

Total Generalized Variation Based Denoising Models for Ultrasound Images

Myeongmin Kang¹ · Myungjoo Kang¹ · Miyoun Jung²

Received: 9 September 2016 / Revised: 3 December 2016 / Accepted: 2 January 2017 /
Published online: 20 January 2017
© Springer Science+Business Media New York 2017

Abstract In this paper, we introduce a class of variational models for the restoration of ultrasound images corrupted by noise. The proposed models involve the convex or nonconvex total generalized variation regularization. The total generalized variation regularization ameliorates the staircasing artifacts that appear in the restored images of total variation based models. Incorporating total generalized variation regularization with nonconvexity helps preserve edges in the restored images. To realize the proposed convex model, we adopt the alternating direction method of multipliers, and the iteratively reweighted ℓ_1 algorithm is employed to handle the nonconvex model. These methods result in fast and efficient optimization algorithms for solving our models. Numerical experiments demonstrate that the proposed models are superior to other state-of-the-art models.

Keywords Ultrasound image denoising · Total generalized variation · Nonconvex regularization · Alternating direction method of multipliers · Iteratively reweighted algorithm

1 Introduction

In various image systems, images are often corrupted by noise during the image acquisition process. Therefore, image denoising is an important and elemental problem in image processing. It aims to remove noise in the observed images and find the best possible restored

✉ Miyoun Jung
mjung@hufs.ac.kr
Myeongmin Kang
wjdjr1@snu.ac.kr
Myungjoo Kang
mkang@snu.ac.kr

¹ Department of Mathematical Sciences, Seoul National University, Seoul, Korea

² Department of Mathematics, Hankuk University of Foreign Studies, Yongin, Korea

image corresponding to the unknown clean image. In this article, we consider the denoising problem for ultrasound images. The removal of additive Gaussian noise has been well-studied and there are some methods that produce remarkable denoising results. However, ultrasound images involve more complicated noise such as multiplicative noise [1, 42, 43].

Let Ω be a bounded open domain with a compact Lipschitz boundary. Experimental measurements in [28] show that an observed ultrasound image $f : \Omega \rightarrow \mathbb{R}$ can be modeled using the following form:

$$f = u + \sqrt{u}n, \tag{1}$$

where $u : \Omega \rightarrow \mathbb{R}$ is a true image and n is a Gaussian random variable with zero mean and a standard deviation σ , i.e. $n \sim \mathcal{N}(0, \sigma)$.

To reconstruct u from the noisy data f , various filtering-based methods, such as Lee filter [26], Kuan filter [25], and PDE-based methods have been proposed. In particular, from the degradation model (1), Krissian et al. [24] derived the following convex data-fidelity term

$$\int_{\Omega} \frac{(f - u)^2}{u}. \tag{2}$$

In a variational framework, a data-fidelity function is usually derived from the maximum a posteriori (MAP) estimation based on the degradation model and a probability density function of noise. In this case, the degradation model in (1) has a singular form. Hence, finding a proper data-fidelity function according to the MAP estimation is very complicated. In fact, the data-fidelity term in (2) is not derived from the MAP estimation. This term was obtained by adapting the work of Rudin et al. [39] dealing with the Gaussian distribution.

Several variational models involving the term (2) have been proposed for solving the inverse problem in (1). First, Jin et al. [21] proposed the following total variation (TV) regularization based model:

$$\min_{u>0} \int_{\Omega} |Du| + \lambda \int_{\Omega} \frac{(u - f)^2}{u} dx, \tag{3}$$

where $\lambda > 0$ is a tuning parameter, and TV [39] is defined as

$$\int_{\Omega} |Du| = \sup_v \left\{ \int_{\Omega} u \operatorname{div}(v) dx \mid v \in C_c^1(\Omega), \|v\|_{\infty} \leq 1 \right\},$$

where the vector measure Du represents the directional or weak gradient of u , and $\|\cdot\|_{\infty}$ is the essential supremum norm. This regularization has been widely used in image processing owing to its edge-preserving property. The authors in [21] proved the existence and uniqueness of the solution of model (3), and it was shown that the solution belongs to $[\min(f), \max(f)]$ when $f > 0$. From now on, we call the model in (3) the ‘‘TV model’’. More recently, the well-balanced speckle noise reduction (WBSN) model, associated with an edge detection technique, was proposed in [3]:

$$\min_u \int_{\Omega} g|\nabla u|^2 dx + \lambda \int_{\Omega} (1 - g) \frac{(u - f)^2}{u} dx, \tag{4}$$

where g is a smooth non-increasing function, defined as $g(x) = \frac{1}{1+k|(G_{\sigma} * u)(x)|}$ with a Gaussian function G_{σ} and a constant value $k > 0$. To minimize the model (4) with a fixed g , the authors in [3] obtained a numerical solution by solving the Euler–Lagrangian equation for u and then updating g from its definition.

Despite the benefits of TV regularization, TV-based methods cause staircasing artifacts in the restored images. There have been many efforts to improve the performance of TV using

higher-order regularization. In [12,29], the second-order TV regularization was proposed to reduce the staircasing effects. Furthermore, several hybrid regularization methods combining TV with a higher-order regularizer have been considered. The first one is a convex combination of TV and second-order TV [27,37], which takes advantages of both first- and second-order derivatives. In an earlier work, the infimal convolution TV (ICTV) regularizer [10,40] was developed, which takes the infimal convolution of TV and second-order TV. Recently, as a generalization of ICTV, the total generalized variation (TGV) regularizer [6] was proposed. The TGV regularization results in restored images with sufficiently denoised smooth regions and well-preserved edges. In this work, we employ the TGV regularization for the removal of noise in the ultrasound images.

In early works, numerical algorithms [9,39] that directly solve the Euler–Lagrangian equation using the finite difference method were used to find a minimizer of an energy functional. Recently, because of the development of the operating splitting technique [20], many unconstrained minimization problems in image processing can be converted into linearly constrained problems with simple objective functions. This also leads to numerous iterative algorithms [11,46,47] for solving linearly constrained minimization problems, which commonly have some subproblems that can be solved alternately. In particular, the alternating direction method of multipliers (ADMM) [15,20] is one of the most well-known convex optimization algorithms. The ADMM is equivalent to the Douglas-Rachford splitting algorithm [14], and it has been applied to efficiently solve various image processing problems [18,19,41]. We adopt the ADMM here to solve our convex minimization problem.

Since the seminal work of Geman and Geman [17], many studies [4,16,32,38] have demonstrated that nonconvex regularization performs better than convex regularization for image restoration by preserving edges. Specifically, Krishnan et al. [23] demonstrated the superiority of the nonconvex ℓ_q -norm of gradient for $0.5 < q < 0.8$ over its ℓ_1 norm in an image denoising problem. Moreover, several studies [22,36] have utilized a nonconvex hybrid higher-order regularizer and shown its superiority over nonconvex TV regularizers. Recently, Ochs et al. [34,35] introduced a nonconvex TGV for some applications in computer vision. We also utilize a nonconvex TGV for the ultrasound image denoising problem. On the other hand, many efficient algorithms for solving nonconvex minimization problems have been developed [2,30,31,33,44]. Recently, the iteratively convex majorization-minimization methods for solving nonsmooth nonconvex optimization problems have been proposed in [34,35] with convergence analysis. These methods are the generalized versions of the reweighted ℓ_1 algorithm in [8]. In this work, we exploit the algorithm in [34] for solving our nonconvex model.

In this article, we introduce two TGV based minimization models for the restoration of ultrasound images polluted by noise. We make use of the convex or nonconvex TGV as a regularization term in our minimization problems. The TGV regularization allows the smooth transition regions in ultrasound images to be restored naturally. In addition, the nonconvex TGV preserves the edges of the restored images well. We also present a convergence analysis for the convex TGV model and efficient optimization algorithms for solving the proposed convex or nonconvex models.

The rest of our paper is organized as follows. In Sect. 2, we review the TGV regularization and describe the ADMM for the general convex minimization problem with linear constraints. We also present the iteratively reweighted ℓ_1 algorithm for the linearly constrained nonconvex minimization problem. In Sect. 3, we introduce our minimization models with the convex TGV or nonconvex TGV regularization for ultrasound image denoising. Specifically, Sect. 3.1 presents the convergence analysis for our convex TGV model and an optimization algorithm for solving the model. Section 3.2 describes the nonconvex TGV-based model and

its optimization algorithm. In Sect. 4, we provide some numerical experiments with comparisons to existing denoising models. Lastly, in Sect. 5, we summarize this work and discuss future work.

2 Background

2.1 Total Generalized Variation

The total generalized variation (TGV) was proposed by Bredies et al. [6], as a generalization of the infimal convolution of TV and second-order TV regularizers [10], as follows:

$$TGV_{\alpha}^k(u) = \sup_v \left\{ \int_{\Omega} u \operatorname{div}^k(v) \, dx \mid v \in C_c^k(\Omega, \operatorname{Sym}^k(\mathbb{R}^d)), \right. \\ \left. \|\operatorname{div}^j(v)\|_{\infty} \leq \alpha_j, \quad j = 0, \dots, k - 1 \right\}, \tag{5}$$

where $\operatorname{Sym}^k(\mathbb{R}^d)$ is the space of symmetric k -tensors on \mathbb{R}^d as

$$\operatorname{Sym}^k(\mathbb{R}^d) = \{ \xi : \mathbb{R}^d \times \dots \times \mathbb{R}^d \mid \xi \text{ is } k\text{-linear and symmetric} \},$$

and the $C_c^k(\Omega, \operatorname{Sym}^k(\mathbb{R}^d))$ represents the space of compactly supported symmetric tensor fields. Here, $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_{k-1})$ is a fixed positive vector. When $k = 1$ and $\alpha_0 = 1$, the definition of TGV is identical to that of TV, which implies that TGV is a generalization of TV.

According to [6], the TGV_{α}^k can be represented as a k -fold infimal convolution by employing the Fenchel–Rockafellar duality formula, which is given by

$$TGV_{\alpha}^k(u) = \inf_{u_j, \forall j=1, \dots, k} \sum_{j=1}^k \alpha_{k-j} \int_{\Omega} |\mathcal{E}(u_{j-1}) - u_j| \, dx, \tag{6}$$

where $u_0 = u, u_k = 0, u_j \in C_c^{k-j}(\Omega, \operatorname{Sym}^j(\mathbb{R}^d))$ for all $j = 1, \dots, k$, and \mathcal{E} is the distributional symmetrized derivative, i.e., $\mathcal{E}(u_{j-1}) = \frac{\nabla u_{j-1} + (\nabla u_{j-1})^T}{2}$; for instance, when $k = 2$, TGV_{α}^2 involves the l_1 -norm of $\mathcal{E}(u_0) - u_1 = \nabla u - u_1$ and $\mathcal{E}(u_1)$ with appropriate weights. In this case, the minimization in (6) can be interpreted as an optimal balancing between the first- and second-order derivatives of u in terms of sparse penalization (via the Radon norm). Therefore, from this representation (6), one can say that $TGV_{\alpha}^k(u)$ automatically balances the first- and higher-order derivatives (up to the order k), rather than using any fixed combination. As a result, it reduces the staircasing effects of the TV functional.

In this work, we particularly utilize the second order TGV, obtained by letting $k = 2$ in (5). The followings describe several properties of the second order TGV, denoted by TGV_{α}^2 [5, 7]. First, $TGV_{\alpha}^2 : L^p(\Omega) \rightarrow \mathbb{R} \cup \{\infty\}$ is a proper, convex, and lower semicontinuous function for $1 \leq p < \infty$. Moreover, there exist some positive constants c and C that, for any $u \in L^1(\Omega)$, satisfy

$$c(\|u\|_1 + TV(u)) \leq \|u\|_1 + TGV_{\alpha}^2(u) \leq C(\|u\|_1 + TV(u)).$$

Lastly, TGV_{α}^2 satisfies the Poincaré inequality, i.e., there exists a positive constant C' such that for any $u \in L^1(\Omega)$,

$$\|u - P(u)\|_2 \leq C' \cdot TGV_{\alpha}^2(u), \tag{7}$$

where P is a linear projection to the space of affine functions. Using these properties, we will prove the existence and uniqueness of the solutions of our proposed model in the next section.

2.2 Alternating Direction Method of Multipliers

This section reviews a convex optimization algorithm called the alternating direction method of multipliers (ADMM).

Let us consider the following linearly constrained minimization problem:

$$\min_{\mathbf{x}, \mathbf{y}} H(\mathbf{x}) + G(\mathbf{y}) \quad \text{such that} \quad \mathbf{Ax} + \mathbf{By} = b, \tag{8}$$

where $H : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ and $G : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{\infty\}$ are convex, proper, and lower semicontinuous functions, $A \in \mathbb{R}^{k \times n}$, $B \in \mathbb{R}^{k \times m}$, and $b \in \mathbb{R}^k$.

The augmented Lagrangian function for problem (8) is given by

$$\mathcal{L}_A(\mathbf{x}, \mathbf{y}, \lambda; \mu) = H(\mathbf{x}) + G(\mathbf{y}) - \lambda^T (\mathbf{Ax} + \mathbf{By} - b) + \frac{\mu}{2} \|\mathbf{Ax} + \mathbf{By} - b\|_2^2,$$

where λ is a Lagrange multiplier vector or dual variable and $\mu > 0$ is a penalty parameter.

The augmented Lagrangian method minimizes \mathcal{L}_A over \mathbf{x} and \mathbf{y} jointly for fixed λ and then updates λ for fixed \mathbf{x} and \mathbf{y} . Since the augmented Lagrangian function has a coupled penalty term with respect to (\mathbf{x}, \mathbf{y}) , it is hard to solve the subproblem over (\mathbf{x}, \mathbf{y}) in many cases. To overcome this drawback, the ADMM alternately solves the subproblem, by minimizing the augmented Lagrangian function \mathcal{L}_A over one variable (\mathbf{x} or \mathbf{y}) with the other variable fixed and performs only one outer iteration, which yields

$$\begin{cases} \mathbf{x}_{k+1} = \arg \min_{\mathbf{x}} \mathcal{L}_A(\mathbf{x}, \mathbf{y}_{k+1}, \lambda_k; \mu), \\ \mathbf{y}_{k+1} = \arg \min_{\mathbf{y}} \mathcal{L}_A(\mathbf{x}_{k+1}, \mathbf{y}, \lambda_k; \mu), \\ \lambda_{k+1} = \lambda_k - \nu \mu (\mathbf{Ax}_{k+1} + \mathbf{By}_{k+1} - b), \end{cases} \tag{9}$$

where $\nu \in (0, \frac{1+\sqrt{5}}{2})$ is a parameter. The convergence of the ADMM was proved in [13] as follows:

Theorem 1 *Assume that there exists a saddle point $(\mathbf{x}_*, \mathbf{y}_*, \lambda_*)$ of problem (8) and that A and B have full column rank or H and G are coercive. Then, for any $\nu \in (0, \frac{1+\sqrt{5}}{2})$, the tuple $(\mathbf{x}_k, \mathbf{y}_k, \lambda_k)$ generated by the ADMM (9) satisfies*

$$\lambda_k \rightarrow \lambda_*, \quad \mathbf{Ax}_k \rightarrow \mathbf{Ax}_*, \quad \text{and} \quad \mathbf{By}_k \rightarrow \mathbf{By}_*, \quad \text{as } k \rightarrow \infty.$$

The ADMM is adopted to solve our convex minimization problem.

2.3 Iterative Reweighted ℓ_1 Algorithm

In this subsection, we recall the iteratively reweighted ℓ_1 algorithm (IRLA) [34], which was proposed for solving nonconvex optimization problems.

For a convex function f , a vector g is called a *subgradient* at a point \mathbf{x} if for any vector \mathbf{y} , $f(\mathbf{y}) - f(\mathbf{x}) - g^T(\mathbf{y} - \mathbf{x}) \geq 0$ is satisfied. The set of all subgradients at \mathbf{x} is called the *subdifferential* at \mathbf{x} and is denoted by $\partial f(\mathbf{x})$. If \tilde{f} is a concave function, then $-\tilde{f}$ is a convex function. An element of $-\partial(-\tilde{f})$ is called a *limiting-supergradient* of \tilde{f} .

Let us consider the following nonconvex minimization optimization problem with linear constraints:

$$\min_{\mathbf{x}} E_1(\mathbf{x}) + E_2(|\mathbf{x}|), \quad \text{subject to} \quad \mathbf{Ax} = b, \tag{10}$$

where E_1 is a proper, lower semicontinuous, and convex function, E_2 is a concave and coordinatewise nondecreasing function, and $|\cdot|$ is coordinatewise. When the IRLA is applied to problem (10), it leads to the following iterative algorithm,

$$\mathbf{x}_{k+1} = \arg \min_{\mathbf{Ax}=b} E_1(\mathbf{x}) + \langle \tilde{\mathbf{x}}_k, |\mathbf{x}| \rangle, \tag{11}$$

where $\tilde{\mathbf{x}}_k$ is a limiting-supergradient of $E_2(\cdot)$ at $|\mathbf{x}_k|$. If E_2 is a differentiable function, then the limiting-supergradient of E_2 is given by ∇E_2 , from the property of the subdifferential.

The convergence of IRLA was proved in [34]:

Theorem 2 *Let (\mathbf{x}_{k+1}) be the sequence generated by the IRAL algorithm (11), and assume that $E_1(\mathbf{x}) + E_2(|\mathbf{x}|) \rightarrow \infty$ as $\|\mathbf{x}\| \rightarrow \infty$, and $\mathbf{Ax} = b$. Then, (\mathbf{x}_{k+1}) is bounded and has at least one accumulation point.*

In the following section, we apply the IRLA to solve our proposed nonconvex model.

3 Proposed Models and Algorithms

In this section, we present the convex or nonconvex TGV regularization based models for ultrasound image denoising. We also present efficient optimization algorithms to solve our proposed models by utilizing the ADMM and the IRLA.

3.1 Convex TGV Model

First, we propose the following minimization model, which includes the the second-order TGV regularizer and the data fidelity term in (2):

$$\min_{u \in S(\Omega)} E(u) = TGV_\alpha^2(u) + \gamma \int_\Omega \frac{(f - u)^2}{u} dx, \tag{12}$$

where $S(\Omega) = \{u \in L^2(\Omega) : u > 0\}$ and γ is a regularization parameter that controls the balance between the TGV regularization and data fidelity terms. From the definition of $TGV_\alpha^2(u)$ in (6), the TGV model (12) can be reformulated as follows

$$\min_{u \in S(\Omega), v} \alpha_1 \int_\Omega |\nabla u - v| dx + \alpha_0 \int_\Omega |\mathcal{E}(v)| dx + \gamma \int_\Omega \frac{(f - u)^2}{u} dx, \tag{13}$$

where $\alpha_1, \alpha_0 > 0$ are parameters that balance the first- and second-order regularization terms. It is straightforward that the model (12) is convex, due to the convexity of the data-fidelity and TGV terms, as shown in [24] and [7], respectively.

The following theorem exhibits the existence and uniqueness of the solutions for the proposed model in (12).

Theorem 3 *The proposed variational model in (12) admits a unique solution.*

Proof The proof of this theorem is given in the appendix, which is based on the properties of the second-order TGV. □

3.1.1 ADMM for Solving Our Model (12)

In this subsection, we introduce an optimization algorithm for solving our convex TGV model in (12). From its equivalent model in (13), we can derive the discrete TGV_α^2 functional as

$$TGV_\alpha^2(u) = \inf_{v \in C_c^2(\Omega, \mathbb{R}^2)} \alpha_1 \|\nabla u - v\|_1 + \alpha_0 \|\mathcal{E}(v)\|_1,$$

where the operators ∇u and $\mathcal{E}(v)$ are given by

$$\nabla u = \begin{bmatrix} \partial_x u \\ \partial_y u \end{bmatrix} \quad \text{and} \quad \mathcal{E}(v) = \begin{bmatrix} \partial_x v_1 & \frac{1}{2}(\partial_x v_2 + \partial_y v_1) \\ \frac{1}{2}(\partial_x v_2 + \partial_y v_1) & \partial_y v_2 \end{bmatrix}.$$

With this discrete TGV formulation, the discretized version of our proposed model in (13) is as follows:

$$\min_{u > 0, v} \alpha_1 \|\nabla u - v\|_1 + \alpha_0 \|\mathcal{E}(v)\|_1 + \gamma \left\langle \frac{(f - u)^2}{u}, \mathbf{1} \right\rangle. \tag{14}$$

Here, we assume that $f_{i,j} \neq 0$ for all $(i, j) \in \Omega$ (see Remark 1 for further explanation). Furthermore, the set $S(\Omega)$ is not a closed set. Hence, we consider the closed set $\bar{S}(\Omega) = \{u \in L^2(\Omega) : u \geq 0\}$ instead of $S(\Omega)$ for the numerical algorithms. Since the term $\frac{(f-u)^2}{u}$ blows up at $u = 0$, the close set $\bar{S}(\Omega)$ is a simple relaxation of $S(\Omega)$. Therefore, we solve the following minimization problem instead of the original problem in (14):

$$\min_{u \geq 0, v} \alpha_1 \|\nabla u - v\|_1 + \alpha_0 \|\mathcal{E}(v)\|_1 + \gamma \left\langle \frac{(f - u)^2}{u}, \mathbf{1} \right\rangle. \tag{15}$$

We now introduce three auxiliary variables: $z, d = (d_1, d_2)^T$, and $w = (w_1, w_2, w_3)^T$. Using these auxiliary variables, the unconstrained model in (15) can be rewritten as the following constrained form,

$$\begin{aligned} \min_{u, v, z \geq 0, d, w} \quad & \alpha_1 \|d\|_1 + \alpha_0 \|w\|_1 + \gamma \left\langle \frac{(f - z)^2}{z}, \mathbf{1} \right\rangle, \\ \text{subject to} \quad & z = u, \quad d = \nabla u - v, \quad w = \mathcal{E}(v). \end{aligned} \tag{16}$$

Letting $D(z) = \left\langle \frac{(f-z)^2}{z}, \mathbf{1} \right\rangle$ and setting

$$A = \begin{bmatrix} I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} z \\ d \\ w \end{bmatrix}, \quad B = \begin{bmatrix} -I & \mathbf{0} \\ -\nabla & I \\ \mathbf{0} & -\mathcal{E} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} u \\ v \end{bmatrix}, \quad \lambda = \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \lambda^3 \end{bmatrix}, \tag{17}$$

$$H(\mathbf{x}) = \alpha_1 \|d\|_1 + \alpha_0 \|w\|_1 + \gamma D(z) \quad \text{and} \quad G(\mathbf{y}) = 0,$$

the constrained problem in (16) can be rewritten in the form in (8), so we can use the ADMM to solve it. The augmented Lagrangian function for problem (16) is given by

$$\begin{aligned} \mathcal{L}_A(z, d, w, u, v, \lambda^1, \lambda^2, \lambda^3; \mu) = & \alpha_1 \|d\|_1 + \alpha_0 \|w\|_1 + \gamma D(z) \\ & - (\lambda^1)^T (z - u) - (\lambda^2)^T (d - \nabla u + v) - (\lambda^3)^T (w - \mathcal{E}(v)) \\ & + \frac{\mu}{2} \|z - u\|_2^2 + \frac{\mu}{2} \|d - \nabla u + v\|_2^2 + \frac{\mu}{2} \|w - \mathcal{E}(v)\|_2^2. \end{aligned} \tag{18}$$

The ADMM can then applied to problem (16) as follows,

$$\begin{cases} \begin{pmatrix} z_{k+1} \\ d_{k+1} \\ w_{k+1} \end{pmatrix} = \arg \min_{z \geq 0, d, w} \mathcal{L}_A(z, d, w, u_k, v_k, \lambda_k^1, \lambda_k^2, \lambda_k^3; \mu), \\ \begin{pmatrix} u_{k+1} \\ v_{k+1} \end{pmatrix} = \arg \min_{u, v} \mathcal{L}_A(z_{k+1}, d_{k+1}, w_{k+1}, u, v, \lambda_k^1, \lambda_k^2, \lambda_k^3; \mu), \\ \lambda_{k+1}^1 = \lambda_k^1 - \nu\mu(z_{k+1} - u_{k+1}), \\ \lambda_{k+1}^2 = \lambda_k^2 - \nu\mu(d_{k+1} - \nabla u_{k+1} + v_{k+1}), \\ \lambda_{k+1}^3 = \lambda_k^3 - \nu\mu(w_{k+1} - \mathcal{E}(v_{k+1})). \end{cases}$$

The variables $z, d,$ and w in the augmented Lagrangian function are independent of each other. Hence, we can solve the subproblem for $(z_{k+1}, d_{k+1}, w_{k+1})$ in the ADMM separately. We can then obtain the ADMM for model (16) as follows,

$$\begin{cases} d_{k+1} = \arg \min_d \alpha_1 \|d\|_1 - (\lambda_k^2)^T(d) + \frac{\mu}{2} \|d - \nabla u_k + v_k\|_2^2, \\ w_{k+1} = \arg \min_w \alpha_0 \|w\|_1 - (\lambda_k^3)^T(w) + \frac{\mu}{2} \|w - \mathcal{E}(v_k)\|_2^2, \\ z_{k+1} = \arg \min_{z \geq 0} \gamma D(z) - (\lambda_k^1)^T(z) + \frac{\mu}{2} \|z - u_k\|_2^2, \\ \begin{pmatrix} u_{k+1} \\ v_{k+1} \end{pmatrix} = \arg \min_{u, v} (\lambda_k^2)^T(\nabla u - v) + (\lambda_k^3)^T(\mathcal{E}(v)) + \lambda_k^1 T u \\ \quad + \frac{\mu}{2} \|\nabla u - v - d_{k+1}\|_2^2 + \frac{\mu}{2} \|\mathcal{E}(v) - w_{k+1}\|_2^2 + \frac{\mu}{2} \|u - z_{k+1}\|_2^2, \\ \lambda_{k+1}^1 = \lambda_k^1 - \nu\mu(z_{k+1} - u_{k+1}), \\ \lambda_{k+1}^2 = \lambda_k^2 - \nu\mu(d_{k+1} - \nabla u_{k+1} + v_{k+1}), \\ \lambda_{k+1}^3 = \lambda_k^3 - \nu\mu(w_{k+1} - \mathcal{E}(v_{k+1})). \end{cases} \tag{19}$$

From Theorem 1, we can derive the following convergence result.

Theorem 4 For any $\nu \in (0, \frac{1+\sqrt{5}}{2})$, the tuple $(z_k, d_k, w_k, u_k, v_k, \lambda_k^1, \lambda_k^2, \lambda_k^3)$ generated by the ADMM (19) converges to a saddle point $(z_*, d_*, w_*, u_*, v_*, \lambda_*^1, \lambda_*^2, \lambda_*^3)$ of problem (16).

Proof Since $D(z)$ is not convex when $z < 0$, we can rewrite the constrained model in (16) as the following equivalent formula

$$\begin{aligned} & \min_{u, v, z, d, w} \alpha_1 \|d\|_1 + \alpha_0 \|w\|_1 + \gamma \left\langle \tilde{D}(z), \mathbf{1} \right\rangle, \\ & \text{subject to } z = u, d = \nabla u - v \text{ and } w = \mathcal{E}(v), \end{aligned} \tag{20}$$

where $\tilde{D}(z)$ is the extended value function of $\frac{(f-z)^2}{z}$, which is defined as

$$(\tilde{D}(z))_i = \begin{cases} \frac{(f - z_i)^2}{z_i}, & \text{if } z_i > 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is trivial to show that $\tilde{D}(z)$ is convex, proper, and lower semi-continuous.

Moreover, the z -subproblem in the ADMM (the third line in (19)) is equivalent to

$$\arg \min_z \gamma \langle \tilde{D}(z), \mathbf{1} \rangle - (\lambda_k^1)^T(z) + \frac{\mu}{2} \|z - u_k\|_2^2.$$

Hence, the tuple generated by the ADMM (19) is exactly the same as the tuple generated by the ADMM when it is applied to problem (20). Since A and B have full column rank, the assertion of this theorem follows from Theorem 1. \square

Solving the subproblems in (19). We now solve the subproblems in the ADMM (19). First, the subproblems for d_{k+1} and w_{k+1} in (19) have closed form solutions obtained using the shrink operator as follows,

$$d_{k+1} = \text{shrink} \left(\frac{\lambda_k^2}{\mu} + \nabla u_k - v_k, \frac{\alpha_1}{\mu} \right) \text{ and } w_{k+1} = \text{shrink} \left(\frac{\lambda_k^3}{\mu} + \mathcal{E}(v_k), \frac{\alpha_0}{\mu} \right),$$

where the shrink operator $\text{shrink}(s, t)$ is defined as

$$\text{shrink}(s, t) = \frac{s}{\|s\|_1} \max(\|s\|_1 - t, 0).$$

To solve the z -subproblem, we take the partial derivative of the energy with respect to z , which leads to the following normal equation:

$$\gamma \left(1 - \frac{f^2}{z^2} \right) - \lambda_k^1 + \mu(z - u_k) = 0.$$

Thus, the exact solution of z -subproblem can be obtained as a positive solution of the following cubic equation:

$$\mu z^3 + (\gamma - \lambda_k^1 - \mu u_k) z^2 - \gamma f^2 = 0. \tag{21}$$

The following Lemma shows that this cubic equation (21) has one real positive solution.

Lemma 1 *The cubic equation (21) has only one positive solution.*

Proof Let (t_1, t_2, t_3) be the solution of the cubic equation in (21). It follows from (21) that

$$t_1 + t_2 + t_3 = -\frac{(\gamma - \lambda_k^1 - \mu u_k)}{\mu}, \quad t_1 t_2 + t_2 t_3 + t_1 t_3 = 0, \quad t_1 t_2 t_3 = \frac{\gamma f^2}{\mu}. \tag{22}$$

Let $\alpha = t_2 + t_3$ and $\beta = t_2 \cdot t_3$. It is trivial to find that a cubic equation has at least one real solution. Thus, we let t_1 be a real solution of (21) without loss of generality. It is easy to show that $t_i \neq 0$ for all $i = 1, 2, 3$.

Case (1) Assume that t_1 is negative. Since $\frac{\gamma f^2}{\mu} > 0, t_2 \cdot t_3 < 0$. Using this inequality and (22), we have $t_1(t_2 + t_3) > 0$ and thus $t_2 + t_3 < 0$ is satisfied. Thus, we obtain

$$\alpha < 0 \quad \text{and} \quad \beta < 0. \tag{23}$$

Because (t_2, t_3) are the solutions of $t^2 - \alpha t + \beta = 0$, they satisfy the following equations without loss of generality:

$$t_2 = \frac{\alpha + \sqrt{\alpha^2 - 4\beta}}{2} \quad \text{and} \quad t_3 = \frac{\alpha - \sqrt{\alpha^2 - 4\beta}}{2}.$$

From (23), t_2 and t_3 are real numbers, t_3 is negative, and t_2 is positive.

Case (2) Assume that t_1 is positive. It is enough to show that t_2 and t_3 are either complex numbers or negative. Because $\frac{\gamma f^2}{\mu} > 0$, then $t_2 \cdot t_3 > 0$. Similar to the case 1, we have $t_2 + t_3 < 0$, which leads to

$$\alpha < 0 \quad \text{and} \quad \beta > 0. \tag{24}$$

Without loss of generality, (t_2, t_3) can be represented as

$$t_2 = \frac{\alpha + \sqrt{\alpha^2 - 4\beta}}{2} \quad \text{and} \quad t_3 = \frac{\alpha - \sqrt{\alpha^2 - 4\beta}}{2}.$$

If $\alpha^2 - 4\beta < 0$, then (t_2, t_3) are complex numbers. If $\alpha^2 - 4\beta \geq 0$, it follows from (24) that $\alpha \pm \sqrt{\alpha^2 - 4\beta}$ are negative. Thus, in this case, (t_2, t_3) are negative. \square

In practice, we compute the roots of the cubic Eq. (21), using the cubic formula that is the closed-form solution for a cubic equation. Then we choose the positive one from the three solutions.

Lastly, the subproblem for (u_{k+1}, v_{k+1}) in (19) is a least square problem,

$$\arg \min_{\mathbf{y}} \|\mathbf{B}\mathbf{y} + \mathbf{b}\|_2^2$$

with $\mathbf{b} = (z_{k+1} - \lambda_k^1/\mu, d_{k+1} - \lambda_k^2/\mu, w_{k+1} - \lambda_k^3/\mu)^T$. Hence, the solution \mathbf{y}_{k+1} of the above least square problem is the solution of $B^T B \mathbf{y} = -B^T \mathbf{b}$. The matrix $B^T B$ is given by

$$B^T B = \begin{bmatrix} I + \nabla^T \nabla & -\nabla^T \\ -\nabla & I + \mathcal{E}^T \mathcal{E} \end{bmatrix}.$$

The block matrices in $B^T B$ can be diagonalized by the Fourier transform under the periodic boundary condition. Thus, solutions (u_{k+1}, v_{k+1}) in our algorithm can be found explicitly and easily using the Fourier transform and the block matrix inversion formula.

The ADMM for solving model (12) is summarized in Algorithm 1.

Algorithm 1 ADMM for solving model (12)

- 1: **Given parameters** γ, μ, v, α_1 , and α_0 .
 - 2: **Initialize** $k = 0, u_0 = f, v_0 = \mathbf{0}, \lambda_0^1 = \lambda_0^2 = \lambda_0^3 = \mathbf{0}$.
 - 3: **repeat**
 - 4: $d_{k+1} = \text{shrink} \left(\frac{\lambda_k^2}{\mu} + \nabla u_k - v_k, \frac{\alpha_1}{\mu} \right)$,
 - 5: $w_{k+1} = \text{shrink} \left(\frac{\lambda_k^3}{\mu} + \mathcal{E}(v_k), \frac{\alpha_0}{\mu} \right)$,
 - 6: z_{k+1} is the positive solution of the cubic equation in (21),
 - 7: $(u_{k+1}, v_{k+1}) = (B^T B)^{-1} (-B^T (z_{k+1} - \lambda_k^1/\mu, d_{k+1} - \lambda_k^2/\mu, w_{k+1} - \lambda_k^3/\mu)^T)$,
 - 8: $\lambda_{k+1}^1 = \lambda_k^1 - v\mu(z_{k+1} - u_{k+1})$,
 - 9: $\lambda_{k+1}^2 = \lambda_k^2 - v\mu(d_{k+1} - \nabla u_{k+1} + v_{k+1})$,
 - 10: $\lambda_{k+1}^3 = \lambda_k^3 - v\mu(w_{k+1} - \mathcal{E}(v_{k+1}))$,
 - 11: **until** a stopping condition is satisfied
 - 12: **Final output:** u
-

Remark 1 If $f_{i,j} = 0$, the equality $\sqrt{u_{i,j}} = -n_{i,j}$ has to be satisfied by the degraded model in (1). In this case, the intensity value of the original image u at (i, j) must be the square of the Gaussian noise value $n_{i,j}^2$ and moreover, $n_{i,j}$ must have a negative value. However, this phenomenon happens with a very low probability. Hence, we can assume that $f_{i,j} \neq 0$ for all $(i, j) \in \Omega$ in practice.

3.2 Nonconvex TGV Model

In many prior methods, the nonconvex TV-based models minimize the support of the image gradient, which leads to piecewise-constant restored images. They also yield staircase artifacts near smooth transition regions. Therefore, a nonconvex hybrid TV regularizer [36] has been recently proposed, which is a nonconvex version of a convex combination of the first- and

second-order TV. This efficiently deals with the discontinuous and smooth regions, producing visually remarkable solutions.

Here, we utilize a nonconvex version of TGV_α^2 , since TGV_α^2 automatically balances the first- and second-order derivatives, rather than using any fixed combination. Thus, we introduce a nonconvex version of our TGV model in (12), which employs a nonconvex TGV regularization (NTGV), as follows:

$$\min_{u \in S(\Omega), v} \alpha_1 \int_\Omega \phi_1(|\nabla u - v|) \, dx + \alpha_0 \int_\Omega \phi_2(|\mathcal{E}(v)|) \, dx + \gamma \int_\Omega \frac{(f - u)^2}{u} \, dx, \quad (25)$$

where ϕ_i is a nonconvex log function, such as $\phi_i(s) = \frac{1}{\rho_i} \log(1 + \rho_i s)$ with parameter $\rho_i > 0$. The parameter ρ_1 and ρ_2 control the nonconvexity of the first- and second-order regularization terms. Here, we adopt the nonconvex log function among some possible nonconvex functions. It is hard to solve the minimization problem involving the nonconvex ℓ_q -norm regularizer because finding a limiting-supergradient of $\|\cdot\|_q$ at zero is difficult. Moreover, according to the work in [45], the log function is suitable for the reconstruction of piecewise-smooth images, while the fractional function is more proper for the reconstruction of piecewise-constant images. Real ultrasound images are mostly piecewise-smooth, thus, we utilize the log function as our nonconvex function ϕ_i . This log type of NTGV was introduced in [34, 35], and the only difference with ours is that we adopt different values for the parameters ρ_1 and ρ_2 .

The proposed nonconvex model (25) takes advantages of both higher-order regularization and nonconvex regularization for image denoising. That is, it helps sufficiently denoise smooth regions without staircasing effects while preserving edges and details. Comparing with the convex model (12), the NTGV model (25) is expected to be more suitable for denoising images having much structures and strong edges or in the presence of high level of noise, whereas it is expected to provide similar denoising results with the TGV model for images with weak edges or few discontinuous transitions. On the other hand, the TGV model has more computational efficiency since the NTGV model solves the TGV model in its inner loops, which is explained as follows.

We now present an optimization algorithm to solve the NTGV model in (25), by using the IRLA. First, we consider the discretized version of model (25):

$$\min_{u > 0, v} \alpha_1 \langle \phi_1(|\nabla u - v|), \mathbf{1} \rangle + \alpha_0 \langle \phi_2(|\mathcal{E}(v)|), \mathbf{1} \rangle + \gamma D(u). \quad (26)$$

Using the variable splitting technique, we can reformulate the unconstrained model in (26) into the following constrained problem as

$$\begin{aligned} \min_{u, v, z > 0, d, w} \quad & \alpha_1 \langle \phi_1(|d|), \mathbf{1} \rangle + \alpha_0 \langle \phi_2(|w|), \mathbf{1} \rangle + \gamma D(z), \\ \text{subject to} \quad & z = u, d = \nabla u - v \text{ and } w = \mathcal{E}(v). \end{aligned} \quad (27)$$

By letting

$$\mathbf{x} = \begin{bmatrix} u \\ v \\ z \\ d \\ w \end{bmatrix}, \quad A = \begin{bmatrix} -I & \mathbf{0} & I & \mathbf{0} & \mathbf{0} \\ -\nabla & I & \mathbf{0} & I & \mathbf{0} \\ \mathbf{0} & -\mathcal{E} & \mathbf{0} & I & \mathbf{0} \end{bmatrix}, \quad b = \mathbf{0},$$

$$E_1(\mathbf{x}) = \gamma \left\langle \frac{(f-z)^2}{z}, \mathbf{1} \right\rangle, \quad E_2(\mathbf{x}) = \alpha_1 \langle \phi_1(|d|), \mathbf{1} \rangle + \alpha_0 \langle \phi_2(|w|), \mathbf{1} \rangle,$$

the constrained model in (27) can be reformulated as the form in (10).

Hence, the IRLA can be applied to problem (27), which is given by

$$\begin{bmatrix} u_{k+1} \\ v_{k+1} \\ z_{k+1} \\ d_{k+1} \\ w_{k+1} \end{bmatrix} = \arg \min_{\mathbf{Ax}=b} \left\{ \alpha_1 \langle \tilde{d}_k, |d| \rangle + \alpha_0 \langle \tilde{w}_k, |w| \rangle + \gamma D(z) \right\} \tag{28}$$

with $\tilde{d}_k := \frac{1}{\rho_1 |d_k| + 1}$ and $\tilde{w}_k := \frac{1}{\rho_2 |w_k| + 1}$.

The IRLA (28) leads to another convex minimization problem in its inner iteration. It has the same form as the convex TGV model in (16) except for the weighted functions \tilde{d}_k and \tilde{w}_k . Hence, the subproblem in (28) can be solved by the ADMM. That is, by changing the setting in (17) using $H(\mathbf{x}) = \alpha_1 \langle \tilde{d}_k, |d| \rangle + \alpha_0 \langle \tilde{w}_k, |w| \rangle + \gamma D(z)$, we can apply the ADMM to the subproblem in (28), which yields

$$\left\{ \begin{array}{l} d_{k+1} = \arg \min_d \alpha_1 \langle \tilde{d}_k, |d| \rangle - (\lambda_k^2)^T(d) + \frac{\mu}{2} \|d - \nabla u_k + v_k\|_2^2, \\ w_{k+1} = \arg \min_w \alpha_0 \langle \tilde{w}_k, |w| \rangle - (\lambda_k^3)^T(w) + \frac{\mu}{2} \|w - \mathcal{E}(v_k)\|_2^2, \\ z_{k+1} = \arg \min_{z \geq 0} \gamma D(z) - (\lambda_k^1)^T(z) + \frac{\mu}{2} \|z - u_k\|_2^2, \\ \begin{pmatrix} u_{k+1} \\ v_{k+1} \end{pmatrix} = \arg \min_{u,v} (\lambda_k^2)^T(\nabla u - v) + (\lambda_k^3)^T(\mathcal{E}(v)) + \lambda_k^1{}^T u \\ \quad + \frac{\mu}{2} \|\nabla u - v - d_{k+1}\|_2^2 + \frac{\mu}{2} \|\mathcal{E}(v) - w_{k+1}\|_2^2 + \frac{\mu}{2} \|u - z_{k+1}\|_2^2, \\ \lambda_{k+1}^1 = \lambda_k^1 - v\mu(z_{k+1} - u_{k+1}), \\ \lambda_{k+1}^2 = \lambda_k^2 - v\mu(d_{k+1} - \nabla u_{k+1} + v_{k+1}), \\ \lambda_{k+1}^3 = \lambda_k^3 - v\mu(w_{k+1} - \mathcal{E}(v_{k+1})). \end{array} \right. \tag{29}$$

The subproblems for d_{k+1} and w_{k+1} in (29) can be computed explicitly with the shrink operator as follows:

$$d_{k+1} = \text{shrink} \left(\frac{\lambda_k^2}{\mu} + \nabla u_k - v_k, \frac{\alpha_1 \tilde{d}_k}{\mu} \right), \quad w_{k+1} = \text{shrink} \left(\frac{\lambda_k^3}{\mu} + \mathcal{E}(v_k), \frac{\alpha_0 \tilde{w}_k}{\mu} \right).$$

The other solutions can be obtained as described in Sect. 3.1.

Overall, the optimization algorithm for solving our NTGV model in (25) is summarized in Algorithm 2.

Lastly, we can obtain a partial convergence of the IRLA in (28) for solving the constrained problem in (27). In other words, it is trivial to show that the objective function in problem (27) is coercive in $z \in S(\Omega)$. Therefore, from Theorem 2, we can obtain a partial convergence of the IRLA in (28) when it is applied to problem (27): *The sequence $(u_k, v_k, z_k, d_k, w_k)$ generated by (28) is bounded, and it has at least one accumulation point. That is, there exists a converging subsequence to an accumulation point.* However, the nonconvex log function in (27) is not a sum of a convex function; thus, we cannot even assure the local convergence of the algorithm in (28) (see [34] for more details).

Algorithm 2 IRLA with ADMM for solving model (25)

- 1: **Given parameters** $\gamma, \mu, v, \alpha_1, \alpha_0, \rho_1,$ and ρ_2 .
- 2: **Initialize** $\ell = 0, u_0 = f, v_0 = \mathbf{0}, d_j = \nabla u, w_j = \mathbf{0},$ and $\lambda_0^1 = \lambda_0^2 = \lambda_0^3 = \mathbf{0}$.
- 3: **repeat**
- 4: $\tilde{d}_\ell := \frac{1}{\rho_1|d_\ell|+1}$
- 5: $\tilde{w}_\ell := \frac{1}{\rho_2|w_\ell|+1}$
- 6: **for** $k = 1$ to M_{in} **do**
- 7: $d_{\ell,k+1} = \text{shrink}\left(\frac{\lambda_{\ell,k}^2}{\mu} + \nabla u_{\ell,k} - v_{\ell,k}, \frac{\alpha_1 \tilde{d}_\ell}{\mu}\right),$
- 8: $w_{\ell,k+1} = \text{shrink}\left(\frac{\lambda_{\ell,k}^3}{\mu} + \mathcal{E}(v_{\ell,k}), \frac{\alpha_0 \tilde{w}_\ell}{\mu}\right),$
- 9: $z_{\ell,k+1}$ is the positive solution of the cubic equation in (21),
- 10: $\begin{pmatrix} u_{\ell,k+1} \\ v_{\ell,k+1} \end{pmatrix} = (B^T B)^{-1} \left(-B^T \left(\frac{\lambda_{\ell,k}^1}{\mu} - z_{\ell,k+1}, \frac{\lambda_{\ell,k}^2}{\mu} - d_{\ell,k+1}, \frac{\lambda_{\ell,k}^3}{\mu} - w_{\ell,k+1} \right)^T \right),$
- 11: $\lambda_{\ell,k+1}^1 = \lambda_{\ell,k}^1 - v\mu(z_{\ell,k+1} - u_{\ell,k+1}),$
- 12: $\lambda_{\ell,k+1}^2 = \lambda_{\ell,k}^2 - v\mu(d_{\ell,k+1} - \nabla u_{\ell,k+1} + v_{\ell,k+1}),$
- 13: $\lambda_{\ell,k+1}^3 = \lambda_{\ell,k}^3 - v\mu(w_{\ell,k+1} - \mathcal{E}(v_{\ell,k+1})),$
- 14: **end for**
- 15: **until** a stopping condition is satisfied
- 16: **Final output:** u

4 Numerical Experiments

In this section, we present numerical results for our proposed models and compare them with existing variational models such as the TV model [21] and the WBSN model [3]. We test using both natural images and real ultrasound images. All experiments were implemented using MATLAB R2015b on a desktop PC with an Intel CPU at 3.50GHz, 16GB RAM, and a 64-bit Window 10 operating system.

In [21], the authors solved the TV model by using the gradient-descent method. However, this model can be easily implemented by the split-Bregman method, similar to our approach. For fair comparison with our models, we use the split-Bregman method to solve the TV model. The stopping criterion for our TGV/NTGV model and the TV model is given by

$$\frac{\|u_k - u_{k-1}\|_2}{\|u_k\|_2} < 10^{-4} \quad \text{or} \quad k > M, \tag{30}$$

where M is the maximum iteration number; $M = 500$ for our TGV model and the TV model, $M = 100$ for our NTGV model. Besides, the maximum iteration number of the ADMM for the inner subproblem in the NTGV model is fixed as $M_{\text{in}} = 10$. The WBSN model is solved by the gradient-descent based method, so the maximum iteration number M is set as 5000.

Recall that the data fidelity term in (2) is not derived from the MAP estimation based on the degradation model. These result in denoised images with shifted intensity values. That is, variational models using (2) as the data fidelity term are likely to produce a big gap between the mean of a restored image and that of an original clean image. To reduce this gap, we add the bias correction step in the final restored image. Since $\text{mean}(n) = 0$, it follows from the degradation model in (1) that

$$\begin{aligned} \text{mean}(f) &= \text{mean}(u) + \text{mean}(\sqrt{u} \cdot n) \\ &\approx \text{mean}(u) + \text{mean}(\sqrt{u})\text{mean}(n) = \text{mean}(u). \end{aligned}$$

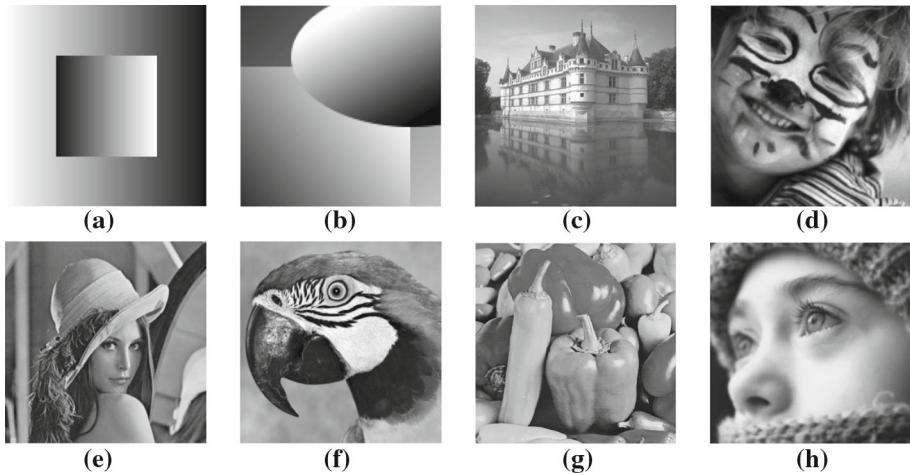


Fig. 1 Original images. **a** Synthetic1 (256 × 256), **b** Synthetic2 (253 × 253), **c** Castle (321 × 481), **d** Child (255 × 255), **e** Lena (256 × 256), **f** Parrot (256 × 256), **g** Pepper (512 × 512), **h** Woman (256 × 254)

Table 1 Comparison of PD algorithm [11] versus ADMM applied to solve model (12)

Model	Computing time/Energy value/Iteration number								
	$\sigma = 2$			$\sigma = 4$			$\sigma = 6$		
PD	5.09	4.9655	295	7.73	10.7610	468	11.00	14.9701	680
ADMM	1.26	4.8042	33	1.42	10.6660	39	1.52	14.8293	40

Best algorithmic performance are given in bold

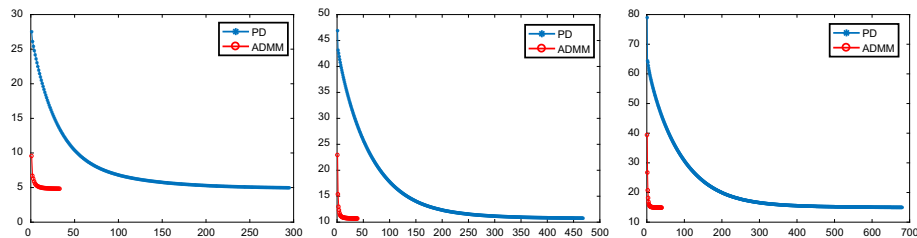


Fig. 2 Plot of energy functional $E(u_k)$ in (12) via iteration number k when using PD algorithm [11] (blue) and ADMM (red), when noise level (1st column) $\sigma = 2$, (2nd) $\sigma = 4$, (3rd) $\sigma = 6$ (Color figure online)

From this approximation and the constraint $u > 0$, the rescaled image \tilde{u} is obtained from the denoised image u as follows:

$$\tilde{u} = \max\{u + \text{mean}(f) - \text{mean}(u), 0\}.$$

We then regard the image \tilde{u} as the final restored image for all models.

The selection of parameters is as follows. First, we select α_0 and α_1 to satisfy $\alpha_0 + \alpha_1 = 1$. We tune the parameter α_1 , which controls the balance between first- and second-order terms. For real ultrasound images, α_1 is fixed as 0.4 or 0.6. The parameter γ is selected depending on the noise level σ ; if the noise level σ is large, a small value of γ is more proper, and vice

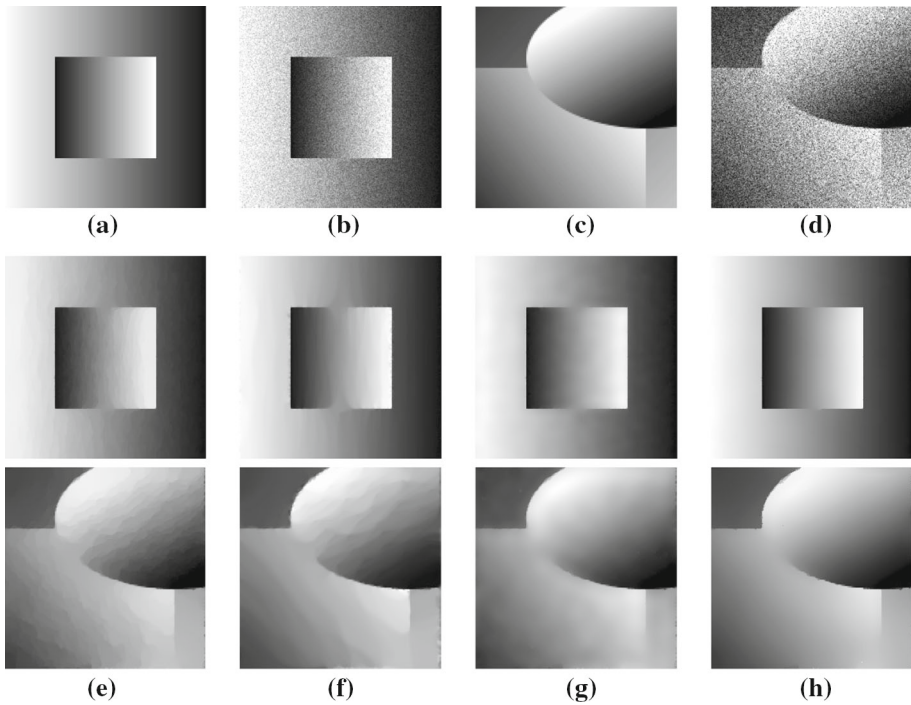


Fig. 3 Denoising results of our proposed models and other models on synthetic images. **a, c** Original images, **b, d** Noisy images (**b** $\sigma = 2$, **d** $\sigma = 4$), **e** TV model [21], **f** WBSN model [3], **g** TGV model (12), **h** NTGV model (25)

versa. Hence, we mainly tune the parameters α_1 and γ to achieve the best restored images. The parameter μ in the ADMM affects on the speed of the algorithm; we set $\mu = 0.01$ for our TGV and the TV models, and $\mu = 0.05$ for our NTGV model. The values for ρ_1 and ρ_2 in the NTGV model are selected in the ranges $[0.001, 0.005]$ and $[0.05, 0.2]$, respectively, for natural images. Since real ultrasound images in general have blurry edges, we tune the parameter ρ_1 in the range $[0.01, 0.05]$.

4.1 Denoising Results of Synthetic and Natural Images

In the experiments, all clean images have intensity values in the range $[0, 255]$. In this numerical test, we used two synthetic images and six natural images, as shown in Fig. 1. The test images were corrupted using (1) as the degradation model, with $\sigma = 2, 4$, and 6 .

In order to measure the quality of the restored images, we compute the peak signal-to-noise ratio (PSNR) value and the structure similarity (SSIM) index, which are defined as follows:

$$PSNR(u_*, \bar{u}) = 10 \log_{10} \left(\frac{255^2 mn}{\|u_* - \bar{u}\|_2^2} \right),$$

$$SSIM(u_*, \bar{u}) = \frac{(2\mu\bar{u}\mu_{u_*} + c_1)(\sigma_{\bar{u}u_*} + c_2)}{(\mu_u^2 + \mu_{u_*}^2 + c_1)(\sigma_u^2 + \sigma_{u_*}^2 + c_2)},$$

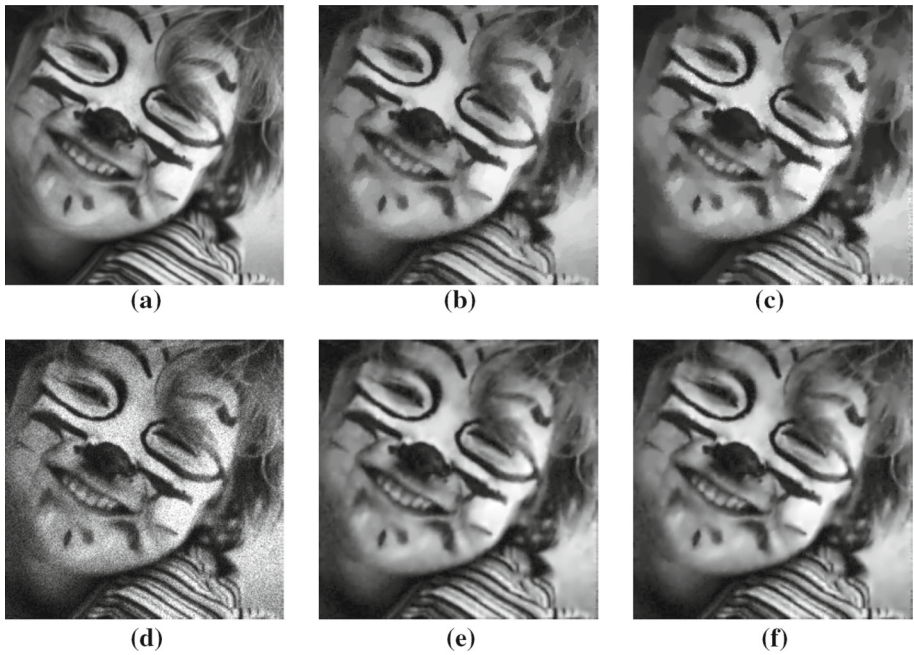


Fig. 4 Denoising results of our proposed models and other models on a natural image when $\sigma = 2$. **a** Original image, **b** TV model [21], **c** WBSN model [3], **d** Noisy image, **e** TGV model (12), **f** NTGV model (25)

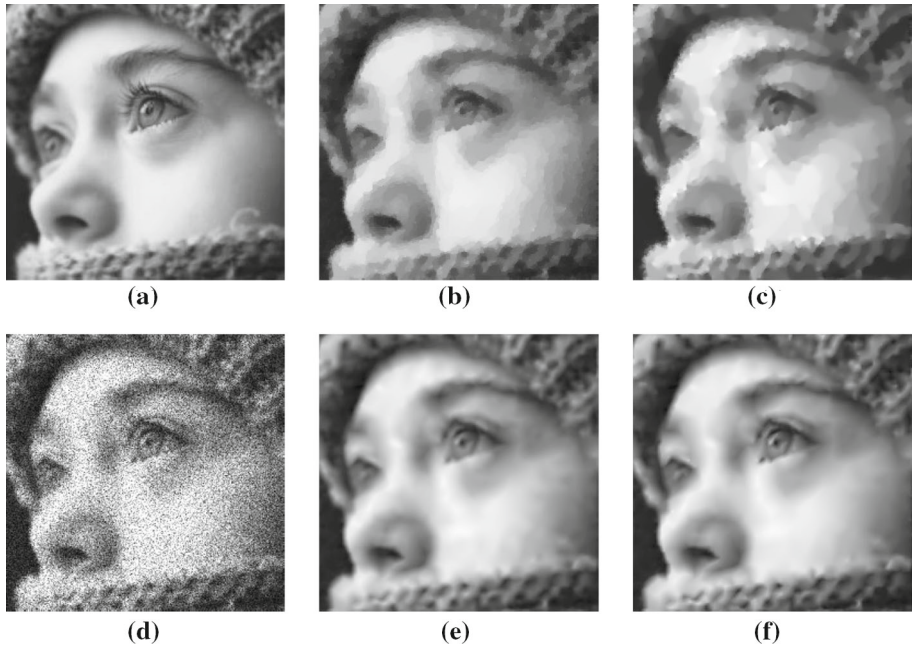


Fig. 5 Denoising results of our proposed models and other models on a natural image when $\sigma = 4$. **a** Original image, **b** TV model [21], **c** WBSN model [3], **d** Noisy image, **e** TGV model (12), **f** NTGV model (25)

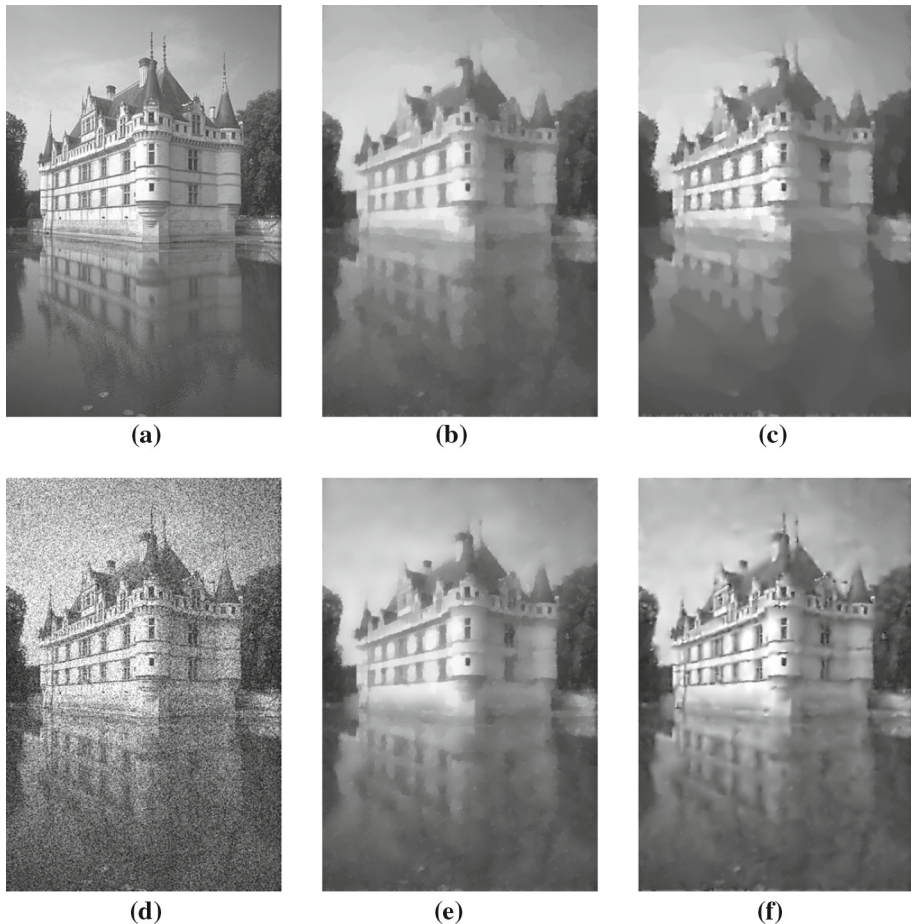


Fig. 6 Denoising results of our proposed models and other models on a natural image when $\sigma = 4$. **a** Original image, **b** TV model [21], **c** WBSN model [3], **d** Noisy image, **e** TGV model (12), **f** NTGV model (25)

where $u_* \in \mathbb{R}^{m \times n}$ is the clean image, $\bar{u} \in \mathbb{R}^{m \times n}$ is the restored image, μ_u is the average of u , σ_u is the standard deviation of u , and c_1 and c_2 are some constants for stability. For all methods, we determined the tuning parameters that obtained the best restored images visually as well as using the PSNR and SSIM values.

Before discussing denoising results, we first compare the ADMM and the primal-dual algorithm [11] that is used for solving the TGV-based denoising model in [6]. We apply both algorithms to our TGV model (12), and the same regularization parameters (α_1, γ) and stopping criterion in (30) are used. Table 1 and Fig. 2 present the denoising results of the “child” image in the presence of noise with $\sigma = 2, 4, 6$. In Table 1, we present the computing time, final energy functional values, and total iteration numbers, and Fig. 2 shows the plots of the energy functional in (12) via iteration number k . It can be seen that the ADMM reaches the stopping criterion much faster and provides lower final energy functional values than the PD algorithm, which implies the ADMM finds more accurate solutions within shorter time. These show the efficiency of the ADMM for solving our TGV model, compared with the PD algorithm.

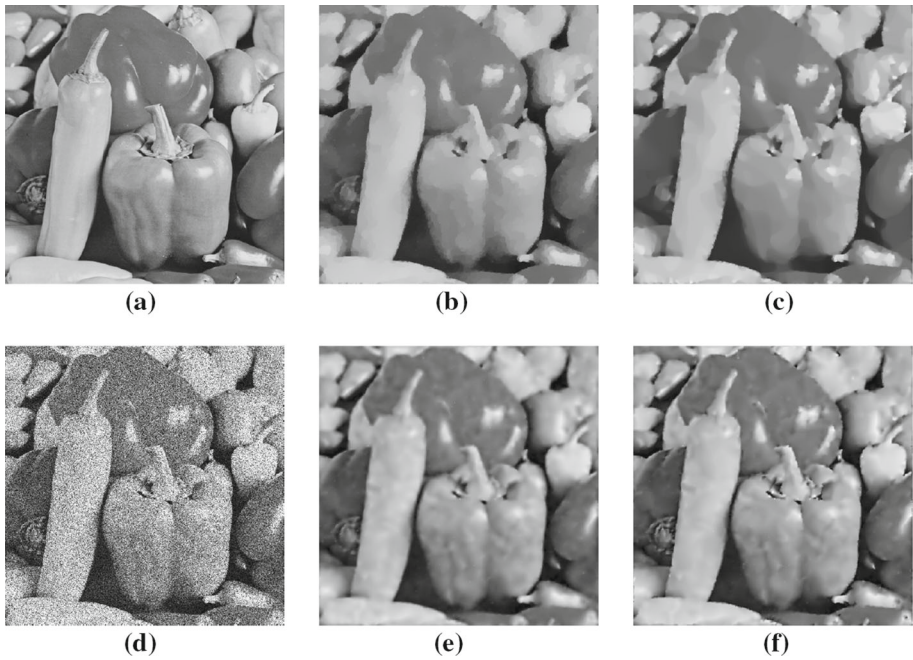


Fig. 7 Denoising results of our proposed models and other models on a natural image when $\sigma = 6$. **a** Original image, **b** TV model [21], **c** WBSN model [3], **d** Noisy image, **e** TGV model (12), **f** NTGV model (25)

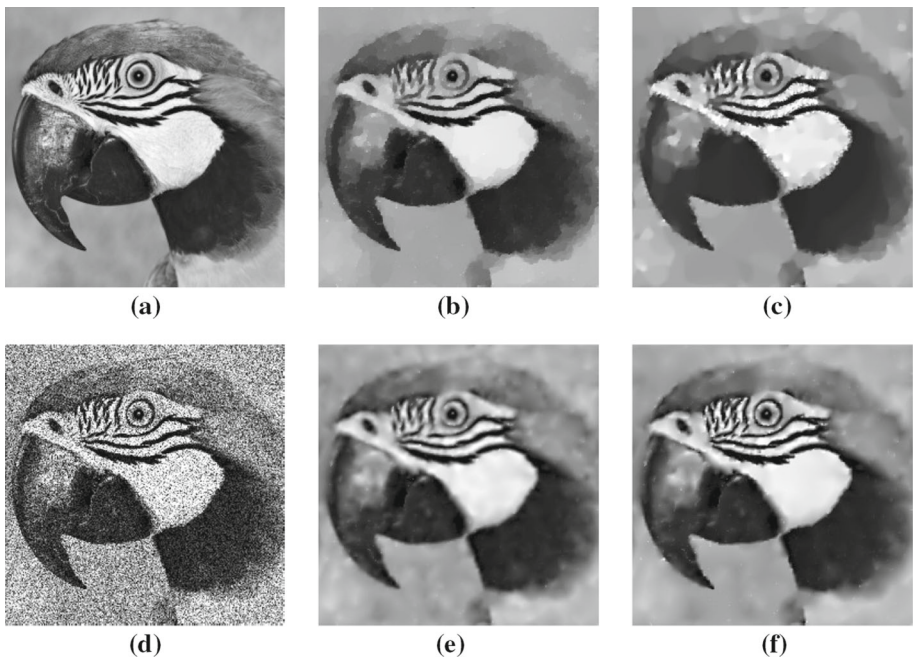


Fig. 8 Denoising results of our proposed models and other models on a natural image when $\sigma = 6$. **a** Original image, **b** TV model [21], **c** WBSN model [3], **d** Noisy image, **e** TGV model (12), **f** NTGV model (25)

Table 2 Comparison results for PSNR and SSIM

Model	TV	WBSN	TGV	NTGV
Image	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR /SSIM
Noise level	$\sigma = 2$			
castle	27.26/0.8258	27.21/0.8237	27.49/0.8280	28.31/0.8445
face	29.38/0.8818	28.46/0.8683	30.97/0.9210	31.28/0.9222
lena	28.64/0.8426	27.87/0.8107	29.28/0.8598	29.94/0.8731
parrot	28.34/0.8418	27.38/0.8228	28.48/0.8433	29.05/0.8600
pepper	30.40/0.8179	29.15/0.8056	30.82/0.8261	31.32/0.8337
woman	30.96/0.8877	30.62/0.8847	33.24/0.9388	33.56/0.9401
syn2	34.20/0.9633	36.45/0.9802	36.62/0.9822	42.35/0.9924
syn1	34.26/0.9646	34.27/0.9742	36.48/0.9780	38.56/0.9861
Noise level	$\sigma = 4$			
castle	24.02/0.7318	24.16/0.7303	24.21/0.7350	24.95/0.7631
face	25.94/0.8059	25.28/0.7732	27.17/0.8571	27.70/0.8624
lena	26.03/0.7699	25.25/0.7333	26.26/0.7806	26.72/0.8021
parrot	24.20/0.7676	24.24/0.7516	24.61/0.7744	25.80/0.7964
pepper	27.29/0.7702	27.26/0.7603	27.51/0.7812	28.43/0.7925
woman	28.00/0.8203	27.24/0.8095	29.63/0.8947	29.96/0.8983
syn2	31.78/0.9470	32.07/0.9637	32.48/0.9722	36.09/0.9864
syn1	31.80/0.9369	31.96/0.9597	32.49/0.9650	32.78/0.9738
Noise level	$\sigma = 6$			
castle	23.08/0.6905	23.29/0.7036	23.18/0.6865	23.65/0.7147
face	24.21/0.7444	23.52/0.7159	25.44/0.8063	25.70/0.8117
lena	23.64/0.6973	23.76/0.6865	24.50/0.7248	24.94/0.7471
parrot	23.18/0.7305	22.68/0.6981	23.17/0.7263	23.93/0.7475
pepper	25.99/0.7428	25.74/0.7353	26.12/0.7506	26.71/0.7622
woman	26.31/0.7741	25.30/0.7430	27.93/0.8615	28.19/0.8657
syn2	29.61/0.9355	28.96/0.9380	30.04/0.9571	31.69/0.9786
syn1	29.52/0.9218	30.09/0.9387	30.27/0.9535	31.36/0.9656

Best denoising performance are given in bold

In Figs. 3, 4, 5, 6, 7, and 8, we display the original images, noisy images, and restored images of all models, tested on synthetic and natural images. Figure 3 presents the denoising results when the noise level $\sigma = 2$ and 4, respectively, tested on piecewise smooth synthetic images. Figures 4, 5, 6, 7, and 8 exhibit the results when $\sigma = 2, 4, 6$, respectively, tested on natural images.

Throughout the examples, we can observe that our proposed models alleviate the stair-casing effects often found in piecewise smooth transition areas in the restored images by the TV and WBSN models. Hence, the restored images of our models look more like natural images. For instance, in Figs. 4 and 5, the restored images of the TV and WBSN models have staircasing artifacts around the chin of the child and woman, while our TGV models conserve the piecewise smooth transition regions very well. This demonstrates the superiority of the TGV-based regularizers over the first-order derivative based regularizers for denoising piecewise smooth images. In addition, the WBSN and NTGV models commonly recover the

Table 3 Comparison results for computing time

Model	TV	WBSN	TGV	NTGV	TV	WBSN	TGV	NTGV
Noise level	$\sigma = 2$				$\sigma = 4$			
castle	2.27	71.42	8.46	37.40	4.24	64.92	13.77	34.62
face	0.69	34.51	1.63	3.93	1.06	29.07	1.76	6.65
lena	0.90	33.12	1.74	10.47	1.34	30.67	2.98	10.63
parrot	0.67	35.72	2.93	9.31	1.41	31.94	1.86	10.41
pepper	3.98	111.43	9.82	30.84	6.57	112.95	11.84	49.38
woman	0.82	35.76	1.27	2.45	1.16	30.56	1.36	3.46
syn2	1.77	36.10	2.49	4.84	2.26	31.13	3.87	8.69
syn1	2.00	35.95	2.89	3.44	2.41	31.88	4.49	10.70
Model	TV		WBSN		TGV		NTGV	
Noise level	$\sigma = 6$							
castle	4.94		62.56		12.80		27.47	
face	1.36		30.65		1.76		6.32	
lena	1.97		27.27		4.22		9.47	
parrot	1.44		25.22		8.06		12.08	
pepper	8.37		105.59		12.67		31.23	
woman	1.52		30.55		1.82		5.47	
syn2	2.99		32.70		5.05		15.97	
syn1	3.33		32.67		15.55		11.58	

denoised images with well preserved edges. However, because the WBSN uses the Tikhonov regularization $\int_{\Omega} |\nabla u|^2 dx$, it smoothes out small features or weak edges in the restored images. Specifically, in Fig. 6, the castle’s reflection in the lake is barely visible because of the oversmoothing. In contrast, the images recovered by our NTGV model include the small details.

Comparing our two models, we can see that the NTGV model provides sharper edges and more preserved fine features than the TGV model, particularly in the presence of strong edges or high level of noise, as seen in Figs. 3, 6, 7, and 8. This shows the merit of the nonconvex regularization. On the other hand, in Figs. 4 and 5, the restored images of our models are visually similar. The NTGV model produces slightly better preserved edges and textures than the TGV model, which can be seen in the regions of the textural patterns in child, and the eyes and the woolen hat in woman. Hence, we can presume that both models produce similar denoising results for the images having smooth regions with weak edges.

Lastly, in Table 2, we present the PSNR values and SSIM values of all models for noise levels $\sigma = 2, 4$, and 6. We also observe from Table 2 that our models obtain restored images that are superior to those obtained by the others. Specifically, the NTGV model outperforms the others when $\sigma = 6$. In terms of the computing time in Table 3, the TV model is the fastest for all cases. In contrast, the WBSN is the slowest because it is solved using the gradient descent method. The NTGV requires much more time than the TGV model because of the additional outer iterations for the IRLA. Thus, the TGV model has an advantage with respect to computing time, while the NTGV model recovers the best quality images.

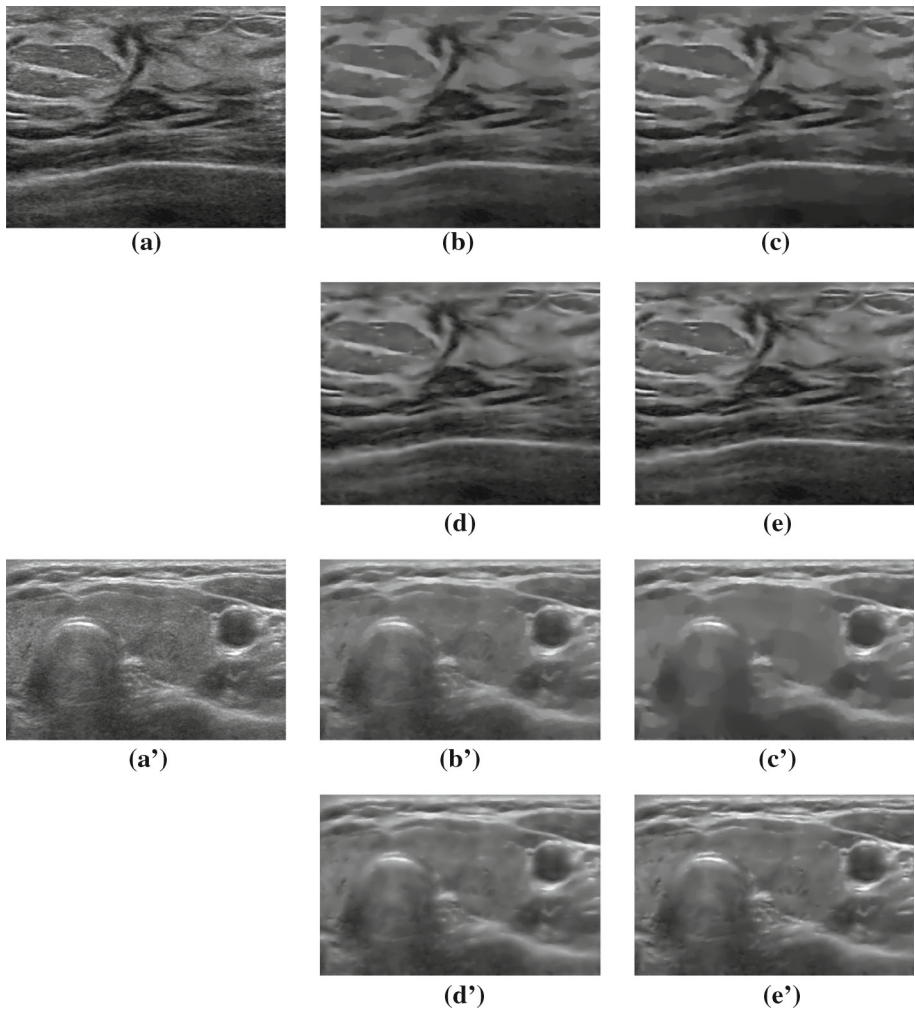


Fig. 9 Denoising results of our proposed models and other models on breast and thyroid images. **a–a'** Noisy image, **b–b'** TV model [21], **c–c'** WBSN model [3], **d–d'** TGV model (12), **e–e'** NTGV model (25)

4.2 Denoising Results of Real Ultrasound Images

In Figs. 9, 10, 11, and 12, we present the denoising results of all models, tested on several real ultrasound images. The TV and the WBSN models give rise to staircasing artifacts, leading to a more blocky, less authentic denoised image. But our models avoid these artifacts and achieve more natural homogenous regions than both state-of-the-arts models. We can also see that the restored images of our models have smooth interiors and clear boundaries, while the others create somewhat rough boundaries. In other words, our proposed models perform better than the other ones.

Moreover, the WBSN and NTGV models find clear corners and features, thanks to their edge-preserving techniques in common. However, the restored images of the WBSN model look somewhat like cartoon images, while the NTGV model produces clearer images with preserved edges. Specifically, the slim vessel in the upper denoised image of our NTGV

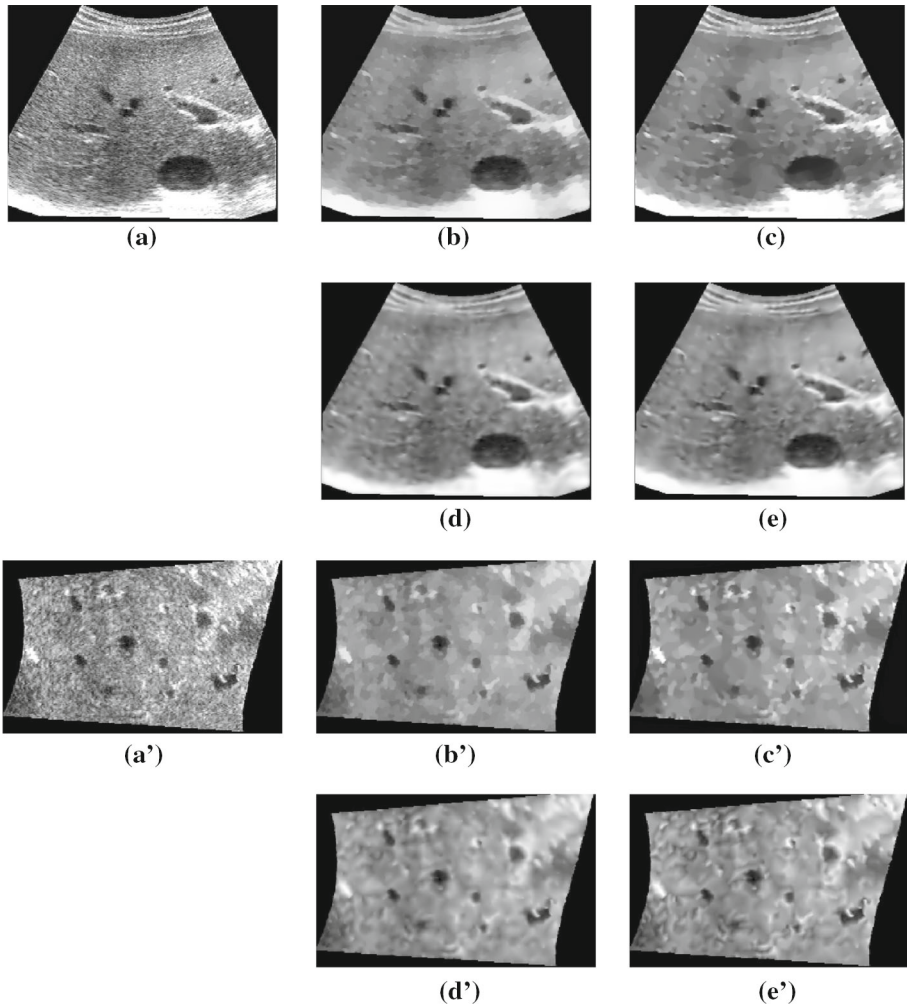


Fig. 10 Denoising results of our proposed models and other models on real ultrasound images. **a–a'** Noisy image, **b–b'** TV model [21], **c–c'** WBSN model [3], **d–d'** TGV model (12), **e–e'** NTGV model (25)

model in Fig. 10 is sleek, with no bumps. Although our two models result in visually similar recovered images, the NTGV model provides slightly clearer slim vessel and edges than the TGV model in an enlarged view. Therefore, these examples also qualitatively verify the superior performance of our models for real ultrasound image denoising.

Lastly, as discussed in Sect. 3.2, we can observe from the numerical results that the NTGV model (25) performs better than the TGV model (12) when denoising images having much structures or strong edges, such as “castle”, “pepper” images, or images with high level of noise. On the other hand, the NTGV model provides similar denoising results with the TGV model for images with weak edges, such as “child”, “woman” images or real ultrasound images. Moreover, the TGV model has a much benefit of computing time, as shown in Table 3. To sum up, these numerical experiments validate the effectiveness of the TGV regularization and superiority of the nonconvex regularization.

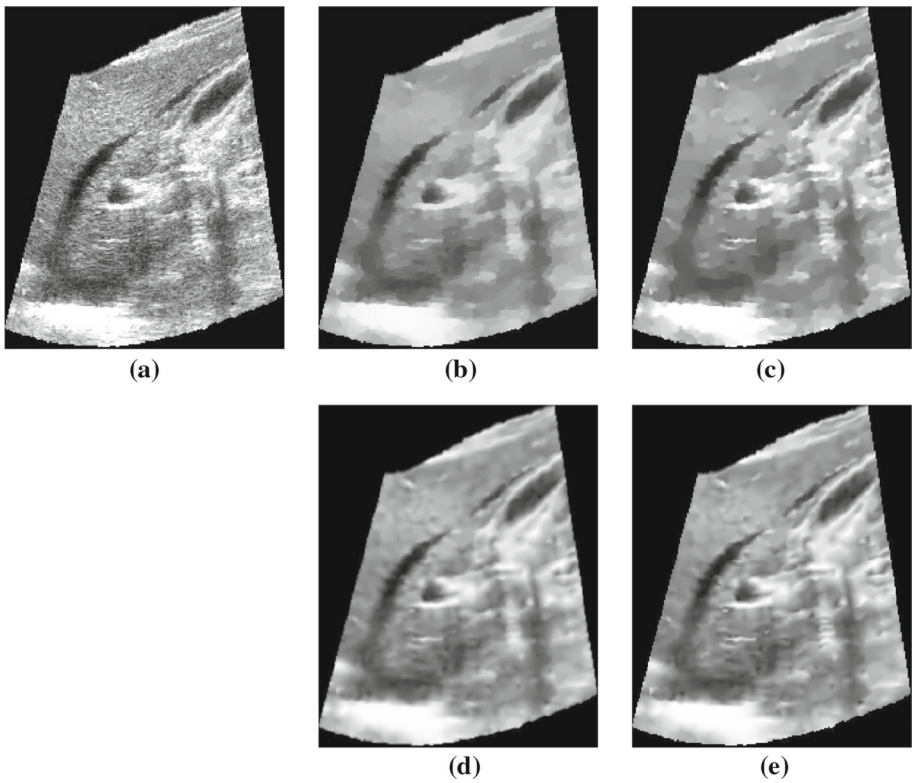


Fig. 11 Denoising results of our proposed models and other models on a real ultrasound image. **a** Noisy image, **b** TV model [21], **c** WBSN model [3], **d** TGV model (12), **e** NTGV model (25)

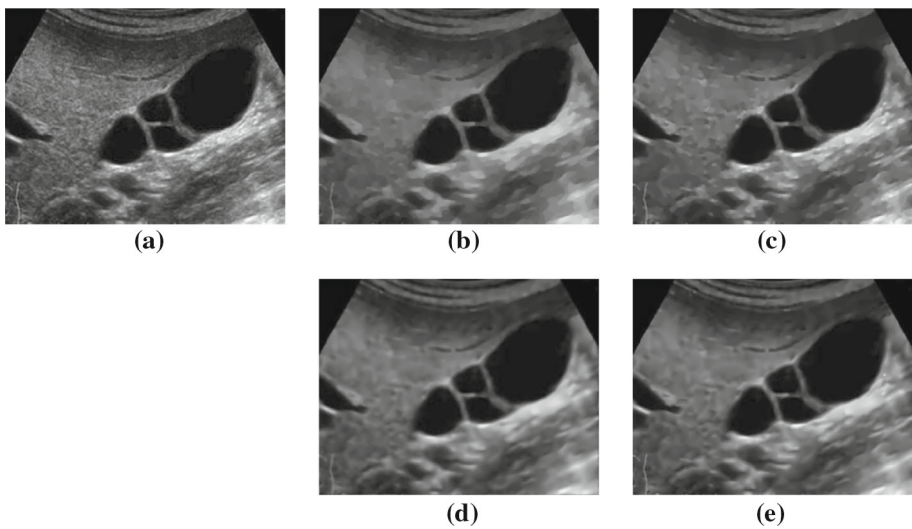


Fig. 12 Denoising results of our proposed models and other models on a real ultrasound image. **a** Noisy image, **b** TV model [21], **c** WBSN model [3], **d** TGV model (12), **e** NTGV model (25)

5 Conclusion

In this paper, we proposed two TGV-based minimization models for denoising ultrasound images. Our models consist of a convex data-fidelity term and the convex or nonconvex TGV regularization terms. The proposed models eliminate the staircasing artifacts commonly appeared in the results of TV-based models. Moreover, the nonconvex TGV model is able to produce better preserved edges, textures, and fine features in the restored images. We proved the existence and uniqueness of the minimizers for the convex TGV model. We also provided efficient iterative algorithms by adopting the ADMM to handle the complicated data fidelity term and nonsmooth convex regularization term, and the IRLA to deal with the nonconvex regularization term. Compared with state-of-the-art methods, the proposed models perform effectively by smoothing homogeneous regions while preserving edges, and thus they provide better denoising results for both natural images and real ultrasound images. Real ultrasound images are usually blurry as well as noisy. Thus, blind deconvolution along with denoising for ultrasound images is a possible target for future work.

Acknowledgements Myeongmin Kang was supported by the NRF (2016R1C1B1009808). Myungjoo Kang was supported by the NRF (2014R1A2A1A10050531, 2015R1A5A1009350), MOTIE (10048720) and IITP-MSIP (B0717-16-0107). Miyoun Jung was supported by the Hankuk University of Foreign Studies Research Fund and NRF (2013R1A1A3010416).

Appendix: The Proof of Theorem 3

Proof For any fixed $x \in \Omega$, we can deduce from the inequality of arithmetic-geometric mean that for $u > 0$,

$$\frac{(f(x) - u)^2}{u} = u + \frac{f(x)^2}{u} - 2f(x) \geq 2|f(x)| - 2f(x).$$

It follows from the above inequality that

$$E(u) \geq \gamma \int_{\Omega} \frac{(f - u)^2}{u} dx \geq \gamma \int_{\Omega} 2|f| - 2f dx,$$

i.e., $E(u)$ is bounded below. Hence, we can choose a minimizing sequence $\{u^n\} \in L^2(\Omega)$ for problem (12), and the sequence $\{TGV_{\alpha}^2(u^n)\}$ is also bounded.

In addition, from the Poincaré inequality of TGV_{α}^2 , we can obtain that

$$\|u^n - Pu^n\|_2 \leq C \cdot TGV_{\alpha}^2(u^n), \quad \text{for some constant } C,$$

where $P : L^2(\Omega) \rightarrow \ker(TGV_{\alpha}^2)$ is a linear projection. It follows that $u^n - Pu^n$ is bounded in $L^2(\Omega)$ for each n .

We can also easily deduce that for $u > 0$,

$$\begin{aligned} \int_{\Omega} \frac{(f - u)^2}{u} dx &= \int_{\Omega} u + \frac{f^2}{u} - 2f dx \geq \|u\|_1 - \int_{\Omega} 2f dx \\ &\geq c \cdot \|Pu\|_1 - \int_{\Omega} 2f dx \geq c \cdot \|Pu\|_2 - \int_{\Omega} 2f dx, \end{aligned}$$

where c is a positive constant that is independent of n . Since $\{u^n\}$ is a minimizing sequence of problem (12), $\|Pu^n\|_2$ is bounded from the above inequalities.

Then, from the following inequalities,

$$\|u^n\|_2 \leq \|u^n - Pu^n + Pu^n\|_2 \leq \|u^n - Pu^n\|_2 + \|Pu^n\|_2,$$

we can conclude that $\{u^n\}$ is bounded in $L^2(\Omega)$. Therefore, there exist a subsequence $\{u^{n_k}\}$ and $u^* \in L^2(\Omega)$ such that the subsequence converges weakly to u^* . Since $u^n > 0$ and $E(u^*)$ must be bounded, $u^* > 0$ and $u^* \in S(\Omega)$.

Because TGV_α^2 is lower semicontinuous, $E(u)$ is also lower semicontinuous. By applying Fatou's Lemma, we have

$$E(u^*) \leq \liminf_{n \rightarrow \infty} E(u^n) = \inf_{u \in S(\Omega)} E(u),$$

which implies that u^* is a solution of problem (12).

If we let $p(t) = \frac{(f(x)-t)^2}{t}$ for any fixed $x \in \Omega$, then its second derivative is given by $p''(t) = \frac{2f(x)^2}{t^3}$ and thus $p''(t) > 0$ for any $t > 0$. Hence, p is a strictly convex function when $t > 0$. Therefore, problem (12) has a unique minimizer. \square

References

1. Aubert, G., Aujol, J.F.: A variational approach to removing multiplicative noise. *SIAM J. Appl. Math.* **68**(4), 925–946 (2008)
2. Aubert, G., Kornprobst, P.: *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, vol. 147. Springer, Berlin (2006)
3. Barcelos, C.A., Vieira, L.E.: Ultrasound speckle noise reduction via an adaptive edge-controlled variational method. In: 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 145–151 (2014)
4. Besag, J.: Digital image processing: towards Bayesian image analysis. *J. Appl. Stat.* **16**(3), 395–407 (1989)
5. Bredies, K., Holler, M.: A TGV regularized wavelet based zooming model. *Scale Space Var. Methods Comput. Vis.* **7893**, 149–160 (2013)
6. Bredies, K., Kunisch, K., Pock, T.: Total generalized variation. *SIAM J. Imaging Sci.* **3**(3), 492–526 (2010)
7. Bredies, K., Kunisch, K., Valkonen, T.: Properties of L^1 - TGV^2 : The one-dimensional case. *J. Math. Anal. Appl.* **389**(1), 438–454 (2013)
8. Candés, E.J., Wakin, M.B., Boyd, S.P.: Enhancing sparsity by reweighted ℓ_1 minimization. *J. Fourier Anal. Appl.* **14**, 877–905 (2008)
9. Chambolle, A.: An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.* **20**, 89–97 (2004)
10. Chambolle, A., Lions, P.: Image recovery via total variation minimization and related problems. *Numer. Math.* **76**(2), 167–188 (1997)
11. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **40**(1), 120–145 (2011)
12. Chan, T., Marquina, A., Mulet, P.: High-order total variation-based image restoration. *SIAM J. Sci. Comput.* **22**(2), 503–516 (2000)
13. Deng, W., Yin, W.: On the global and linear convergence of the generalized alternating direction method of multipliers. *J. Sci. Comput.* **66**, 889–916 (2016)
14. Douglas, J., Rachford, H.H.: On the numerical solution of heat conduction problems in two and three space variables. *Trans. Am. Math. Soc.* **82**(2), 421–439 (1956)
15. Eckstein, J., Bertsekas, D.P.: On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.* **55**, 293–318 (1992)
16. Geman, D., Yang, C.: Nonlinear image recovery with half-quadratic regularization. *IEEE Trans. Image Process.* **4**(7), 932–946 (1995)
17. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell. PAMI* **6**(6), 721–741 (1984)
18. Getreuer, P.: Total variation inpainting using split Bregman. *Image Process. On Line* **2**, 147–157 (2012)

19. Goldstein, T., Bresson, X., Osher, S.: Geometric applications of the split Bregman method: segmentation and surface reconstruction. *J. Sci. Comput.* **45**, 272–293 (2010)
20. Goldstein, T., Osher, S.: The split Bregman method for L1-regularized problems. *SIAM J. Imaging Sci.* **2**(2), 323–343 (2009)
21. Jin, Z., Yang, X.: A variational model to remove the multiplicative noise in ultrasound images. *J. Math. Imaging Vis.* **39**, 62–74 (2011)
22. Kang, M., Jung, M., Kang, M.: Nonconvex higher-order regularization based Rician noise removal with spatially adaptive parameters. *J. Vis. Commun. Image R.* **32**, 180–193 (2015)
23. Krishnan, D., Fergus, R.: Fast image deconvolution using hyper-Laplacian priors. In: *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1033–1041 (2009)
24. Krissian, K., Kikinis, R., Westin, C., Vosburgh, K.: Speckle constrained filtering of ultrasound images. *IEEE Comput. Vis. Pattern Recogn.* **2**, 547–552 (2005)
25. Kuan, D.T., Sawchuk, A.A., Strand, T.C., Chavel, P.: Adaptive restoration of images with speckle. *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-35**, 373–383 (1987)
26. Lee, J.S.: Speckle suppression and analysis for synthetic aperture radar. *Opt. Eng.* **25**(5), 636–643 (1999)
27. Li, F., Shen, C., Fan, J., Shen, C.: Image restoration combining a total variational filter and a fourth-order filter. *J. Vis. Commun. Image R.* **18**(4), 322–330 (2007)
28. Loupas, T., McDicken, W.N., Allan, P.L.: An adaptive weighted median filter for speckle suppression in medical ultrasonic images. *IEEE Trans. Circuits Syst.* **36**(1), 129–135 (1989)
29. Lysaker, M., Lundervold, A., Tai, X.C.: Noise removal using fourth-order partial differential equation with application to medical magnetic resonance images in space and time. *IEEE Trans. Image Process.* **12**(12), 1579–1590 (2003)
30. Nesterov, Y.: *Introductory Lectures on Convex Optimization: A Basic Course*, vol. 87. Kluwer, Boston (2004)
31. Nikolova, M., Ng, M.K., Tam, C.P.: Fast nonconvex nonsmooth minimization methods for image restoration and reconstruction. *IEEE Trans. Image Process.* **19**(12), 3073–3088 (2010)
32. Nikolova, M., Ng, M.K., Zhang, S., Ching, W.: Efficient reconstruction of piecewise constant images using nonsmooth nonconvex minimization. *SIAM J. Imaging Sci.* **1**(1), 2–25 (2008)
33. Nocedal, J., Wright, S.J.: *Numerical Optimization*. Springer, New York (2006)
34. Ochs, P., Dosovitskiy, A., Brox, T., Pock, T.: An iterated ℓ_1 algorithm for non-smooth non-convex optimization in computer vision. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2013)
35. Ochs, P., Dosovitskiy, A., Brox, T., Pock, T.: On iteratively reweighted algorithm for nonsmooth nonconvex optimization in computer vision. *SIAM J. Imaging Sci.* **8**(1), 331–372 (2015)
36. Oh, S., Woo, H., Yun, S., Kang, M.: Non-convex hybrid total variation for image denoising. *J. Vis. Commun. Image R.* **24**(3), 332–344 (2013)
37. Papafitsoros, K., Schönlieb, C.B.: A combined first and second order variational approach for image restoration. *J. Math. Imaging Vis.* **48**, 308–338 (2014)
38. Robini, M., Lachal, A., Magnin, I.: A stochastic continuation approach to piecewise constant reconstruction. *IEEE Trans. Image Process.* **16**(10), 2576–2589 (2007)
39. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Phys. D Nonlinear Phenom.* **60**(1), 259–268 (1992)
40. Setzer, S., Steidl, G.: Variational methods with higher-order derivatives in image processing. In: *Approximation Theory XII: San Antonio*, pp. 360–386 (2008)
41. Setzer, S., Steidl, G., Teuber, T.: Deblurring Poissonian images by split Bregman techniques. *J. Vis. Commun. Image R.* **21**(3), 193–199 (2010)
42. Shi, J., Osher, S.: A nonlinear inverse scale space method for a convex multiplicative noise model. *SIAM J. Imaging Sci.* **1**(3), 294–321 (2008)
43. Steidl, G., Teuber, T.: Removing multiplicative noise by Douglas-Rachford splitting methods. *J. Math. Imaging Vis.* **36**(2), 168–184 (2010)
44. Teboul, S., Blanc-Feraud, L., Aubert, G., Barlaud, M.: Variational approach for edge-preserving regularization using coupled PDEs. *IEEE Trans. Image Process.* **7**(3), 387–397 (1998)
45. Vese, L., Chan, T.: Redced non-convex functional approximations for image restoration and segmentation. *UCLA CAM Report 97-56* (1997)
46. Woo, H., Yun, S.: Proximal linearized alternating direction method for multiplicative denoising. *SIAM J. Sci. Comput.* **35**(2), B336–B358 (2013)
47. Yang, J., Zhang, Y.: Alternating direction algorithms for ℓ_1 problems in compressive sensing. *SIAM J. Sci. Comput.* **33**(1), 250–278 (2011)