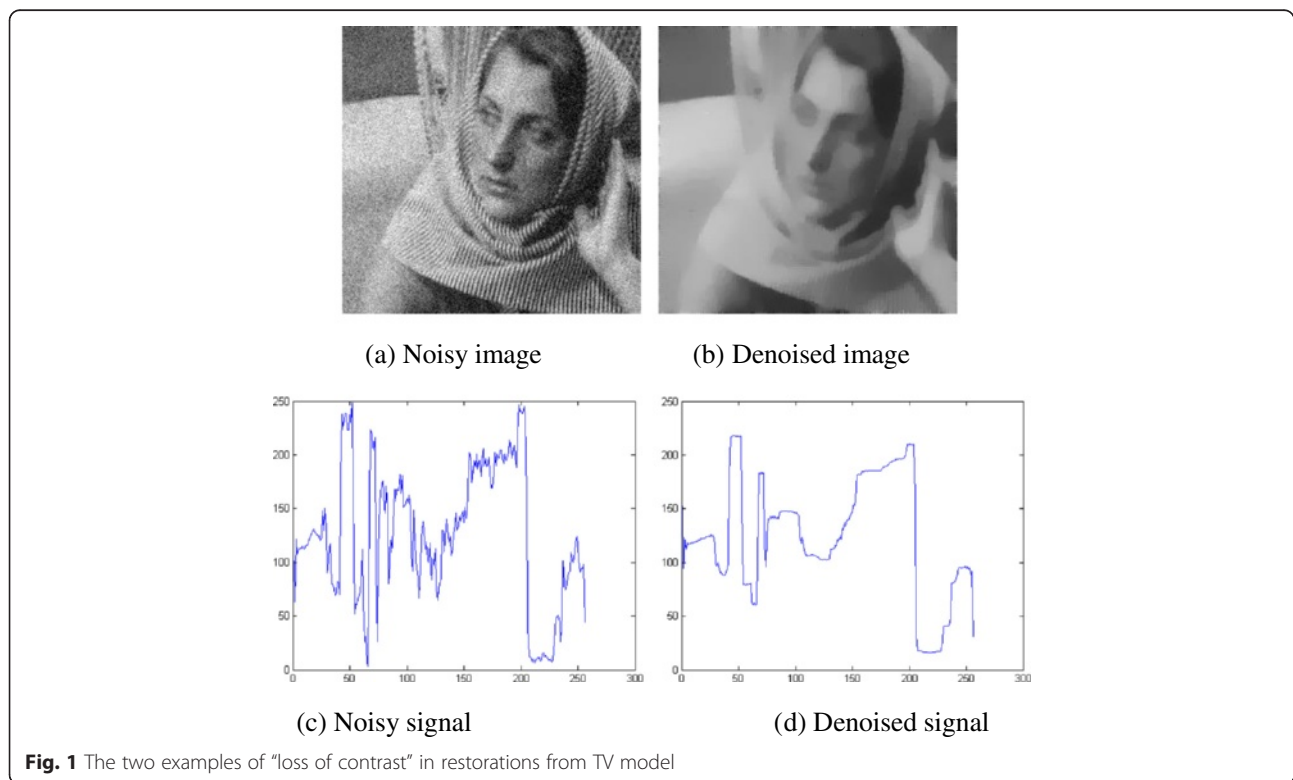(a) Noisy image   (b) Denoised image

(c) Noisy signal   (d) Denoised signal

Fig. 1 The two examples of "loss of contrast" in restorations from TV model

$$\begin{cases} u_t = \mathrm{div}(\boldsymbol{D}\nabla u) \\ u(x, t = 0) = u_0 \end{cases}$$

where $\boldsymbol{D}$ is diffusion tensor. Guy Gilboa et al. [23] proposed forward and backward diffusion (FBD) to simultaneously remove the noise and enhance the contrast.

$$\begin{cases} u_t = \mathrm{div}(-g(|\nabla u|)\nabla u) \\ u(x, t = 0) = u_0 \end{cases}$$

where $g(|\nabla u|)$ is a decreasing function of the gradient. Actually, the above PDE-based methods all adapt the backward diffusion to enhance the contrast. i.e., the diffusion coefficients in these diffusion equations are negative. In addition, there also have some methods that implement the classical image enhancement methods by PDE. For example, in [24], the authors proposed a PDE-based approach to perform global histogram equalization. In [25], the authors implemented the scale transformation by PDE.

So far, variation methods achieve noise removing and contrast preserving mainly by adapting the fidelity term in TV regularization. For example, in [26], a stretching function f is employed in fidelity term in TV regulation to enhance the contrast (TV-FF), where the new fidelity term is defined as $\|g(u) - u_0\|_{L^2(\Omega)}^2$ with $g = f^{-1}$. In [27], the authors proposed a fidelity term based on image gradient, which is defined as $\||\nabla u| - k|\nabla u_0|\|_{L^2(\Omega)}^2$ where k is a constant and satisfies $k \geq 1$. In [28], the authors proposed a common formulation using the image gradient, $\||\nabla u| - k T_\varepsilon(|\nabla G_\sigma * u_0|)\|_{L^2(\Omega)}^2$ where T_ε is a piecewise stretching function and G_σ is a Gaussian kernel with the standard deviation σ.

In this paper, based on variation method and backward diffusion, we propose a forward-backward diffusion model in the framework of TV (called TV-FBD). Compared to the tradition forward-backward diffusion models, our model has the following characteristics and advantages.

(1) Tradition forward-backward diffusion models are PDE based, while our diffusion model is based on variation. So, our model has the better extensibility than the tradition PDE based models. Some image information (such as gradient, direction of edges, textures, and so on) can easily be incorporated into the energy to improve the restoration quality, while it is hard to do that in PDE based models.

(2) Tradition forward-backward diffusion models are often implemented by finite difference scheme or multi-grid technique whose computing efficiency is always low in practice. While our variation model can be solved by some faster modern convex optimization algorithms that have higher efficiency.

The rest of this paper is organized as follows. In Section 2, we theoretically analyze the loss of contrast in TV regularization for piecewise constant functions. In Section 3, we present our forward-backward diffusion model in the framework of TV (i.e., TV-FBD model). In Section 4, we introduce a two-step splitting method for the proposed model. The numerical results showing the performance of the proposed model are given in Section 5. This paper is summarized in Section 6.

2 The loss of contrast in TV regularization

In this section, we analyze the loss of contrast in TV regularization for radially symmetric piecewise constant functions. We do this because image features are often partially or entirely composed by piecewise constant, or the 'limit' of piecewise constant functions. Another important reason is that we can find exact results for radially symmetric piecewise constant case, which are impossible to derive in the general case.

Firstly, with radially symmetry, we rewrite the TV regularization problem as a general mode,

$$\min_u \int \frac{1}{2}(u(r) - u_0(r))^2 + \alpha |u_r(r)| d\Omega(r) \qquad (3)$$

where $u_r(r)$ is the directional derivative of $u(r)$ with respect to r; and $d\Omega(r)$ is the infinitesimal element which satisfies $d\Omega(r) = dr$ in R^1, $d\Omega(r) = 2\pi r dr$ in R^2 and $d\Omega(r) = 4\pi r^2 dr$ in R^3, respectively. We note that the true image and its corresponding noisy version $u_0(r)$ are radially symmetric. So in this case, the noise present in the image is also supposed as radially symmetry. In general, noise is not radially symmetric. The reason why we make this assumption is to make the mathematical analysis and results possible.

Any piecewise constant functions are comprised of two types of features: "steps" and "extrema" regions, as illustrated in Fig. 2. So in what follows, we only study the effects of TV regularization when $u_0(r)$ is a monotonic steps function and a unimodal function, separately.

Proposition 1 (Monotonic steps function) Let $u_0(r)$ be defined on $[r_0, r_3]$. In addition, suppose that:

1. $u_0(r)$ is a monotonic steps function, i.e., $U_i \geq U_{i+1}$ for any $1 \leq i \leq 3$, where

$$U_i = \frac{\displaystyle\int_{r_{i-1}}^{r_i} u_0(r) d\Omega(r)}{\displaystyle\int_{r_{i-1}}^{r_i} d\Omega(r)}$$

that is actually the mean of $u_0(r)$ in the region of $\Omega_{r_{i-1}, r_i} = \Omega_{r_i} - \Omega_{r_{i-1}}$.

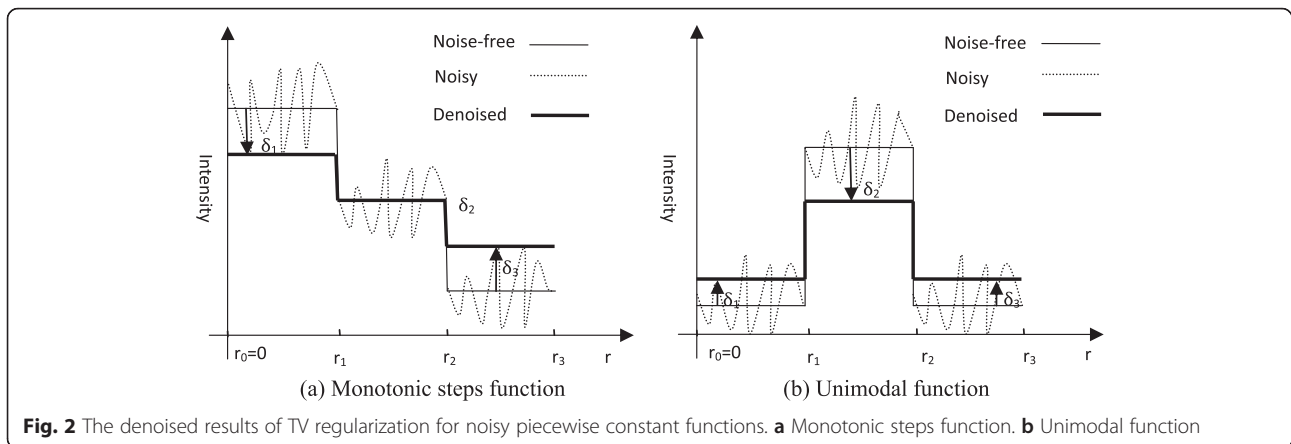2. $U_i + \delta_i \geq U_{i+1} + \delta_{i+1}$ for any $0 \leq i \leq 3$.

Fig. 2 The denoised results of TV regularization for noisy piecewise constant functions. **a** Monotonic steps function. **b** Unimodal function

3. $\max\{|u_0(r) - U|\} \geq |\delta_i| \geq \max_{0 \leq i \leq 3}\{|u_0(r) - U_i|\}$ for any $1 \leq i \leq 3$, where

$$U = \frac{\int_{r_0}^{r_3} u_0(r)d\Omega(r)}{\int_{r_0}^{r_3} d\Omega(r)}$$

that is actually the mean of $u_0(r)$ in the entire region Ω.

Then the solution of (3) can be written as

$$u(r) = U_i + \delta_i \text{ for } r \in [r_{i-1}, r_i] \text{ with } 1 \leq i \leq 3.$$

where

$$\delta_i = \begin{cases} \frac{\alpha(|\partial\Omega_{r_{i-1}}| - |\partial\Omega_{r_i}|)}{|\Omega_{r_{i-1},r_i}|}, & \text{if } \frac{|\alpha(|\partial\Omega_{r_{i-1}}| - |\partial\Omega_{r_i}|)|}{|\Omega_{r_{i-1},r_i}|} \leq \max\{|u_0(r) - U|\} \\ U - U_i, & \text{if } \frac{|\alpha(|\partial\Omega_{r_{i-1}}| - |\partial\Omega_{r_i}|)|}{|\Omega_{r_{i-1},r_i}|} \geq \max\{|u_0(r) - U|\} \end{cases}$$

(4)

We note that Ω is the symmetry interval in R^1, the circle in R^2, or the sphere in R^3; $|\partial\Omega|$ is the boundary of Ω; $|\Omega|$ is the length of the symmetry interval in R^1, the area of the circle in R^2, or the volume of the sphere in R^3.

Remark 1 For Proposition 1, we make a few notes. (1) We refer the reader to Theorem 1 in [18] for a similar proof of this proposition. (2) If the mean of noise added to each region is zero, then the most ideal restoration (i.e., true image) is a monotonic steps function whose discontinuities are at $\{r_i\}$, and the value in region Ω_{r_{i-1},r_i} is U_i. (3) Proposition 1 only shows the results for monotonically decreasing step function. Actually, the results for monotonically increasing step functions are analogous. The only difference between them is that the changes in intensity over each region are of opposite sign. (4) For α sufficiently large (but finite), from the Eq. (4), we note that the regularized image $u(r)$ is

constant and is simply the mean of the observed image $u_0(r)$ over the entire domain Ω. In this case, the contrast of the regularized image $u(r)$ is zero.

Example 1 In the simplest case of R^1 function, we assume that (1) $r_0 = 0$; (2) $|\partial\Omega_{r_0}| = 0$ and $|\partial\Omega_{r_3}| = 0$ (Neumann boundary conditions); and (3) $|r_0r_1| = |r_1r_2| = |r_2r_3| = r$ (equidistant interval). In this case, the change in function intensity is given by $\delta_1 = -\alpha/r$, $\delta_2 = 0$ and $\delta_3 = \alpha/r$; and the regularized image is represented as

$$u(r) = U_i + \delta_i = \begin{cases} U_1 - \dfrac{\alpha}{r}, & i = 1 \\ U_2, & i = 2 \\ U_3 + \dfrac{\alpha}{r}, & i = 3 \end{cases}$$

It is obvious that the contrasts (i.e., the scale of the discontinuities at $\{r_i\}$) are less than they were in the true image, as illustrated in Fig. 2a. And the total loss of contrasts is $2\alpha/r$.

Example 2 We extend our results in Example 1 to the R^2 functions. In this case, Ω_{r_i}s are the concentric circles centered at $r = 0$ and with radius of r_i. Under the same conditions as Example 1, by Proposition 1, we have $\delta_1 = -2\alpha/r$, $\delta_2 = -2\alpha/3r$, and $\delta_3 = 4\alpha/5r$. Then the regularized image is represented as

$$u(r) = U_i + \delta_i = \begin{cases} U_1 - \dfrac{2\alpha}{r}, & i = 1 \\ U_2 - \dfrac{2\alpha}{3r}, & i = 2 \\ U_3 + \dfrac{4\alpha}{5r}, & i = 3 \end{cases}$$

From the last equation, we obtain that the total loss of contrasts for concentric circles in R^2 is $14\alpha/5r$. Similarly, we deduce that the total loss of contrasts for concentric spheres in R^3 is $69\alpha/19r$.

Proposition 2 (Unimodal function) Let $u_0(r)$ be defined in $[r_0, r_3]$. In addition, suppose that:

1) $u_0(r)$ is a step function with a single extremum, i.e., $U_1 \le U_2 \ge U_3$.
2) $U_2 + \delta_2 \ge U_i + \delta_i$ for $i = 1,3$.
3) $U_2 + \delta_2 \le \min_{r \in [r_1, r_2]} u_0(r)$

4) $U_i + \delta_i \ge \max_{r \in [r_{i-1}, r_i]} u_0(r)$ for $i = 1,3$.
5) $\max\{|u_0(r) - U|\} \ge |\delta_i|$ for any $1 \le i \le 3$.

Then the solution of (3) can be written as

$$u(r) = U_i + \delta_i \text{ for } r \in [r_{i-1}, r_i] \text{ with } 1 \le i \le 3.$$

where

$$\delta_i = \begin{cases} (-1)^{i-1} \dfrac{\alpha(|\partial\Omega_{r_{i-1}}| + |\partial\Omega_{r_i}|)}{|\Omega_{r_{i-1}, r_i}|}, & \text{if } \dfrac{|\alpha(|\partial\Omega_{r_{i-1}}| + |\partial\Omega_{r_i}|)|}{|\Omega_{r_{i-1}, r_i}|} \le \max\{|u_0(r) - U|\} \\ U - U_i, & \text{if } \dfrac{|\alpha(|\partial\Omega_{r_{i-1}}| + |\partial\Omega_{r_i}|)|}{|\Omega_{r_{i-1}, r_i}|} \ge \max\{|u_0(r) - U|\} \end{cases}$$

$$(5)$$

This proposition can be proved by dividing the unimodal function $u_0(r)$ into two types of component $u_0^1(r)$ and $u_0^2(r)$ with the splitpoint $\frac{r_1 + r_2}{2}$, where $u_0^1(r)$ and $u_0^2(r)$ are monotonically increasing and decreasing step functions, respectively, and then using the conclusion in Proposition 1.

Example 3. For the simplest case of R^1 unimodal function, we still assume that (1) $r_0 = 0$; (2) $|\partial\Omega_{r_0}| = 0$

and $|\partial\Omega_{r_3}| = 0$; and (3) $|r_0 r_1| = |r_1 r_2| = |r_2 r_3| = r$. In this case, the change in function intensity is given by $\delta_1 = \alpha/r$, $\delta_2 = -2\alpha/r$, and $\delta_3 = \alpha/r$; and the regularized image is represented as

$$u(r) = U_i + \delta_i = \begin{cases} U_1 + \dfrac{\alpha}{r}, & i = 1 \\ U_2 - \dfrac{2\alpha}{r}, & i = 2 \\ U_3 + \dfrac{\alpha}{r}, & i = 3 \end{cases}$$

We can clearly see that the contrasts in the restoration are less than they were in the true image, as illustrated in Fig. 2b. And the total loss of contrasts is $3\alpha/r$.

Example 4. We extend our results to the unimodal functions in R^2 and R^3. Under the same conditions as Example 3, by Proposition 2, we obtain that the total loss of contrasts is $4\alpha/r$ in R^2 and $36\alpha/7r$ in R^3.

Examples 1–4 indicate that in TV regularization, the loss of contrast for piecewise constant functions is exactly inversely proportional to scale of local feature measured by r (which explains why TV regularization can remove smaller scaled noise, while preserving larger scaled features essentially intact), is independent of original intensity, and is directly proportional to the regularization parameter α.
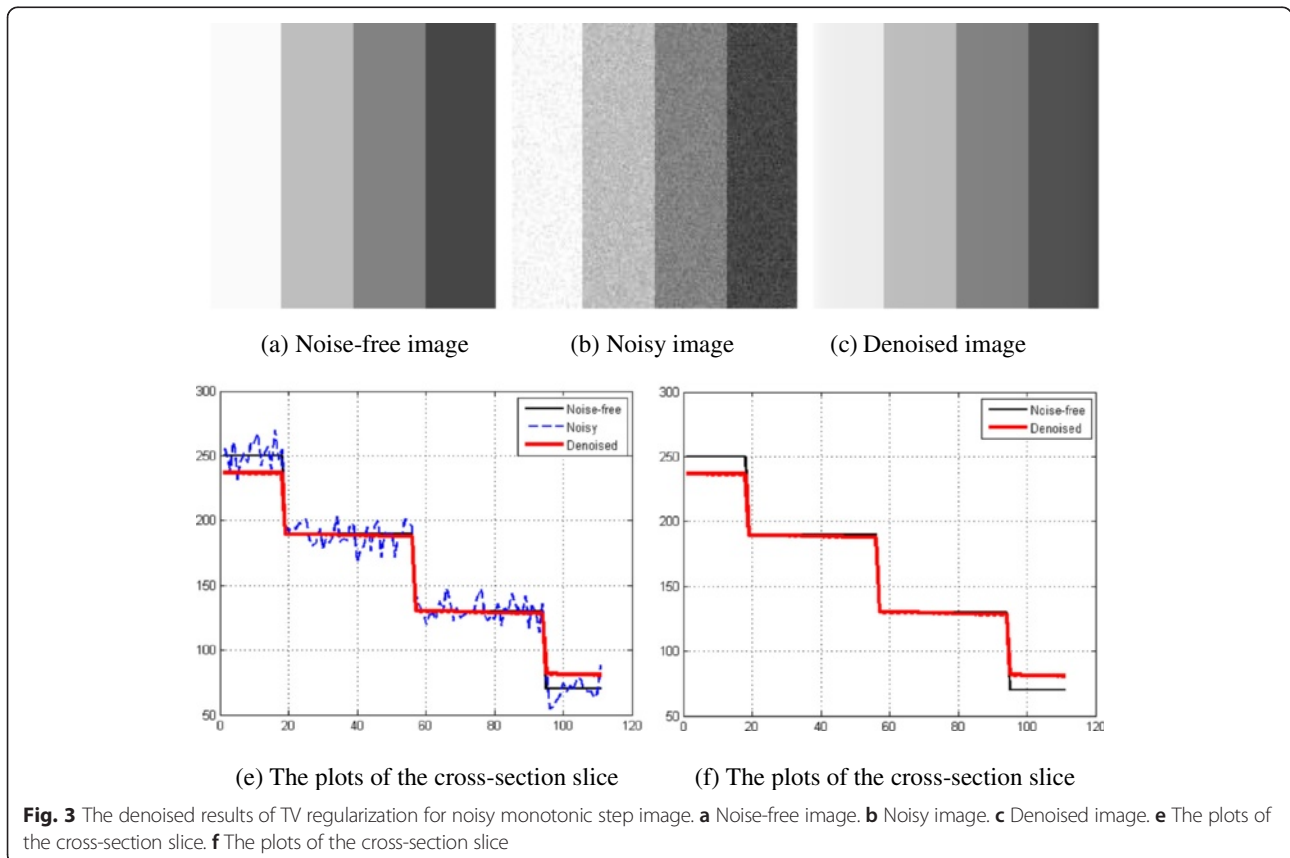

Fig. 3 The denoised results of TV regularization for noisy monotonic step image. **a** Noise-free image. **b** Noisy image. **c** Denoised image. **e** The plots of the cross-section slice. **f** The plots of the cross-section slice

To further show the loss of contrast in TV regularization, we use TV regulation model (3) to two synthetic images, one is a noisy monotonic step image, and another is a noisy step image with some extremums. The noisy images are obtained by adding the Gaussian noise with zero mean and 10 standard deviation to the corresponding clean versions. The results are show in the Figs. 3 and 4, respectively. From the results, it is obvious that the restorations (Figs. 3c and 4c) have a loss in contrast. The plots of the cross-section slice further illustrate this point. Overall, the agreement between the theory and the experiment results is compatible.

3 Forward-backward diffusion in the framework of TV (TV-FBD)

3.1 Backward diffusion model

Let $f \in L^2(\Omega)$, we construct an energy with respect to u.

$$\int_\Omega \varphi(|\nabla u|)dx + \frac{1}{2}\|f-u\|_2^2 \tag{6}$$

where $\|\cdot\|_2^2$ denotes the square of the L^2-norm; ∇ represents gradient operator; and potential function φ is defined in $[0, +\infty)$ and satisfies that:

(C.1) $\varphi(s) \geq 0$ for any $s \in [0, +\infty)$;

(C.2) $\varphi(s)$ is a monotony decreasing function in $[0, +\infty)$;

(C.3) $\varphi(0) = 1$ and $\lim_{s \to +\infty}\varphi(s) = 0$.

The corresponding gradient descent flow of energy (6) is

$$\frac{\partial u}{\partial t} = \text{div}\left(\frac{\varphi'(|\nabla u|)}{|\nabla u|}\nabla u\right) + (u-f) \tag{7}$$

In the Eq. (7), the term $\text{div}(\varphi'(|\nabla u|)/|\nabla u| \cdot \nabla u)$ is the diffusion term with the diffusion velocity $\varphi'(|\nabla u|)/|\nabla u|$. Because $\varphi(s)$ is a monotonically decreasing function in $[0, +\infty)$, we have $\varphi'(s) < 0$ in $[0, +\infty)$. With the nonnegativity of $|\nabla u|$, it is obvious that $\varphi'(|\nabla u|)/|\nabla u| < 0$. In this case, the gradient descent flow

$$\frac{\partial u}{\partial t} = \text{div}\left(\frac{\varphi'(|\nabla u|)}{|\nabla u|}\nabla u\right) \tag{8}$$

is actually a backward diffusion equation with a negative diffusion velocity. Next, we study the diffusion behavior in Eq.(8) in theory. Eq.(8) can be rewritten as

$$\frac{\partial u}{\partial t} = \frac{\varphi'(|\nabla u|)}{|\nabla u|}\Delta u + \nabla\left(\frac{\varphi'(|\nabla u|)}{|\nabla u|}\right) \cdot \nabla u \tag{9}$$

More precisely, Eq.(9) is a reaction diffusion equation, in which the term $\varphi'(|\nabla u|)/|\nabla u| \cdot \Delta u$ is the diffusion term, and $\nabla(\varphi'(|\nabla u|)/|\nabla u|) \cdot \nabla u$ is the reaction term.

(a) Noise-free image   (b) Noisy image   (c)Denoised image

(e) The plots of the cross-section slice   (f) The plots of the cross-section slice
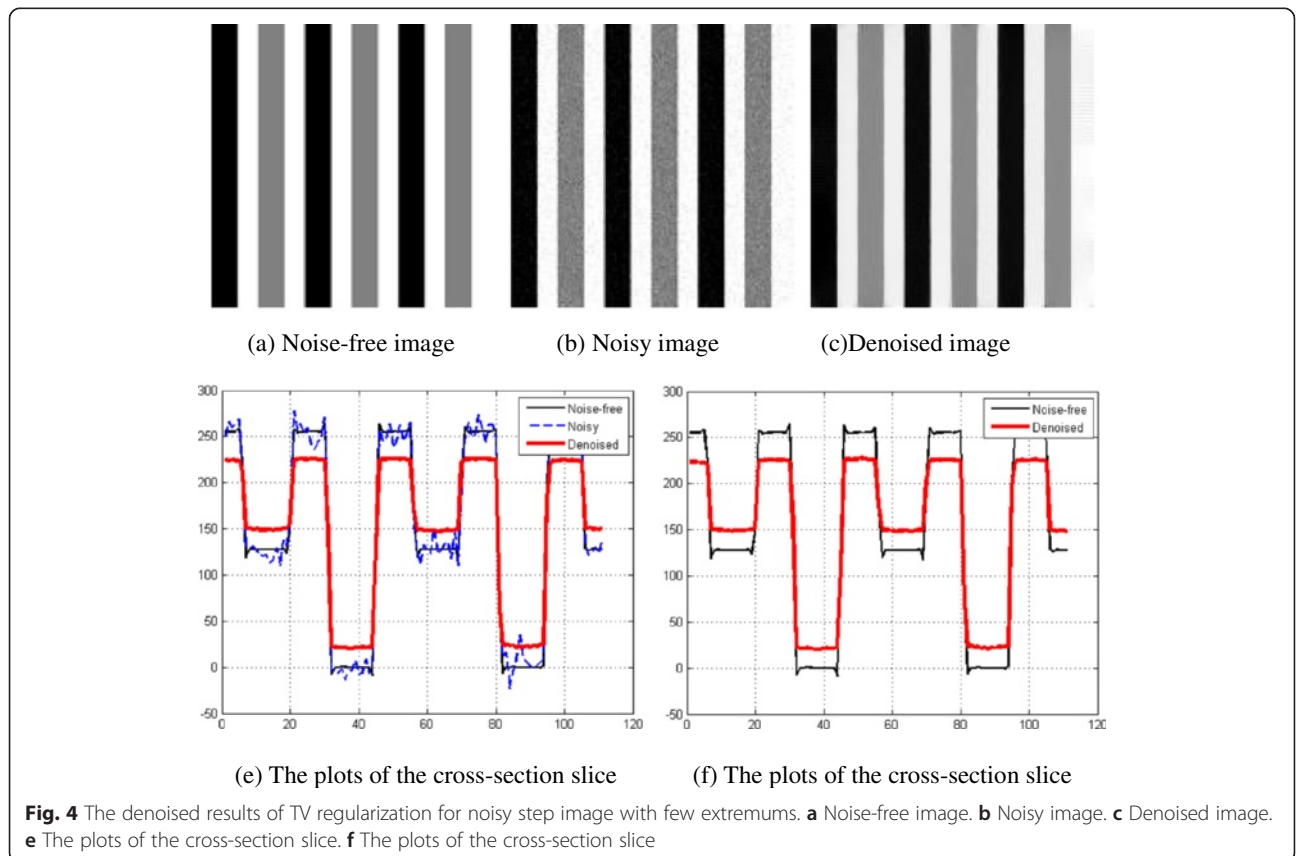
Fig. 4 The denoised results of TV regularization for noisy step image with few extremums. **a** Noise-free image. **b** Noisy image. **c** Denoised image. **e** The plots of the cross-section slice. **f** The plots of the cross-section slice

Then, studying the diffusion behavior of Eq.(8) is equivalent to study the diffusion term $\varphi'(|\nabla u|)/|\nabla u| \cdot \nabla u$ in Eq.(9), where Δ is Laplacian operator. To simplify the notation, we denote $\varphi'(|\nabla u|)/|\nabla u| = \psi(u)$. With the conclusion stated in above, we have

$$\Psi(u) < 0$$

We next study the diffusion behavior of $\partial u/\partial t = \varphi'(|\nabla u|)/|\nabla u| \cdot \Delta u = \Psi(u)\Delta u$ in Eq.(9) in the framework of difference. Using forward difference to approximate $\partial u/\partial t$, and successively using backward difference and forward difference once to approximate Laplacian, we obtain

$$u_s^{t+1} = u_s^t + \frac{\Psi\left(u_s^t\right)}{|N_s|} \sum_{p \in N_s} \nabla u_{s.p}^t \qquad (10)$$

where u_s^t represents value of function u at the "s" point after t times iterations; N_s represents the neighborhood of s point; $|N_s|$ is the cardinality of the neighborhood; and

$$\nabla u_{s.p}^t = u_p - u_s^t, \ p \in N s \qquad (11)$$

In combining Eq.(10) with Eq.(11), we obtain

$$u_s^{t+1} = u_s^t + \frac{\Psi\left(u_s^t\right)}{|N_s|} \sum_{p \in N_s} \nabla u_{s.p}^t = u_s^t + \frac{\Psi\left(u_s^t\right)}{|N_s|} \sum_{p \in N_s} \left(u_p - u_s^t\right)$$

From the last equation, we deduce that

$$u_s^{t+1} - u_s^t = \left|\Psi\left(u_s^t\right)\right| \left(u_s^t - \frac{1}{|N_s|} \sum_{p \in N_s} u_p\right) \qquad (12)$$

where $\frac{1}{|N_s|} \sum_{p \in N_s} u_p$ is the mean of u the neighborhood N_s.

We suppose that $u_s^0 > \frac{1}{|N_s|} \sum_{p \in N_s} u_p$, where u_s^0 is the initial value of u before iteration. By Eq.(12), we have

$$u_s^1 > u_s^0;$$

$$u_s^2 - u_s^1 = \left|\Psi\left(u_s^1\right)\right| \left(u_s^1 - \frac{1}{|N_s|} \sum_{p \in N_s} u_p\right)$$
$$> \left|\Psi\left(u_s^1\right)\right| \left(u_s^0 - \frac{1}{|N_s|} \sum_{p \in N_s} u_p\right) > 0$$

In successively using the last inequality, we deduce that

$$u_s^{t+1} > u_s^t$$

Similarly, if $u_s^0 < \frac{1}{|N_s|} \sum_{p \in N_s} u_p$, we have

$$u_s^{t+1} < u_s^t$$

From the above analysis, we can conclude that if the function value of a point is larger (smaller) than the mean of its neighborhood, its function value is getting larger (smaller) and larger (smaller) during the iteration, which may further increase the difference between itself and its neighbors. So, backward diffusion model can enhance the contrast and sharpen edges in image processing. Two examples of contrast enhancement by backward diffusion model can be observed in Fig. 5 for the "monotonic step" and "extrema" cases.

3.2 The selection of φ

With the conditions (C.1)–(C.3), there are several different choices for potential function ϕ. We here only consider two typical real-valued functions: one is an exponential function, and the other is a rational function, which are defined as following,

$$\varphi_1(s) = e^{-s}; \quad \varphi_2(s) = \frac{1}{1+s}$$

Figure 6a shows the plots of these two functions. It is obvious that both the functions meet the conditions (C.1)–(C.3), and also have the same variation trend in $[0, +\infty)$. A little difference between them is that $\varphi_1(s)$ has a faster descent velocity than $\varphi_2(s)$. In the backward diffusion equation, the diffusion velocities with potential function φ_1 and φ_2 are
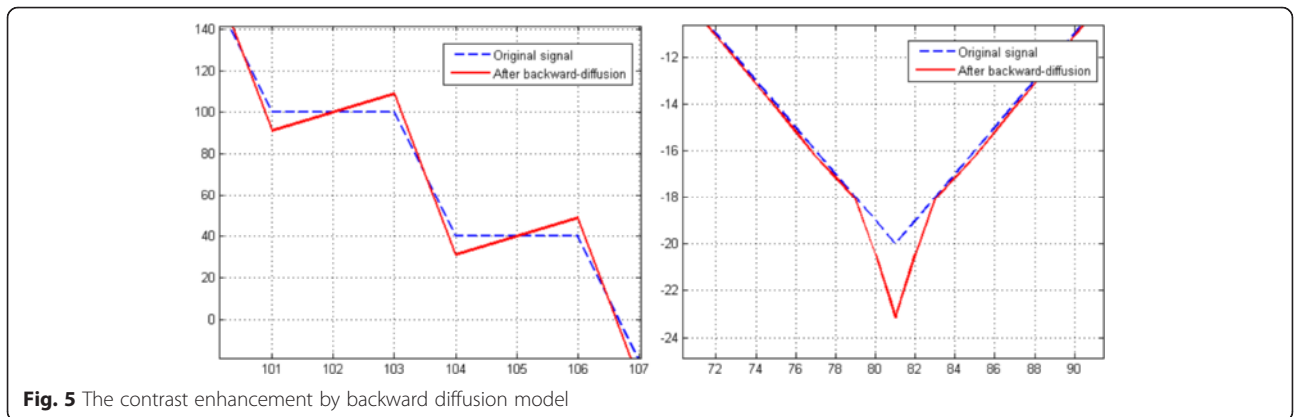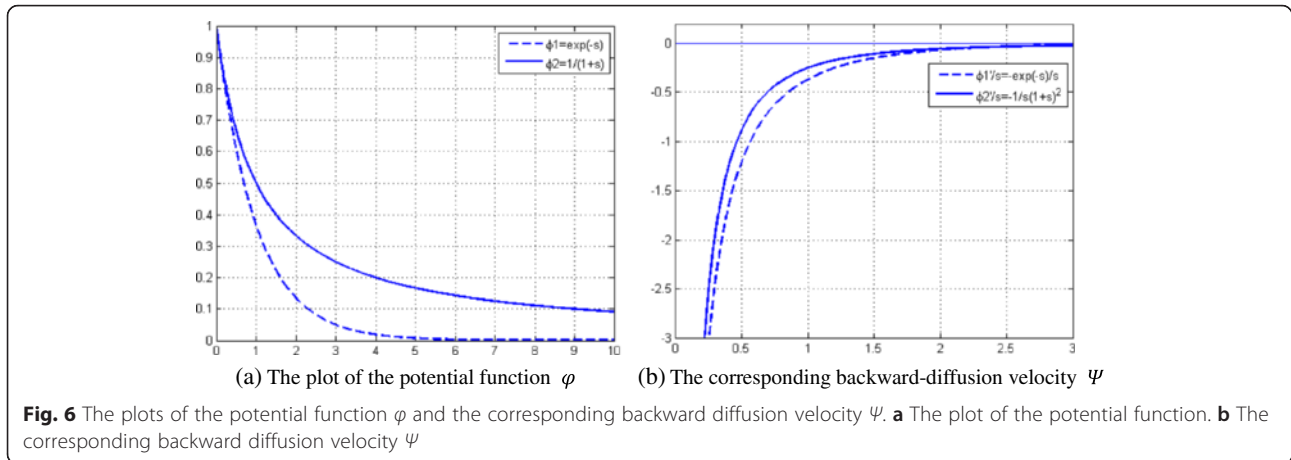

Fig. 5 The contrast enhancement by backward diffusion model

Fig. 6 The plots of the potential function φ and the corresponding backward diffusion velocity Ψ. **a** The plot of the potential function. **b** The corresponding backward diffusion velocity Ψ

$$\Psi_1(s) = \frac{\varphi_1'(s)}{s} = \frac{-e^{-s}}{s}$$

$$\Psi_2(s) = \frac{\varphi_2'(s)}{s} = \frac{-1}{s(1+s)^2}$$

, respectively. Figure 6b shows the plots of these two diffusion velocity functions Ψ_1 and Ψ_2. We have the following observations:

(1) Both plots are below the line of $s = 0$, which implies that $\Psi_1(s) < 0$ and $\Psi_2(s) < 0$ for any $s \in [0, +\infty)$. This further demonstrates that we can achieve the backward diffusion by using the above potential functions φ_1 and φ_2.

(2) The plots show that the diffusion velocity decreases with the increase of the value of s, and the smaller the value of s is, the faster the diffusion viscosity is. That is to say, in image diffusion, the area with small variation in intensity has a large backward diffusion velocity. Conversely, the area with large intensity variation has a small backward diffusion velocity. In this case, we can enhance the contrast in the regions with small intensity variation, and preserve the sharp intensity variation in the other regions.

(3) Ψ_1 and Ψ_2 have the same variation trend on $[0, +\infty)$, and providing nearly the same function values at the same point of s. So, in image diffusion, using these two potential function φ_1 and φ_2, we can obtain nearly the same backward diffusion velocity. In what follows, we set potential function as $\varphi(s) = e^{-s}$.

3.3 The TV-FBD model

In combining backward diffusion energy with TV energy, we have

$$\alpha \int_\Omega |\nabla u| dx + \beta \int_\Omega \varphi(|\nabla u|) dx + \frac{1}{2}\|f - u\|_2^2 \qquad (13)$$

where α and β are two nonnegative tuning parameters which balance the strength of backward diffusion and forward diffusion. The corresponding gradient descent flow of energy (13) is

$$\frac{\partial u}{\partial t} = \alpha \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) + \beta \mathrm{div}\left(\frac{\varphi'(|\nabla u|)}{|\nabla u|}\nabla u\right) + (u - f) \qquad (14)$$

Because the div is a linear operator, the Eq.(14) can be rewritten as

$$\frac{\partial u}{\partial t} = \mathrm{div}\left(\frac{\alpha + \beta\varphi'(|\nabla u|)}{|\nabla u|}\nabla u\right) + (u - f) \qquad (15)$$

Equation (15) is actually a reaction diffusion equation, in which $(\alpha + \beta\varphi'(|\nabla u|))/|\nabla u|$ is diffusion velocity. In fact, the term $\alpha/|\nabla u| > 0$ is forward diffusion velocity, which measures the ability of denoising for diffusion equation (15); and $\beta\varphi'(|\nabla u|)/|\nabla u| < 0$ is backward diffusion velocity, which controls the ability of contrast enhancing for diffusion equation (15). Here, the parameters α and β play the key role in our forward-backward diffusion to determinate magnitudes and directions of diffusion velocity. We next give a few observations.

First, we notice that if we take $\alpha = \beta$, then $\alpha/s > \beta|\varphi'(s)|/s$, i.e., the forward diffusion velocity is larger than the backward diffusion velocity. Figure 7a shows the plots of α/s and $\beta|\varphi'(s)|/s$ with $\alpha = \beta = 1$, which further demonstrates this point. So in this case, the forward-backward diffusion velocity satisfies that $(\alpha + \beta\varphi'(|\nabla u|))/|\nabla u| > 0$, which implies that the forward diffusion dominates the diffusion directions, and the forward-backward diffusion only turns into a pure forward diffusion (see Fig. 7b).

Second, we notice that if we take $\alpha < \beta$, then the forward-backward diffusion velocity $(\alpha + \beta\varphi'(|\nabla u|))/|\nabla u|$

is not always greater than zero, or always less than zero. Let s_0 be a zero point of $(\alpha + \beta\varphi'(s))/s$, i.e., $(\alpha + \beta\varphi'(s_0))/s_0 = 0$. It is obvious that

(1) when $|\nabla u| < s_0$, we have $(\alpha + \beta\varphi'(|\nabla u|))/|\nabla u| < 0$, Eq.(15) is a backward diffusion equation which can enhance the contrast in the regions where u satisfies $|\nabla u| < s_0$;

(2) when $|\nabla u| > s_0$, we have $(\alpha + \beta\varphi'(|\nabla u|))/|\nabla u| > 0$, Eq.(15) is a forward diffusion equation which can remove the noise from the regions where u satisfies $|\nabla u| > s_0$.

Figure 8 further demonstrates the above conclusions. In addition, from this figure, we deduce that the larger the difference between α and β is, the larger the value of s_0 is.

4 Numerical implementation

In this paper, we employ the two-step splitting (TSS) method to implement the proposed model (13), which allows for further fine tuning of strength of backward diffusion and forward diffusion. In addition, the TSS method splits the mixed diffusion model (13) into a backward diffusion and a forward diffusion, which enables us to seek the suitable fast algorithm for them, separately.

Firstly, we split the mixed diffusion model (13) into two sub-problems:

$$\min_{u}\left\{\alpha \int_{\Omega}|\nabla u|dx + \frac{1}{2}\|f-u\|_2^2\right\} \qquad (16)$$

and

$$\min_{u}\left\{\beta \int_{\Omega}\varphi(|\nabla u|)dx\right\} \qquad (17)$$
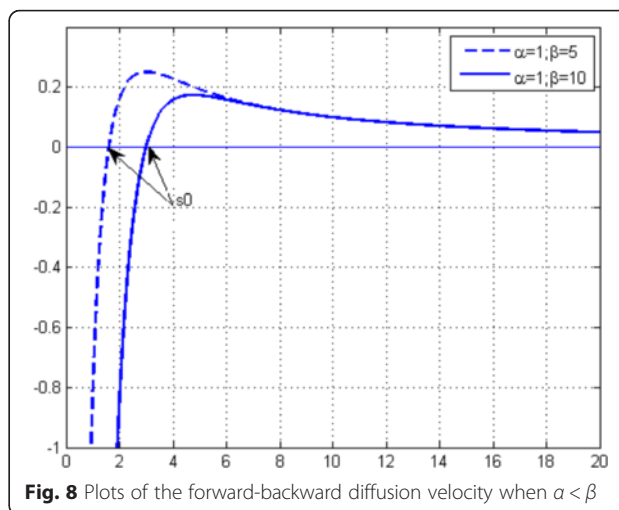
Then, the TSS method is stated as follows:

Fig. 8 Plots of the forward-backward diffusion velocity when $\alpha < \beta$

Step 1: Solve the forward diffusion term u with initial condition $u(x, t = 0) = u^n$ in problem (16) till some time T_f to obtain the intermediate solution, denoted by $u^{n+1/2} = u(x, T_f)$;

Step 2: Solve the backward diffusion term u with initial condition $u(x, t = 0) = u^{n+1/2}$ in problem (17) till some time T_b to obtain the final solution, denoted by $u^n = u(x, T_b)$.

The first step is to remove the noise from the observation, while leading to a reduction of contrast in intensity. And the second step is a correction for the first step, which can make up the losses in contrast. Because the backward diffusion is ill-posed, the second step may have the risk of moving the diffusion term u far away from observation f. By choosing a small enough T_b compared to the spatial resolution (i.e., the number of grid points), and iteratively using the first step to smooth the backward diffusion term u during the iteration process, the deviation will fall within acceptable limits. So, by using forward diffusion (step 1) and backward diffusion

(a) Forward and backward diffusion velocity   (b) Forward-backward-diffusion velocity
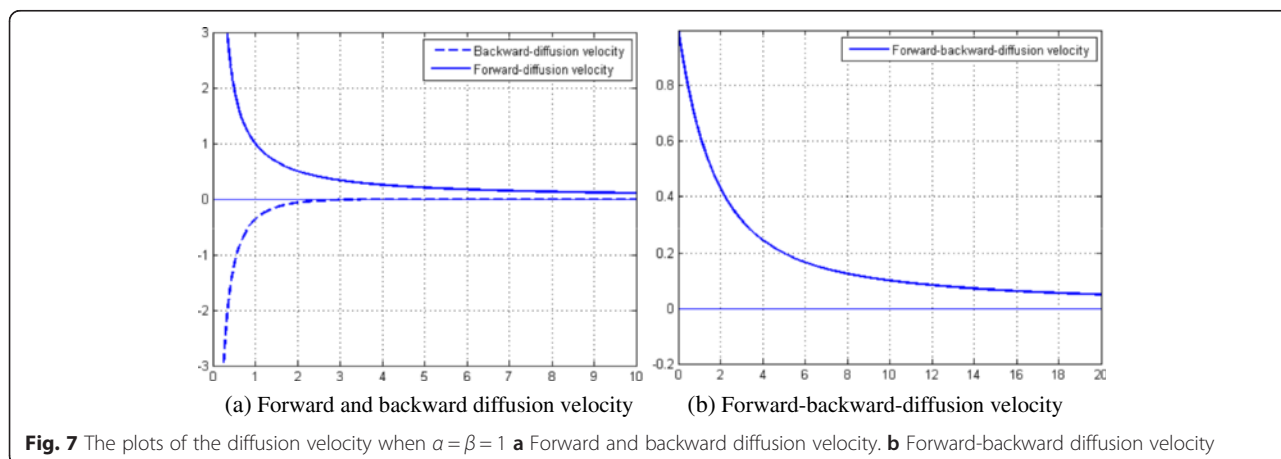
Fig. 7 The plots of the diffusion velocity when $\alpha = \beta = 1$ **a** Forward and backward diffusion velocity. **b** Forward-backward diffusion velocity

(step 2) alternatively, we can remove the noise, simultaneously preserve the contrast.

We here adopt the projection algorithm in the dual framework proposed by Chambolle [29] to solve minimization problem (16). The solution can be represented as:

$$u = f - \text{Proj}_{G_\alpha}(f)$$

where $\text{Proj}_{G_\alpha}(f)$ is the orthogonal projection of f on the closed convex set $G_\alpha = \{v : ||v||_G \leq \alpha\}$. In the discrete case, setting $\text{Proj}_{G_\alpha}(f) = \text{div}(\boldsymbol{g})$, the computation of this nonlinear projection amounts to solve the following constrained minimization problem with inequality constraints:

$$\min_{\boldsymbol{g} \in C_c^1(\Omega, R^2)} \left\{ ||\alpha \text{div}(\boldsymbol{g}) - f||_2^2, \; |\boldsymbol{g}_{i,j}| \leq 1, \; i = 1, 2, \cdots, M; j = 1, 2, \cdots, N \right\}$$

where $M \times N$ is the image size; $|\boldsymbol{g}_{i,j}| = \sqrt{g_1^2 + g_2^2}$ with $\boldsymbol{g} = (g_1, \; g_2)$. The necessary condition (Euler-Lagrange equation) of the minimization problem (27) getting an extremum is:

$$-\nabla(\alpha \text{div}(\boldsymbol{g}) - f)_{i,j} + \lambda_{i,j} \boldsymbol{g}_{i,j} = 0 \qquad (18)$$

where $\lambda_{i,j}$ is Lagrange multiplier. By the complementary slackness condition, we have

$$\lambda_{i,j} = \left| \nabla(\alpha \text{div}(\boldsymbol{g}) - f)_{i,j} \right|$$

Then, using semi-implicit fixed point iteration strategy to solve Eq.(18) with respect to $\boldsymbol{g}$, we obtain the following iteration scheme:

$$\boldsymbol{g}_{i,j}^0 = 0; \; \boldsymbol{g}_{i,j}^{n+1/2}$$
$$= \frac{\boldsymbol{g}_{i,j}^n + \Delta t_1 \left(\nabla \left(\text{div}(\boldsymbol{g}^n) - \frac{f}{\alpha} \right)_{i,j} \right)}{1 + \Delta t_1 \left| \left(\nabla \left(\text{div}(\boldsymbol{g}^n) - \frac{f}{\alpha} \right)_{i,j} \right) \right|} \qquad (19)$$

The forward diffusion term u is represented as

$$u^{n+1/2} = \alpha \text{div} \left(\boldsymbol{g}_{i,j}^{n+1/2} \right) \qquad (20)$$

By variation theory and gradient descent scheme, the minimum of (17) is the steady-state solution of the following PDE.

$$\begin{cases} \dfrac{\partial u}{\partial t} = \beta \text{div} \left(\dfrac{\varphi'(|\nabla u|)}{|\nabla u|} \nabla u \right) \\ u(\boldsymbol{x}, t = 0) = u_0(\boldsymbol{x}) \end{cases} \qquad (21)$$

In evolution Eq.(21), $|\nabla u|$ is in the denominator. In order to avoid the singularity, it is common to use a slightly perturbed norm $|\nabla u|_\varepsilon = \sqrt{|\nabla u|^2 + \varepsilon}$, where ε is

a small positive constant, to replace $|\nabla u|$. This is equivalent to minimize the functional

$$\beta \int_\Omega \varphi(|\nabla u|_\varepsilon) dx \qquad (22)$$

In [30], it is shown that the solutions of the perturbed problems (22) converge to the solution of (17) when $\varepsilon \to 0$. In our experiment, we set $\varepsilon = 10^{-5}$. In this case, by variation theory and gradient descent scheme, the minimum of (17) is the steady-state solution of the following PDE.

$$\begin{cases} \dfrac{\partial u}{\partial t} = \beta \text{div} \left(\dfrac{\varphi'(|\nabla u|_\varepsilon)}{|\nabla u|_\varepsilon} \nabla u \right) \\ u(\boldsymbol{x}, t = 0) = u_0(\boldsymbol{x}) \end{cases} \qquad (23)$$

Using finite difference method to solve Eq.(23) numerically, we obtain the following iteration scheme:

$$u_{i,j}^{n+1} = u_{i,j}^{n+1/2} + \Delta t_2 \left(\beta \text{div} \left(\frac{\varphi'\left(|\nabla u^{n+1/2}|_\varepsilon\right)}{|\nabla u^{n+1/2}|_\varepsilon} \nabla u^{n+1/2} \right) \right)_{i,j} \qquad (24)$$

where the discrete version of the gradient predator $(\nabla u)_{i,j} = ((\partial_x u)_{i,j}, (\partial_y u)_{i,j})$ is computed by:

$$(\partial_x u)_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j}, & j < N \\ 0, & j = N \end{cases} \text{ and } (\partial_y u)_{i,j}$$
$$= \begin{cases} u_{i+1,j} - u_{i,j}, & i < M \\ 0, & i = M \end{cases}$$

And the discrete version of the divergence predator $\text{div}(\xi^1, \xi^2)_{i,j}$ is computed by:

$$\text{div}(\xi^1, \xi^2)_{i,j} = \begin{cases} \xi_{i,j}^1 - \xi_{i,j-1}^1, & 1 < j < N \\ \xi_{i,j}^1, & j = 1 \\ -\xi_{i,j-1}^1, & j = N \end{cases}$$
$$+ \begin{cases} \xi_{i,j}^2 - \xi_{i-1,j}^2, & 1 < i < M \\ \xi_{i,j}^2, & i = 1 \\ -\xi_{i-1,j}^2, & i = M \end{cases}$$

In the first step, we smooth the observation by forward diffusion Eqs. (19)–(20). While in the second step, we enhance the contrast in the smoothed version obtained from the previous step by backward diffusion Eq. (24). Then, the enhanced version obtained from the second step is used as the next input of the forward diffusion step Eqs. (19)–(20). We proceed with successive application of the above alternate steps, and obtain the alternate iteration algorithm for our TV-FBD model (see Algorithm 1).

Algorithm 1: alternate iteration algorithm for FBD model

---

1. Initialize $u^0 = f$, $n = 0$.

2. Compute $u^{n+1/2}$

$$g_{i,j}^0 = 0; \quad g_{i,j}^{n+1/2} = \frac{g_{i,j}^n + \Delta t_1 \left(\nabla \left(\mathrm{div}(g^n) - \frac{u^n}{\alpha} \right)_{i,j} \right)}{1 + \Delta t_1 \left| \left(\nabla \left(\mathrm{div}(g^n) - \frac{u^n}{\alpha} \right)_{i,j} \right) \right|}$$

$$u^{n+1/2} = \alpha \mathrm{div}\left(g_{i,j}^{n+1/2}\right)$$

3. Compute u^{n+1}

$$u^{n+1} = u_{i,j}^{n+1/2} + \Delta t_2 \left(\beta \mathrm{div} \left(\frac{\varphi'\left(\left| \nabla u^{n+1/2} \right|_\varepsilon \right)}{\left| \nabla u^{n+1/2} \right|_\varepsilon} \nabla u^{n+1/2} \right) \right)_{i,j}$$

4. If u^{n+1} satisfies stationary condition, stop; otherwise, set $n = n + 1$ and return to Step 2.

---

5 Experimental results

In this section, we show the experimental results of image denoising on several synthetic and real images. The comparisons with TV model [12], TSM model [20], FBD model [23], and TV-FF model [26] are also performed in a forthcoming paper to show the superiority of the proposed model. The reason why we choose these models to compare in that the TV model [12] is the most original variational model in image denoising, which is the source of our study in this paper, and the other three models are representations of the three major classes of contrast preserving in TV regularization, respectively.

The algorithms are implemented by MATLAB software on a PC with an Intel Core (I5), CPU (2.50 GHz) and RAM (2.00 GB). We give also the peak signal-to-noise ratio (PSNR) and mean structural similarity (MSSIM) index [31] for quantitative analysis, which is defined as

$$\mathrm{PSNR}(u, u_0) = 10 \cdot \log_{10}\left(\frac{255^2}{\mathrm{MSE}} \right) \quad \text{with } \mathrm{MSE} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} (u-u_0)^2}{M \times N}$$

and

$$\mathrm{MSSIM}(u, u_0) = \frac{1}{\mathrm{M}} \sum_{i=1}^{M} \mathrm{SSIM}\left(u^i, u_0^i \right)$$

with

$$\mathrm{SSIM}\left(u^i, u_0^i \right) = \frac{\left(2\mu_{u^i}\mu_{u_0^i} + C_1 \right)\left(2\sigma_{u^i \, u_0^i} + C_2 \right)}{\left(\mu_{u^i}^2 + \mu_{u_0^i}^2 + C_1 \right)\left(\sigma_{u^i}^2 + \sigma_{u_0^i}^2 + C_2 \right)},$$

respectively. Here, u is the restored image from the observation u_0, μ_{u^i} and σ_{u^i} are the mean and the standard deviation of u at the i-th local image window, respectively; $\sigma_{u^i \, u_0^i}$ is the covariance between u^i and u_0^i; M is the number of local windows in the image; C_1 and C_2 are two constants. In all experiments, we set $C_1 = (0.01 * 255)^2$ and $C_2 = (0.03 * 255)^2$.

5.1 The choice of parameters

The parameter α is to adjust the degree of smoothing. If it is too small, the model cannot effectively remove the noises. Conversely, if α is too large, a large amount of image details will be erased due to the oversmoothing of

(a) Noise-free image   (b) Noisy image   (c) Denoised image

(d) The plots of the cross-section slice   (e) The plots of the cross-section slice

Fig. 9 The denoised results of our model for noisy step images shown in Figs. 3 and 4. **a** Noise-free image. **b** Noisy image. **c** Denoised image. **d** The plots of the cross-section slice. **e** The plots of the cross-section slice
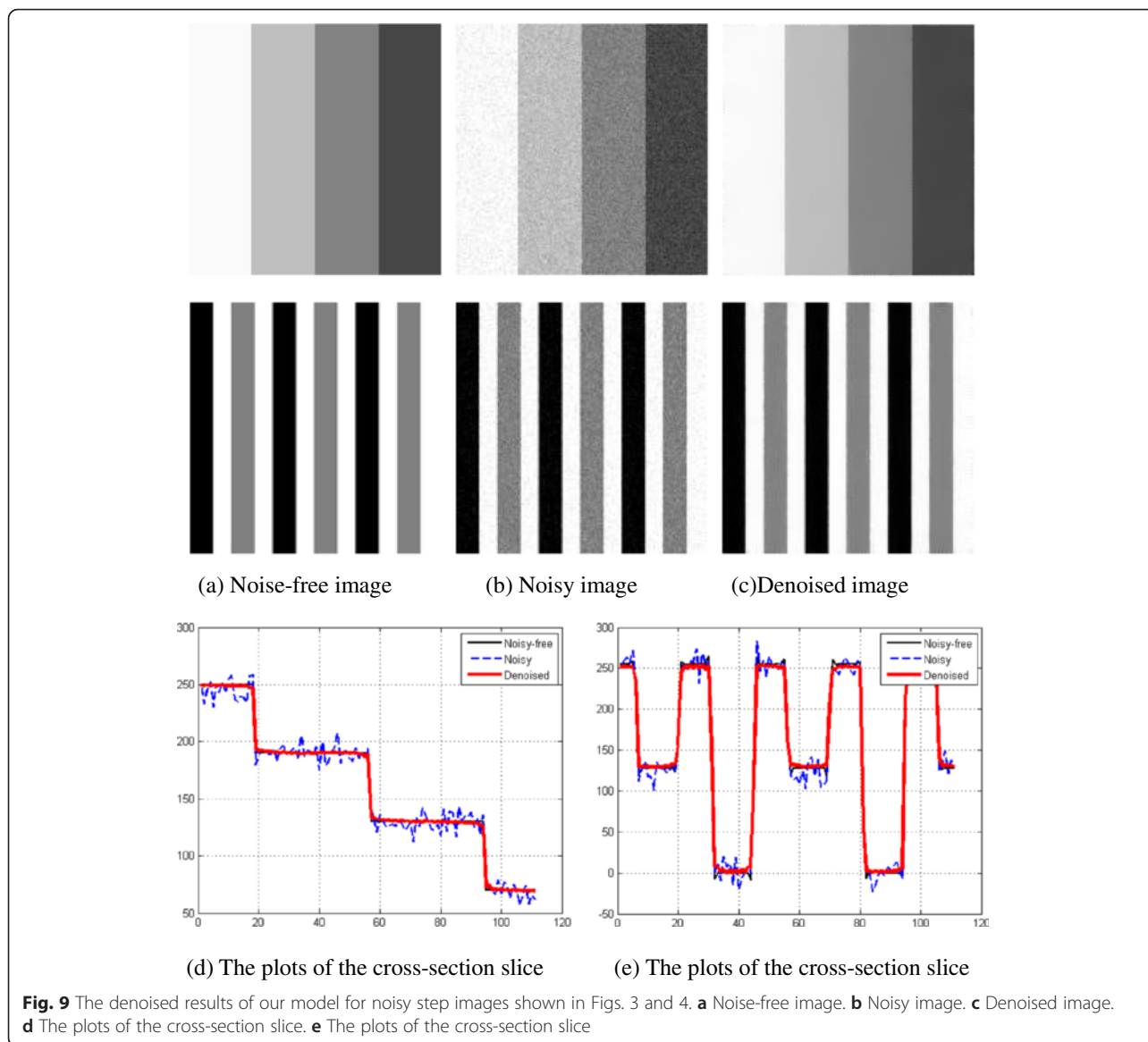
the image. How to choose an optimal smoothing parameter α in TV regularization is still an "open question". In this paper, we use the "trial and error" technique to determine the value of smoothing parameter. The parameter β controls the velocity of backward diffusion, as discussed in Section 3.3, which may be larger than α to achieve forward and backward diffusion. In all experiments, we set $\beta = 5\alpha$.

In [29], the authors introduced a sufficient condition to ensure the convergence of the iterative formula (13): if $\Delta t_1 \leq 1/8$, then $\alpha \mathrm{div}(\boldsymbol{g}^n) \to \mathrm{Proj}_{G_\alpha}(f)$ as $n \to +\infty$. With consideration of the numerical stability, convergence, and diffusion efficiency, we set $\Delta t_1 = 0.12$ in all experiments. Actually, the time step Δt_2 also controls the

backward diffusion in Step 2 as the weighting parameter β, and a large Δt_2 has the risk of increasing the range of image intensity excessively, and resulting in intensity distortion. Therefore, we should use a small enough Δt_2 compared to the time step Δt_1 of the forward diffusion so that the range of image intensity is within acceptable limits. In our experiments, we set $\Delta t_2 = 0.01$.

5.2 Test of synthetic images

In our first experiment, we demonstrate contrast preserving in denoising application for different synthetic images. The test noisy images are obtained by adding the Gaussian noise with standard deviation $\sigma = 10$ into the clean versions.
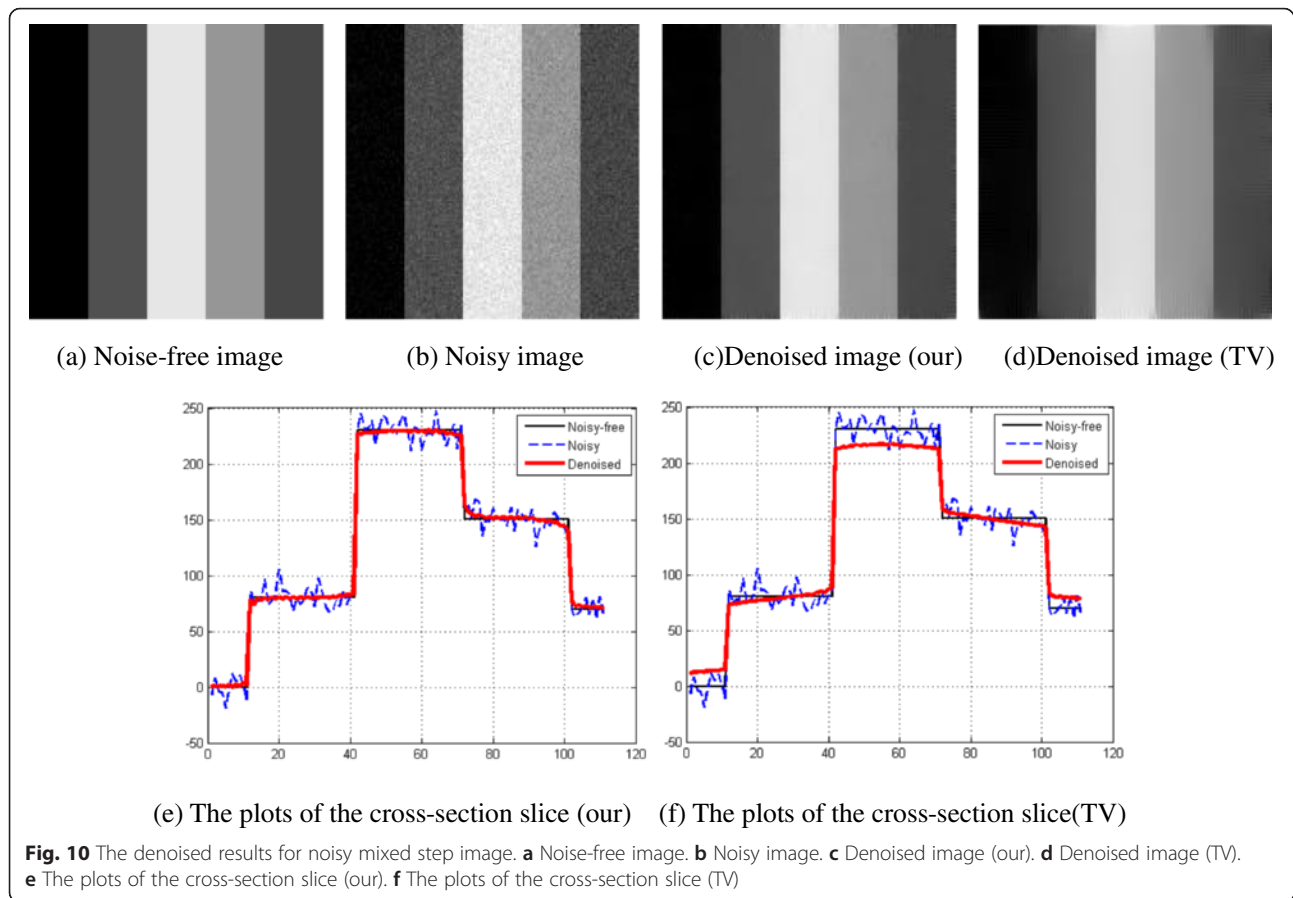
(a) Noise-free image (b) Noisy image (c)Denoised image (our) (d)Denoised image (TV)

(e) The plots of the cross-section slice (our) (f) The plots of the cross-section slice(TV)

Fig. 10 The denoised results for noisy mixed step image. **a** Noise-free image. **b** Noisy image. **c** Denoised image (our). **d** Denoised image (TV). **e** The plots of the cross-section slice (our). **f** The plots of the cross-section slice (TV)

Figure 9 shows the denoised results of our model for two noisy step images shown in Figs. 3 and 4. The first row shows the denoised results for the monotonic steps function shown in Fig. 3, the corresponding plots of the cross-section slice are shown in Fig. 9d. The second row shows the denoised results for the extremal steps function shown in Fig. 4, and the corresponding plots of the cross-section slice are shown in Fig. 9e. Here, compared with the noise-free images, we observe that there is no significant loss of contrast in the denoised images. We can see this much more clearly by the plots of the cross-section slice.

Figure 10 shows the denoised results of our model and TV regularization for a mixed step image comprised of monotonic steps and extremal steps. And Fig. 11 shows the denoised results of our model and TV regularization for a piecewise smooth image that can be seen as a "limit" of the piecewise constant. We can clearly see that in the both case, our model and TV regularization can successfully remove the noise and prevent the edges. But TV regularization obviously reduces the contrast in the denoised images, mainly in extremum regions, and starting and ending

of the monotonic steps. Due to the case that the backward diffusion term is incorporated into the energy, our model can compensate the loss of contrast caused by TV regularization.

5.3 Test of real images

In our second experiment, we evaluate the performance of the proposed model using the Barbara image contaminated by Gaussian noise with different standard deviations. In addition, we compare our model to TV, FBD, TSM, and TV-FF. We note that the parameters of each model are optimized to achieve the best restoration with respect to the PSNR.

Figure 12 shows the denoised results for noisy Barbara obtained by adding Gaussian noise with standard deviation $\sigma = 5$ to the clean one. Figure 12a shows the noise-free Barbara image; Fig. 12b shows the corresponding noisy version. The black solid line in the noisy Barbara represents the cross-section slice that will be plotted in the following experiments to show how noise is removed and how contrast changes. Figure 12c–g shows the denoised results using the five models, and Fig. 12h–l
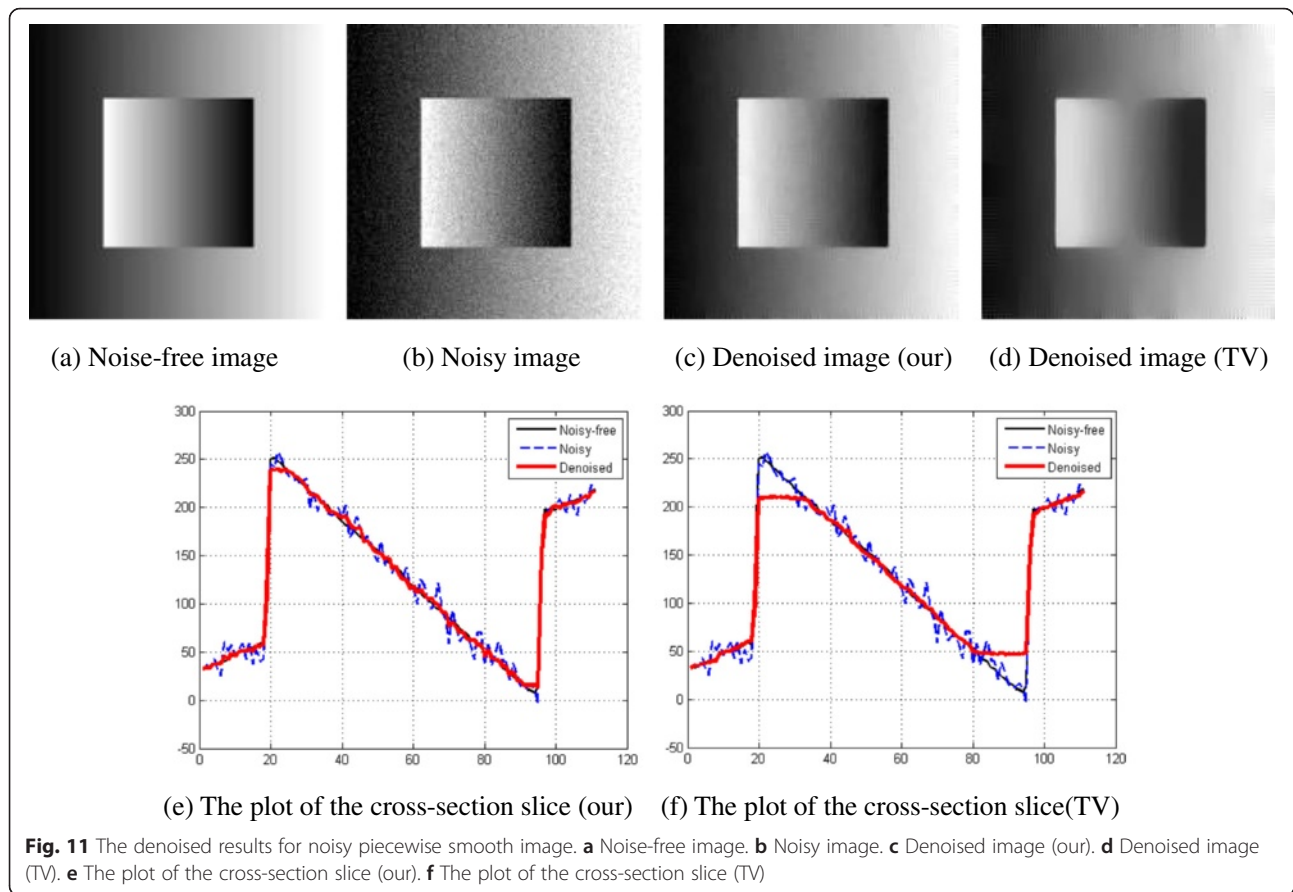
(a) Noise-free image   (b) Noisy image   (c) Denoised image (our)   (d) Denoised image (TV)

(e) The plot of the cross-section slice (our)   (f) The plot of the cross-section slice(TV)

Fig. 11 The denoised results for noisy piecewise smooth image. **a** Noise-free image. **b** Noisy image. **c** Denoised image (our). **d** Denoised image (TV). **e** The plot of the cross-section slice (our). **f** The plot of the cross-section slice (TV)

shows the corresponding local zoom-in of the plots of the cross-section slice (range is of horizontal axis $\in [140, 160]$; intensity $\in [100, 200]$). Figure 13 shows denoised results for noisy Barbara contaminated by Gaussian noise with standard deviation $\sigma = 10$.

From the results shown in Figs. 12 and 13, we see that the five models can be ranked according to the restoration quality: TSM < TV < TV-FF < FBD < TV-FBD. The TSM model performs worst since it applies histogram equalization to enhance the restoration obtained by TV regularization in the previous step. Histogram equalization can perform well in contrast enhancement and improve the visual quality. But it has the drawback of increasing the range of intensity, and resulting in visual distortion. So, the restoration obtained by TSM may be worse than that obtained by TV. The TV-FF model performs better but also has some loss of contrast near steps. The FBD model performs significantly better since it adopts a linear backward diffusion in the diffusion PDE, which can compensate the loss of contrast caused by forward diffusion. The proposed TV-FBD model performs best, which leads to a best approximation of the original data with appropriate contrast near steps. The reason why the TV-FBD model performs better than FBD model is that the TV-FBD model adopts a nonlinear backward diffusion, which allows for further fine tuning of velocity in backward diffusion (see Fig. 6). And the FBD model adopts linear backward diffusion which has a constant velocity at any position. The quantitative evaluation of the restoration emphasizes the visual impression of the results. See Table 1 for the exact PSNR and MSSIM values.

Next, we test the denoising capabilities of the five models in case of severe noise. Figure 14 shows the denoised results for noisy Barbara containing Gaussian noise of standard deviation $\sigma = 15$; and Fig. 15 shows results for noisy Barbara containing Gaussian noise of standard deviation $\sigma = 20$. It is obvious that these five models can remove noise very well and preserve the sharp edges in the restorations. However, TV regularization leads to a darker result due to the loss of contrast. Due to the use of the contrast enhancement scheme, the other four models lead to significantly better results. Qualitatively, these models perform equally well, but the quantitative evaluation shows that the proposed TV-FBD model has higher PSNR and MSSIM values (see Table 1).

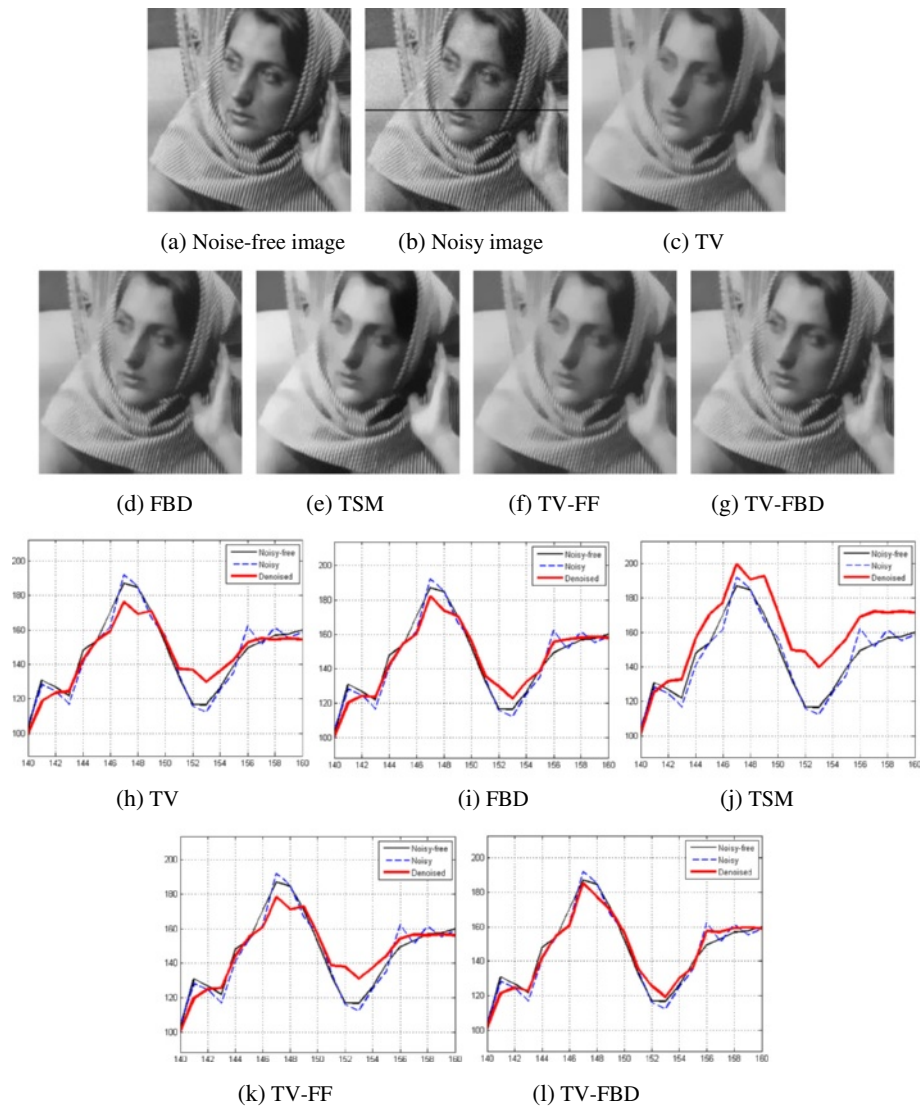Finally, to further show the effectiveness and adaptability of the proposed model, we apply it to denoise

Fig. 12 The denoised results of noisy Barbara ($\sigma = 5$). **a** Noise-free image. **b** Noisy image. **c** TV. **d** FBD. **e** TSM. **f** TV-FF. **g** TV-FBD. **h** TV. **i** FBD. **j** TSM. **k** TV-FF. **l** TV-FBD

four images of size 128×128 contaminated by Gaussian noise with different standard deviations ($\sigma = 5,10,15$, and 20). The first two test images are relatively simple. One is a panda image that contains large white areas and black areas (see Fig. 16a); the second is a camera-man in front of a blurry background, which contains large black areas and gray areas (see Fig. 17a). The other two are butterfly and Lenna images, respectively (see Figs. 18a and 19a, respectively). Compared to the first two images, these two are more complex, which contain a large amount of details, textures, and features of low contrast. Again, we compare our model to TV, FBD, TSM, and TV-FF.

We here only show the restoration results of the im-ages containing Gaussian noise with standard devia-tions $\sigma = 10$ (see Figs. 16, 17, 18, 19). For the other cases, we only give the PSNR and MSSIM values in Table 2. From these figures, we can clearly see that all models can successfully remove noise and simultan-eously preserve edges. But the quantitative evaluation shows that the proposed TV-FBD model has the highest PSNR and MSSIM values (see Table 2), which further demonstrates that our model has the best performance in these five models.

From the above denoised results, we note that forward-backward diffusion is a good tool for noise
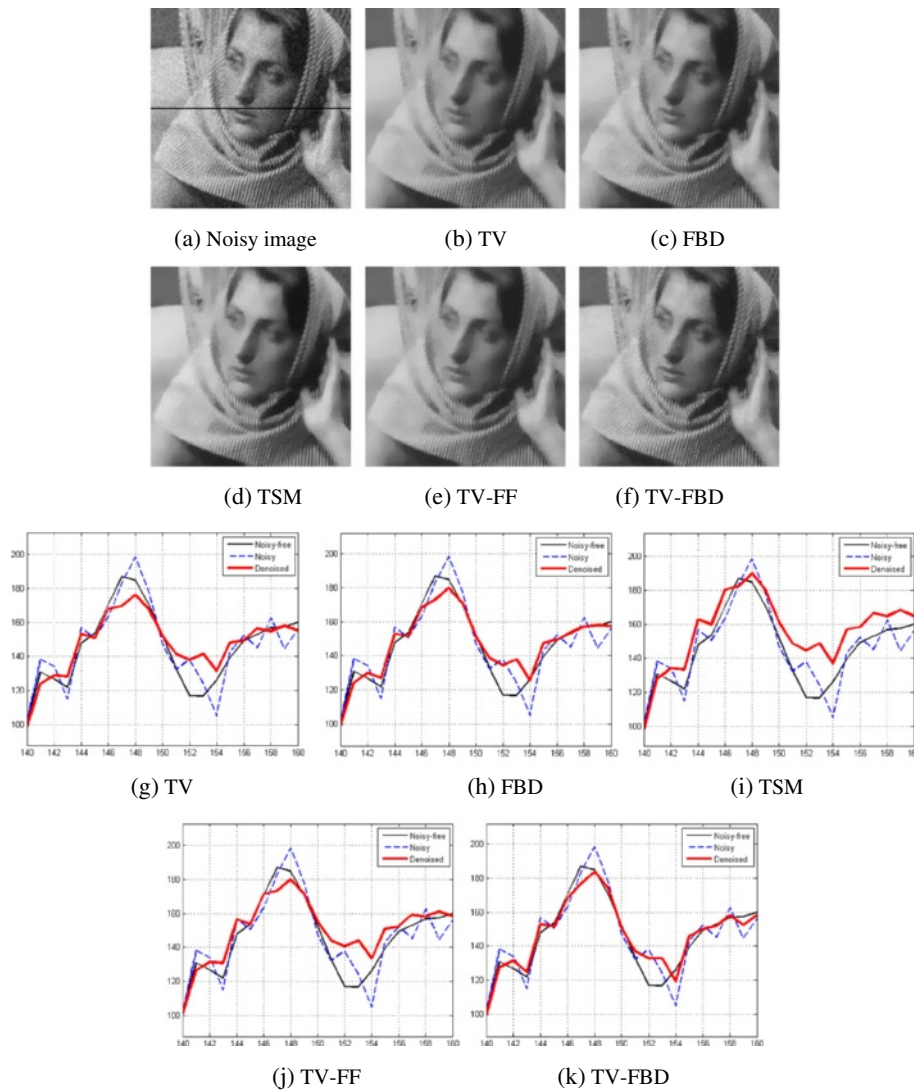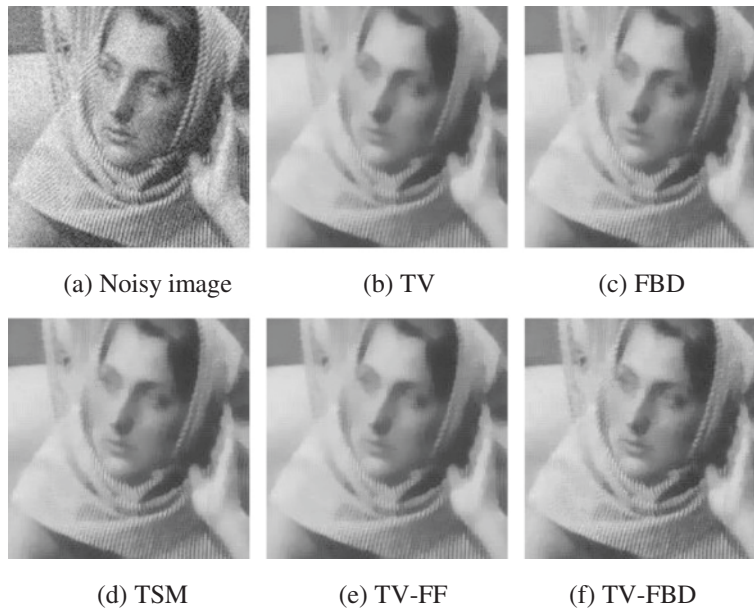
Fig. 13 The denoised results of noisy Barbara ($\sigma = 10$). **a** Noisy image. **b** TV. **c** FBD. **d** TSM. **e** TV-FF. **f** TV-FBD. **g** TV. **h** FBD. **i** TSM. **j** TV-FF. **k** TV-FBD

removing and contrast preserving. Table 2 shows that the traditional FBD model and TV-FBD have the best performance in terms of PSNR and MSSIM indexes among the five models. We here compare the efficiency of the proposed TV- FBD to the traditional FBD. The

CPU time for FBD and TV- FBD models are listed in Table 3, where the CPU time is the cost time that the restoration takes from the initiation to first achieve the local maximum of PSNR. One can clearly see that our model performs faster than traditional FBD since it

Table 1 PSNR and MSSIM for different models

Noise level	TV		FBD		TSM		TV-FF		TV-FBD	
	PSNR	MSSIM	PSNR	MSSIM	PSNR	MSSIM	PSNR	MSSIM	PSNR	MSSIM
σ=5	28.0192	0.9135	29.5753	0.9207	27.0321	0.8921	28.0813	0.9147	31.8464	0.9824
σ=10	27.7714	0.8963	28.9002	0.9247	26.5600	0.8546	27.8249	0.9045	29.9547	0.9628
σ=15	25.9548	0.8387	26.5473	0.8912	25.6347	0.8364	25.9754	0.8425	27.1542	0.8847
σ=20	25.4821	0.8242	25.7425	0.8298	23.9268	0.7845	25.4967	0.8256	26.1264	0.8796

(a) Noisy image   (b) TV   (c) FBD

(d) TSM   (e) TV-FF   (f) TV-FBD

Fig. 14 The denoised results of noisy Barbara ($\sigma = 15$). **a** Noisy image. **b** TV. **c** FBD. **d** TSM. **e** TV-FF. **f** TV-FBD
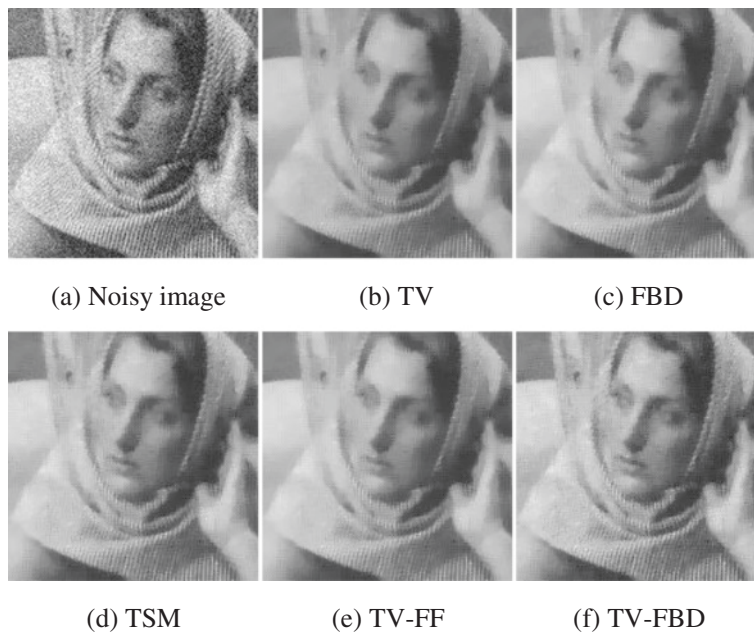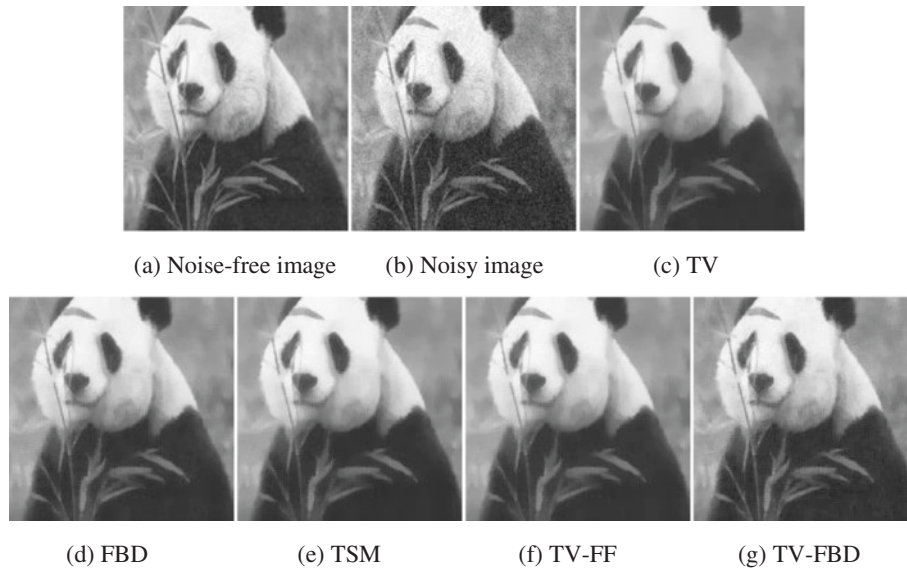

(a) Noisy image   (b) TV   (c) FBD

(d) TSM   (e) TV-FF   (f) TV-FBD

Fig. 15 The denoised results of noisy Barbara ($\sigma = 20$).**a** Noisy image. **b** TV. **c** FBD. **d** TSM. **e** TV-FF. **f** TV-FBD

(a) Noise-free image   (b) Noisy image   (c) TV

(d) FBD   (e) TSM   (f) TV-FF   (g) TV-FBD

Fig. 16 The denoised results of noisy panda ($\sigma = 10$). **a** Noise-free image. **b** Noisy image. **c** TV. **d** FBD. **e** TSM. **f** TV-FF. **g** TV-FBD
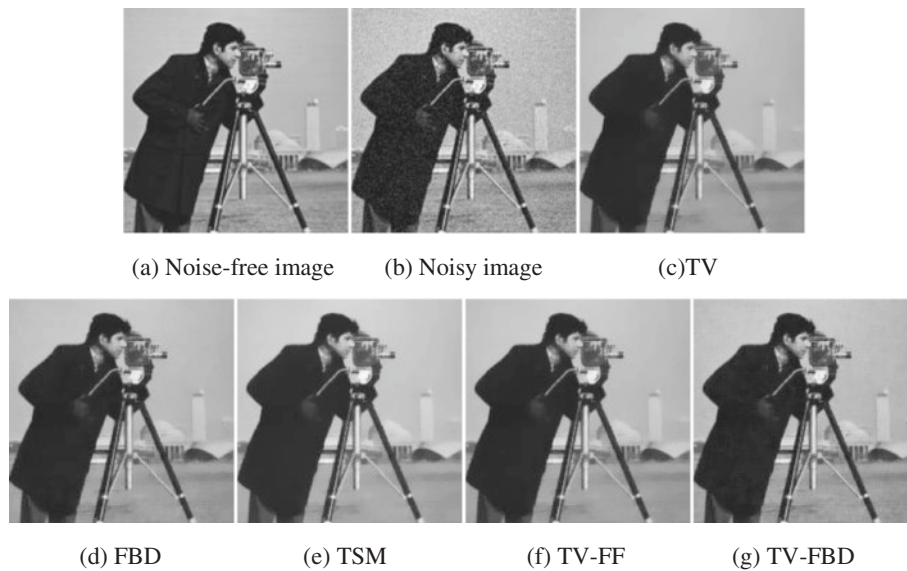

(a) Noise-free image   (b) Noisy image   (c)TV

(d) FBD   (e) TSM   (f) TV-FF   (g) TV-FBD

Fig. 17 The denoised results of noisy cameraman ($\sigma = 10$). **a** Noise-free image. **b** Noisy image. **c** TV. **d** FBD. **e** TSM. **f** TV-FF. **g** TV-FBD
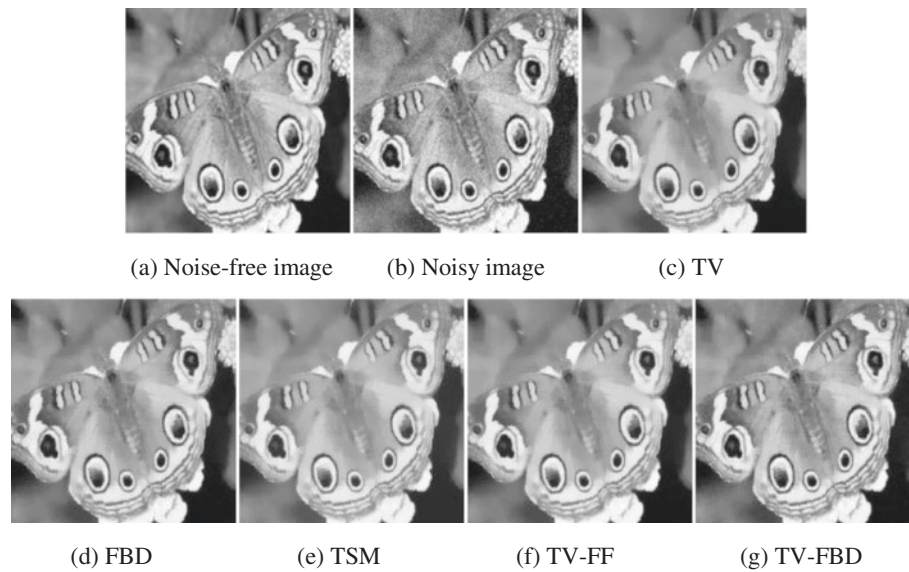
Fig. 18 The denoised results of noisy butterfly ($\sigma = 10$). **a** Noise-free image. **b** Noisy image. **c** TV. **d** FBD. **e** TSM. **f** TV-FF. **g** TV-FBD

adopt the more efficient dual projection algorithm rather than the finite differences scheme.

6 Conclusions
In this paper, we proposed a forward-backward diffusion model in the framework of total variation (i.e., TV-FBD), which can effectively solve the problem of the contrast loss in TV regularization. New model was obtained by introducing a nonconvex and monotony

decrease function with respect to total variation into the TV energy. A two-step splitting method was then proposed to effectively solve the TV-FBD model. We adopted the efficient projection algorithm in the dual framework to solve the forward diffusion in the first step, and then employed the simple finite differences scheme to solve the backward diffusion to compensate the loss of contrast occurred in the previous step. The experiments in both synthetic and real images demonstrated
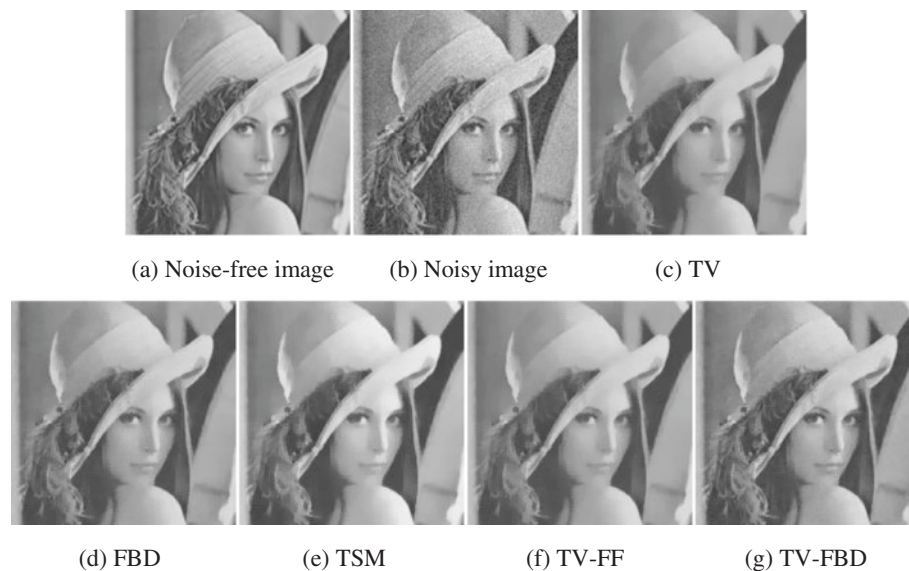

Fig. 19 The denoised results of noisy Lenna ($\sigma = 10$). **a** Noise-free image. **b** Noisy image. **c** TV. **d** FBD. **e** TSM. **f** TV-FF. **g** TV-FBD

Table 2 PSNR and MSSIM for different models

Image	TV		FBD		TSM		TV-FF		TV-FBD	
	PSNR	MSSIM	PSNR	MSSIM	PSNR	MSSIM	PSNR	MSSIM	PSNR	MSSIM
Panda ($\sigma = 5$)	31.5289	0.9887	32.8156	0.9915	29.7899	0.9575	31.4687	0.9827	34.3858	0.9996
Panda ($\sigma = 10$)	29.9145	0.9541	30.7611	0.9786	29.9093	0.9624	29.7495	0.9526	31.2835	0.9879
Panda ($\sigma = 15$)	28.0571	0.9127	28.4116	0.9247	26.0272	0.8547	28.6672	0.9356	28.8940	0.9245
Panda ($\sigma = 20$)	26.4142	0.8684	27.0674	0.8914	20.9034	0.6013	26.1909	0.8456	27.4173	0.8975
Cameraman ($\sigma = 5$)	31.7429	0.9876	33.2583	0.9990	29.4312	0.9489	31.6916	0.9915	35.1524	0.9998
Cameraman ($\sigma = 10$)	29.9190	0.9568	30.9482	0.9875	25.6738	0.8240	29.7772	0.9524	31.6007	0.9904
Cameraman ($\sigma = 15$)	28.8345	0.9247	28.8403	0.9287	24.7083	0.8004	28.6645	0.9147	28.8666	0.9314
Cameraman ($\sigma = 20$)	26.3152	0.8645	27.1729	0.9005	22.7043	0.7524	26.1041	0.8432	27.5059	0.9010
Butterfly ($\sigma = 5$)	29.5958	0.9527	31.2972	0.9916	25.2441	0.8225	29.7328	0.9489	33.3647	0.9991
Butterfly ($\sigma = 10$)	27.4500	0.9046	28.7401	0.9230	24.8330	0.8123	27.5483	0.9102	29.9052	0.9468
Butterfly ($\sigma = 15$)	26.1095	0.8603	27.0424	0.8996	21.4993	0.7056	27.0781	0.9014	27.5441	0.9032
Butterfly ($\sigma = 20$)	24.9835	0.8113	25.1705	0.8245	21.4317	0.6957	25.0673	0.8241	25.7847	0.8498
Lenna ($\sigma = 5$)	29.9093	0.9648	31.4282	0.9935	23.4201	0.7658	29.8653	0.9426	33.4752	0.9992
Lenna ($\sigma = 10$)	29.4348	0.9450	30.4161	0.9878	24.2896	0.8114	29.3260	0.9437	31.1194	0.9840
Lenna ($\sigma = 15$)	27.4370	0.9083	27.9907	0.9116	25.8678	0.8339	27.4010	0.9016	28.5369	0.9402
Lenna ($\sigma = 20$)	25.1589	0.8245	26.6747	0.8756	23.7794	0.7892	26.1157	0.8604	26.7442	0.8765

the promising performance of the proposed model. Compared with the classical TV, FBD, TSM, and TV-FF, our TV-FBD model has the highest PSNR and MSSIM values.

It should point out that our model cannot very well recover the texture in the restoration due to the use of TV regularization. And the TV minimization favors the solutions that are piecewise constant, which implies that the oscillatory components such as textures are removed. Actually, we have proposed a multi-scale variational image decomposition model to extract texture in our previous work [32]. This implies that if the extracted texture can be incorporated into the restoration, the quality of the restoration can be improved. Our successive research will focus on how to incorporate texture representation into our forward-backward diffusion function based on total variation.

Table 3 CPU time for FBD and TV-FBD models (second)

Image	FBD	TV-FBD	Image	FBD	TV-FBD
Panda ($\sigma = 5$)	2.4380	1.7895	Butterfly ($\sigma = 5$)	2.4546	1.7863
Panda ($\sigma = 10$)	2.7710	1.8215	Butterfly ($\sigma = 10$)	2.7864	1.8264
Panda ($\sigma = 15$)	2.7135	1.7886	Butterfly ($\sigma = 15$)	2.7657	1.7756
Panda ($\sigma = 20$)	2.7642	1.7541	Butterfly ($\sigma = 20$)	2.7689	1.7787
Cameraman ($\sigma = 5$)	2.4346	1.7878	Lenna ($\sigma = 5$)	2.4789	1.7826
Cameraman ($\sigma = 10$)	2.6879	1.7996	Lenna ($\sigma = 10$)	2.7883	1.7864
Cameraman ($\sigma = 15$)	2.6450	1.7231	Lenna ($\sigma = 15$)	2.7687	1.7725
Cameraman ($\sigma = 20$)	2.6661	1.7325	Lenna ($\sigma = 20$)	2.7965	1.7896

References
1. L Shao, R Yan, X Li et al., From heuristic optimization to dictionary learning: a review and comprehensive comparison of image denoising algorithms. Cyb. IEEE Trans. **44**(7), 1001–1013 (2014)
2. C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in *Proc. of 1998 Sixth International Conference on Computer Vision*, IEEE press, Piscataway, N.J. 1998, pp. 839–846
3. A Buades, B Coll, JM Morel, A review of image denoising algorithms, with a new one. Multiscale Model. Simul. **4**(2), 490–530 (2005)
4. R Yan, L Shao, L Liu et al., Natural image denoising using evolved local adaptive filters. Signal Process. **103**, 36–44 (2014)
5. Y Abe, Y liguni, *Fast Computation of the High Resolution Image Restoration by Using the Discrete Cosine Transform in Proc. Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on IEEE*, 2007, pp. 745–748
6. R Yan, L Shao, Y Liu, Nonlocal hierarchical dictionary learning using wavelets for image denoising. IEEE Trans. Image Process. **22**(12), 4689–4698 (2013)
7. MA Kutay, HM Ozaktas, Optimal image restoration with the fractional Fourier transform. J. Opt. Soc. Am. A **5**(4), 825–833 (1998)
8. P Perona, J Malik, Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**(7), 629–639 (1990)
9. J Weicker, Anisotropic diffusion in image processing. B.g.teubner Stuttgart **16**(1), 272 (1997)
10. D Zhao, HQ Du, WB Mei, Hybrid weighted l_1-total variation constrained reconstruction for MR image. Chinese J. Electron. **23**(4), 747–752 (2014)
11. A Buades, TM Le, JM Morel et al., Fast cartoon + texture image filters. IEEE Trans. Image Process. **19**(8), 1978–1986 (2010)
12. LI Rudin, S Osher, E Fatemi, Nonlinear total variation based noise removal algorithms. Physica D **60**(1), 259–268 (1992)

13. A Chambolle, PL Lions, Image recovery via total variation minimization and related problems. Numerische Mathematik **76**(2), 167–188 (1997)
14. Z Ren, C He, Q Zhang, Fractional order total variation regularization for image super-resolution. Signal Process. **93**(9), 2408–2421 (2013)
15. JF Aujol, A Chambolle, Dual norms and image decomposition models. Int. J. Comput. Vision **63**(1), 85–104 (2005)
16. JL Starck, M Elad, DL Donoho, Image decomposition via the combination of sparse representations and a variational approach. IEEE Trans. Image Process. **14**(10), 1570–1582 (2005)
17. J. Xu, Z. Chang, J. Fan et al. Noisy image magnification with total variation regularization and order-changed dictionary learning. EURASIP J. Adv. Signal Process, **2015**, 1–13 (2015)
18. D Strong, T Chan, Edge-preserving and scale-dependent properties of total variation regularization. Inverse Probl. **19**(6), 165–187 (2003)
19. L Li, W Feng, J Zhang, Contrast enhancement based single image dehazing VIA TV-l$_1$ minimization, in *Proc. 2014 IEEE International Conference on Multimedia and Expo (ICME)IEEE Computer Society*, 2014, pp. 1–6
20. LX Chen, *Study on Image Restoration Models Based on PDE and Image Enhancement and Segmentation Algorithms, Xi'an, Xidian University*, 2010
21. S Osher, L Rudin, Feature oriented image enhancement using shock filters. SIAM Num. Anal. **27**(4), 919–940 (1990)
22. L Alvarez, L Mazorra, Signal and image restoration using shock filters and anisotropic diffusion. SIAM J. Numer. Anal. **31**(2), 590–605 (1994)
23. G Gilboa, N Soehen, Y Zeevi, Forward and backward diffusion processes for adaptive image enhancement and denoising. IEEE Trans. Image Process. **11**(7), 689–703 (2002)
24. G Sapiro, *Geometric Partial Differential Equations and Image Analysis* (Cambridge Universit Press, London, 2001)
25. KE Jing, YQ Hou, DK Wang et al., An improved algorithm for shape preserving contrast enhancement. Acta Photonica Sinica **38**(1), 2014–219 (2009)
26. R Highnam, M Brady, Model-based image enhancement of far infrared images. IEEE Trans. Pattern Anal. Mach. Intell. **19**(4), 410–415 (1997)
27. C Wang, ZF Ye, Variational enhancement for infrared images. J. Infrared Millimeter Waves **25**(4), 306–310 (2006)
28. M Tang, SD Ma, J Xiao, Model-based adaptive enhancement of far infrared image sequences. Pattern Recogn. Lett. **21**(00), 827–835 (2000)
29. A Chambolle, An algorithm for total variation minimization and applications. J. Math. Imaging Vis. **20**(1–2), 89–97 (2004)
30. R Acar, CR Vogel, Analysis of total variation penalty methods for ill-posed problems. Inverse Probl. **10**(6), 1217–1229 (1994)
31. Z Wang, AC Bovik, HR Sheikh et al., Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
32. LM Tang, CJ He, Multiscale texture extraction with hierarchical (BV,G$_p$,L^2) decomposition. J. Math. Imaging Vis. **45**(2), 148–163 (2013)