Research Article

New Regularization Models for Image Denoising with a Spatially Dependent Regularization Parameter

Tian-Hui Ma, Ting-Zhu Huang, and Xi-Le Zhao

School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China

Correspondence should be addressed to Tian-Hui Ma; nkmth0307@126.com

Received 31 July 2013; Accepted 2 October 2013

Academic Editor: Peilin Shi

Copyright © 2013 Tian-Hui Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We consider simultaneously estimating the restored image and the spatially dependent regularization parameter which mutually benefit from each other. Based on this idea, we refresh two well-known image denoising models: the LLT model proposed by Lysaker et al. (2003) and the hybrid model proposed by Li et al. (2007). The resulting models have the advantage of better preserving image regions containing textures and fine details while still sufficiently smoothing homogeneous features. To efficiently solve the proposed models, we consider an alternating minimization scheme to resolve the original nonconvex problem into two strictly convex ones. Preliminary convergence properties are also presented. Numerical experiments are reported to demonstrate the effectiveness of the proposed models and the efficiency of our numerical scheme.

1. Introduction

Image denoising is a fundamental problem in image processing and computer vision. In many real-world applications, it forms a significant preliminary step for subsequent image processing operations, such as object recognition, medical image analysis, surveillance, and many more. During acquisition and transmission, images are often corrupted by Gaussian noise. In this problem, the degradation process is modeled as

$$f = u_{\text{exact}} + \eta, \tag{1}$$

where f, u_{exact} , and η represent the observed image, the original image, and the additive white Gaussian noise, respectively.

The objective of image denoising is to compute a good estimate of u_{exact} from f. To obtain a reasonable approximated solution from (1), the regularization method, generally used as a numerical technique for stabilizing inverse problems, has been increasingly applied to image denoising over the past decades. A large class of regularization based denoising methods unifies under the following framework:

$$\min_{u} E(u) = F(u) + \alpha R(u), \qquad (2)$$

where u denotes the restored image to be estimated. In the right-hand side of (2), F(u) represents the fidelity term,

which measures the closeness of the estimate to the data. The functional R(u) is called the regularization term pushing u to exhibit some a priori expected features. Parameter α in (2) is known as the regularization parameter, which balances the tradeoff between the two terms.

One key topic in regularization methods is the choice of the regularizer. Among many regularization based denoising models, the total variation regularization, proposed by Rudin, Osher, and Fatemi (ROF) [1], has won tremendous success due to its edge-preserving property. In the ROF model, the restoration result u is generated by solving the following minimization problem:

$$\min_{u} E(u) = \int_{\Omega} \left(u - f \right)^2 dx + \alpha R_{\text{ROF}}(u), \qquad (3)$$

where the TV-term $R_{\text{ROF}}(u) = \int_{\Omega} |Du|$ denotes the total variation of u (see [2] for more details) and Ω denotes the image domain. A remarkable aspect of the ROF model is that the TV-term does not penalize the discontinuities in u; see, for example, [3]. This property allows us to restore the edges of the original image. However, the main disadvantage of the ROF model is the so-called staircase effect (smooth regions are transformed into piecewise constant regions), a phenomenon long observed in the literature [4–6]. As



FIGURE 1: The numerical behavior of g_1^* and g_2^* ; here $\alpha = 0.05$, M = 0.05, and r = 7: (a) the original image, (b) g_1^* , and (c) g_2^* .



FIGURE 2: Original images: (a) synthetic image 1 (128×128), (b) synthetic image 2 (256×256), (c) "Lena" (256×256), and (d) "Barbara" (256×256).

a consequence, the restoration result is unsatisfactory to the eye due to the loss of texture details and the generation of artificial edges that do not exist in the true image. To overcome this effect, models based on high-order PDEs have been proposed in the literature; see, for example, [7–11]. For instance, Lysaker, Lundervold, and Tai [8] proposed a fourthorder PDE model (termed as the LLT model) with the form

$$\min_{u} E(u) = \int_{\Omega} \left(u - f \right)^2 dx + \alpha R_{\text{LLT}}(u), \qquad (4)$$

where $R_{\text{LLT}}(u) = \int_{\Omega} |\nabla^2 u| dx$ is a fourth-order filter. From atheoretical point of view [12], it has been shown that fourthorder PDEs are superior to second-order PDEs in some aspects, including avoiding the staircase effect. However, this type of filters usually blurs the edges of the original image and suffers from the speckle effect in homogeneous regions. Since the ROF model and the LLT model are of both merits and drawbacks, it may be desirable to promote solutions that simultaneously exhibit properties that are enforced by both regularizers. This is the basic idea of [13] proposed by Lysaker and Tai, in which they studied an iterative algorithm based on combining the results generated by (3) and (4). But their method needs to solve two separate PDEs and their combination is not quite intuitive. Li et al. [14] proposed a hybrid model in the following form:

$$\min_{u} E(u) = \int_{\Omega} (u - f)^{2} dx + \alpha \left(\int_{\Omega} (1 - g) |Du| + \int_{\Omega} g |\nabla^{2}u| dx \right),$$
(5)

where the function *g* is chosen as an edge detector $1/(1 + \gamma + k|\nabla G_{\sigma} * f|^2)$ to control the relative weights of the two regularizers (see [14] for details). By their selection of *g*, the second-order filter plays the dominant role where $|\nabla G_{\sigma} * f|$ is large (regions including sharp features), whereas the fourth-order filter plays the dominant role where $|\nabla G_{\sigma} * f|$ is small (homogeneous regions). Therefore, (5) reaps the benefits of both regularizers. More related works on the combination of the ROF model and the LLT model can be found in [15–17].

Another crucial issue in the regularization process is the suitable selection of the regularization parameter. In regularization models, the regularization parameter controls the relative weights of the data fidelity and regularization terms. However, due to the inhomogeneous distribution of cartoon, textures, and small details in an image (in terms of variance), a global constant parameter may not be suitable for these features of different scales. According to the cartoon pyramid model studied in [18], as the regularization parameter goes from small to large values,



FIGURE 3: Results of different models for the synthetic image 1: (a) the noisy image, the restored images by (b) the ROF model, (c) the LLT model, (d) the hybrid model, (e) our Model 1 (α = 0.1 and r = 1), and (f) our Model 2 (α = 0.18 and r = 1).

	Figure 2(a)	Figure 2(b)	Figure 2(c)	Figure 2(d)
ReErr				
Noisy image	0.1064	0.1847	0.1913	0.1643
ROF model	0.0522	0.1279	0.0784	0.0925
LLT model	0.0633	0.1273	0.0807	0.0894
Hybrid model	0.0509	0.1224	0.0758	0.0879
Model 1	0.0361	0.1151	0.0751	0.0842
Model 2	0.0298	0.1132	0.0729	0.0827
SNR				
Noisy image	9.00	6.43	5.49	6.42
ROF model	15.19	9.62	13.24	11.41
LLT model	13.51	9.66	12.99	11.71
Hybrid model	15.41	10.00	13.53	11.85
Model 1	18.39	10.53	13.61	12.23
Model 2	20.06	10.68	13.87	12.38
Time				
ROF model	0.28	1.22	1.26	1.05
LLT model	1.33	4.63	1.73	3.06
Hybrid model	2.23	8.44	4.48	5.46
Model 1	2.87	6.36	5.18	4.65
Model 2	2.50	9.63	5.93	6.08

TABLE 1: Summarized results of the comparative experiment.



FIGURE 4: Results of different models for the synthetic image 2: (a) the noisy image, the restored images by (b) the ROF model, (c) the LLT model, (d) the hybrid model, (e) our Model 1 (α = 0.21 and *r* = 7), and (f) our Model 2 (α = 0.35 and *r* = 7).

Table	2: Summ	arized result	ts of the p	proposed	models	for	different
values	of L with	our initializ	ation and	zero initi	alizatior	1.	

	L	Model 1			Model 2		
		ReErr	SNR	Time	ReErr	SNR	Time
	3	0.0754	13.57	4.66	0.0731	13.85	7.10
Our initialization	6	0.0752	13.60	4.21	0.0730	13.86	5.71
Our mitialization	10	0.0751	13.61	5.51	0.0729	13.87	6.19
	15	0.0752	13.60	6.27	0.0729	13.86	7.74
	10	0.0780	13.28	6.27	0.0768	13.42	5.87
Zero initialization	20	0.0758	13.54	15.93	0.0740	13.74	14.21
	30	0.0753	13.58	29.83	0.0733	13.83	23.87
	40	0.0753	13.59	49.03	0.0730	13.85	33.28

the corresponding solutions generated by the ROF model range from undersmoothed (textures are preserved, while noise remains almost unchanged in homogeneous regions) to oversmoothed (noise is reduced well, but significant details are lost). This suggests that a spatially dependent regularization parameter including, instead of a single constraint, a group of constraints adapted to different regions of the image is desired to obtain higher quality results. To this end, denoising methods with a spatially dependent regularization parameter have been extensively studied; see, for example, [18–22]. In these works, the automated selection strategies of the regularization parameter are based on local variance measures [18, 20–22] and the local statistical characteristics of the noise [21, 22].

Note that, in models (3), (4), and (5), the selection of the regularization parameter never accommodates the information of the current restored image. This triggers us to seek a new approach for regularization parameter estimation. Our main motivation is to simultaneously estimate both the restored image and the regularization parameter which mutually benefit from each other during the denoising procedure. We propose a general model with the following form:

$$\min_{u,g} E\left(u,g\right) = F\left(u\right) + F\left(g\right) + R\left(u,g\right),\tag{6}$$

where g denotes a spatially dependent regularization parameter, F(u) and F(g) are, respectively, fidelity terms for the two variables, and the binary function R(u, g) represents a regularizer. The advantages of (6) are presented as follows. First, instead of a global constant regularization parameter, a spatially dependent regularization parameter is more flexible to image features of different scales. Second, since the estimation of the two variables is derived simultaneously, the regularization parameter is changing more reasonablly by exploiting the more accuracy restored image instead of



FIGURE 5: Results of different models for the image "Lena": (a) the noisy image, the restored images by (b) the ROF model, (c) the LLT model, (d) the hybrid model, (e) our Model 1 (α = 0.26 and *r* = 3), and (f) our Model 2 (α = 0.36 and *r* = 3).

the observed images corrupted by noise. Here we remark that (6) is a general-purpose model which can incorporate various classical methods. Specially we focus on the fourthorder filter in (4) and the hybrid regularizer in (5). The refresh models are named as Models 1 and 2, respectively. Model 1 can suppress the speckle effect caused by the LLT model while less overregularizing textures. Model 2 has the advantage of better restoring textures and homogeneous regions while preserving edges. To overcome the nonconvexity of our models, we utilize an alternating minimization scheme to resolve the original nonconvex problem into two strictly convex ones. Thus, our models can be asymptotically solved.

The outline of the rest of the paper is as follows. In the next section, we give notations and discretizations. In Section 3, we introduce our models with some discussions. The following Section 4 presents the numerical scheme for solving the proposed models. In Section 5, numerical experiments are given to demonstrate the performance of the proposed models. Finally, we conclude the paper in Section 6.

2. Notations and Discretizations

From now on, we will restrict our attention to the discrete setting. We first introduce some notations. Without loss of generality, we assume that all the images in this paper are grayscale and have a square domain. Then we represent an image u as an $n \times n$ matrix, where $u_{i,j}$ represents the intensity value of u at pixel (i, j), for i, j = 1, ..., n. For the sake of simplicity, we assume that the image is periodically extended, and then the FFT can be adopted in our algorithm. It should be pointed out that adaption to other boundary conditions is not difficult in principle. In the rest of this paper, we let $\|\cdot\|_2$, $\|\cdot\|_F$, and \circ denote the 2-norm, the Frobenius norm, and the Hadamard product, respectively. Let X denote the Euclidean space $\mathbb{R}^{n \times n}$. The usual inner product and Euclidean norm of X are denoted as $\langle \cdot, \cdot \rangle_X$ and $\|\cdot\|_X$, respectively. Denote by Y the space $X \times X$ equipped with the inner product $\langle \cdot, \cdot \rangle_Y$ leading canonically to the norm $\|\cdot\|_Y$, that is; for $p = (p^1, p^2) \in Y$ and $q = (q^1, q^2) \in Y$,

$$\langle p,q\rangle_{Y} = \langle p^{1},q^{1}\rangle_{X} + \langle p^{2},q^{2}\rangle_{X}, \qquad \|p\|_{Y} = \sqrt{\langle p,p\rangle_{Y}}.$$
(7)

Moreover, for $y = (y^1, y^2) \in Y$, |y| denotes the $n \times n$ matrix whose element $|y|_{i,j}$ is equal to $||y_{i,j}||_2$ with $y_{i,j} = (y_{i,j}^1, y_{i,j}^2)$. We denote the space $Y \times Y$ as Z. The definitions of the inner product $\langle \cdot, \cdot \rangle_Z$, the norm $|| \cdot ||_Z$, and |z| are analogous to those of Y.



FIGURE 6: Results of different models for the image "Barbara": (a) the noisy image, the restored images by (b) the ROF model, (c) the LLT model, (d) the hybrid model, (e) our Model 1 ($\alpha = 0.28$ and r = 7), and (f) Model 2 ($\alpha = 0.39$ and r = 7).

Now we introduce a discretized version of some necessary operators. For $u \in X$, we introduce forward and backward difference operators as follows:

$$\begin{pmatrix} \mathring{D}_{x}^{+} u \end{pmatrix}_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j}, & 1 \le j \le n-1, \\ u_{i,1} - u_{i,n}, & j = n, \end{cases}$$

$$\begin{pmatrix} \mathring{D}_{y}^{+} u \end{pmatrix}_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j}, & 1 \le i \le n-1, \\ u_{1,j} - u_{n,j}, & i = n, \end{cases}$$

$$\begin{pmatrix} \mathring{D}_{x}^{-} u \end{pmatrix}_{i,j} = \begin{cases} u_{i,1} - u_{i,n}, & j = 1, \\ u_{i,j} - u_{i,j-1}, & 2 \le j \le n, \end{cases}$$

$$\begin{pmatrix} \mathring{D}_{y}^{-} u \end{pmatrix}_{i,j} = \begin{cases} u_{1,j} - u_{n,j}, & i = 1, \\ u_{i,j} - u_{i-1,j}, & 2 \le i \le n. \end{cases}$$

$$(8)$$

Second-order difference operators can be expressed by using a recursive application of first-order difference operators; that is, the operator \dot{D}_{xx}^{-+} is defined by

$$\left(\overset{\circ}{D}_{xx}^{-+}u\right)_{i,j} := \left(\overset{\circ}{D}_{x}^{-}\left(\overset{\circ}{D}_{x}^{+}u\right)\right)_{i,j}.$$
(9)

Other second order difference operators used in this paper such as $\overset{\circ}{D}_{xx}^{+-}$, $\overset{\circ}{D}_{xy}^{++}$, $\overset{\circ}{D}_{yx}^{++}$, $\overset{\circ}{D}_{xy}^{--}$, $\overset{\circ}{D}_{yy}^{--}$, $\overset{\circ}{D}_{yy}^{-+}$, and $\overset{\circ}{D}_{yy}^{+-}$ can be

similarly defined. Based on the definitions above, we can introduce the discrete gradient operator. For $u \in X$, ∇u is a vector of *Y*, which is given by

$$\nabla u = \left(\overset{\circ}{D}_{x}^{+} u, \overset{\circ}{D}_{y}^{+} u \right).$$
 (10)

The discrete total variation of *u* is then given by

$$R_{\text{ROF}}(u) = \sum_{1 \le i, j \le n} |\nabla u|_{i,j}.$$
(11)

To discretize the fourth-order filter in the LLT model (4), we have to introduce the discrete Hessian operator. Here we adopt the definition in [23]. The discrete Hessian operator is a mapping $H: X \rightarrow Z$, and for $u \in X$, Hu is defined by

$$Hu = \begin{pmatrix} \mathring{D}_{xx}^{-+} u & \mathring{D}_{xy}^{++} u \\ \\ \mathring{D}_{yx}^{++} u & \mathring{D}_{yy}^{-+} u \end{pmatrix}.$$
 (12)

Then, $R_{\text{LLT}}(u)$ can be discretized as

$$R_{\text{LLT}}\left(u\right) = \sum_{1 \le i, j \le n} |Hu|_{i,j}.$$
(13)

We also introduce two important operators: div : $Y \rightarrow X$ and $H^* : Z \rightarrow X$, that is, the adjoint operator of $-\nabla$ and

FIGURE 7: Zoomed-in local results corresponding to Figures 4(b)-4(f): (a) the original image, results of (b) the ROF model, (c) the LLT model, (d) the hybrid model, (e) our Model 1, and (f) our Model 2.

H, respectively. By analogy with the continuous setting, for $u \in X$, $y \in Y$, and $z \in Z$, we want them to satisfy

$$\langle \operatorname{div} y, u \rangle_X = \langle y, -\nabla u \rangle_Y, \qquad \langle H^* z, u \rangle_X = \langle z, Hu \rangle_Z.$$
(14)

Then they are formulated as follows:

$$\operatorname{div} y = \overset{\circ}{D_{x}} \overset{-}{y^{1}} + \overset{\circ}{D_{y}} \overset{-}{y^{2}},$$

$$H^{*}z = \overset{\circ}{D_{xx}} \overset{+}{z^{11}} + \overset{\circ}{D_{yx}} \overset{-}{z^{12}} + \overset{\circ}{D_{xy}} \overset{-}{z^{21}} + \overset{\circ}{D_{yy}} \overset{+}{z^{22}}.$$
(15)

Finally, the composite operators $\Delta = \operatorname{div} \cdot \nabla$ and H^*H are also used.

3. The Proposed Models and Discussion

3.1. The Proposed Models. Arising from model (6) and exploiting the benefits of the LLT model (4) and the hybrid model (5), we consider and study the following models.

Model 1. One has

$$\min_{u,g\in X} E_1(u,g) = \frac{1}{2} \|u - f\|_X^2 + \alpha \|g - \frac{M}{\alpha} \cdot \mathbf{1}\|_X^2 + \sum_{1\leq i,j\leq n} g_{i,j}^2 (K(|Hu|))_{i,j}.$$
(16)

Model 2. One has

$$\begin{aligned} \min_{u,g_1,g_2 \in X} E_2(u,g_1,g_2) \\ &= \frac{1}{2} \| u - f \|_X^2 \\ &+ \alpha \left(\left\| g_1 - \frac{M}{\alpha} \cdot \mathbf{1} \right\|_X^2 + \left\| g_2 - \frac{M}{\alpha} \cdot \mathbf{1} \right\|_X^2 \right) \\ &+ \sum_{1 \le i,j \le n} (g_1)_{i,j}^2 (K(|\nabla u|))_{i,j} \\ &+ (g_2)_{i,j}^2 (K(|Hu|))_{i,j}. \end{aligned}$$
(17)

In the above two models, α and M are positive parameters, 1 represents the matrix whose elements are equal to 1, and $K : X \rightarrow X$ denotes a discrete mean filter. More precisely, we choose an odd integer r and define an r-by-r window



FIGURE 8: Zoomed-in local results corresponding to Figures 6(b)-6(f): (a) the original image, results of (b) the ROF model, (c) the LLT model, (d) the hybrid model, (e) our Model 1, and (f) our Model 2.

centered at pixel (i, j) (with a periodic extension at the boundary); that is,

$$\Omega_{i,j}^{r} = \left\{ (s,t) : \min(|s-i|, n-|s-i|) \le \frac{r-1}{2}, \\ \min(|t-j|, n-|t-j|) \le \frac{r-1}{2} \right\},$$
(18)

where s, t = 1, ..., n. Then for $u \in X$, K(u) is given by

$$(K(u))_{i,j} = \frac{1}{r^2} \sum_{(s,t) \in \Omega_{i,j}^r} u_{s,t}.$$
 (19)

It is immediately clear that if r = 1, K coincides with the identity operator. We start our investigations of the proposed models by the following lemma.

Lemma 1. For any $u, v \in X$, one has

$$\sum_{1 \le i,j \le n} u_{i,j}(K(v))_{i,j} = \sum_{1 \le i,j \le n} (K(u))_{i,j} v_{i,j}.$$
 (20)

Proof. We define the following function $k : \mathbb{R}^2 \to \mathbb{R}$:

$$k(x, y) = \begin{cases} \frac{1}{r^2} & \text{if } \min(|x|, n - |x|) \le \frac{r-1}{2}, \\ & \min(|y|, n - |y|) \le \frac{r-1}{2}, \\ 0 & \text{else.} \end{cases}$$
(21)

From the definition of *k*, it is immediate to see that k(x, y) = k(-x, -y). Thus, we have

$$\sum_{1 \le i,j \le n} u_{i,j}(K(v))_{i,j} = \sum_{1 \le i,j \le n} u_{i,j} \left(\sum_{1 \le s,t \le n} k\left(i-s,j-t\right)v_{s,t} \right)$$
$$= \sum_{1 \le s,t \le n} v_{s,t} \left(\sum_{1 \le i,j \le n} k\left(s-i,t-j\right)u_{i,j} \right)$$
$$= \sum_{1 \le s,t \le n} (K(u))_{s,t}v_{s,t}$$
$$= \sum_{1 \le i,j \le n} (K(u))_{i,j}v_{i,j}.$$
(22)

According to Lemma 1, we can rewrite the regularizers in our models in another equivalent form. In (16), for example, we rewrite

$$\sum_{1 \le i, j \le n} g_{i,j}^2 (K(|Hu|))_{i,j} = \sum_{1 \le i, j \le n} (K(g \circ g))_{i,j} |Hu|_{i,j}.$$
 (23)

Next we study the existence result for the solution of our models.

Abstract and Applied Analysis



FIGURE 9: Contour plots corresponding to Figures 5(b)–5(f): (a) the ideal contour plot, results of (b) the ROF model, (c) the LLT model, (d) the hybrid model, (e) our Model 1, and (f) our Model 2.



FIGURE 10: Final values of the regularization parameters in the proposed models and weighting functions in the hybrid model corresponding to Figures 4(d)–4(f) and 6(d)–6(f). (a) g in the hybrid model, (b) g^* in Model 1, (c) g_1^* in Model 2, (d) g_2^* in Model 2, (e) g in the hybrid model, (f) g^* in Model 1, (g) g_1^* in Model 2, and (h) g_2^* in Model 2.

Theorem 2. Assume that $\alpha > 0$, M > 0, and $f \in X$. Then both Models 1 and 2 have a global minimizer.

Proof. First, under our assumption, both models are bounded from below, which implies that the minimization problems are the correct setting.

It is immediate to see that the functions $E_1 : X \times X \rightarrow \mathbb{R}$ and $E_2 : X \times X \times X \rightarrow \mathbb{R}$ are proper, coercive, and continuous. Then, the existence of a global minimizer is deduced by Weierstrass' theorem [24].

We end this subsection with the following proposition which describes the relationship between the variables in a solution of our models.

Proposition 3. (i) Assume that $(u^*, g^*) \in X \times X$ is a (local) minimizer of Model 1. Then (u^*, g^*) satisfies

$$u^{*} = \arg \min_{u} E_{1}(u, g^{*}),$$

$$g^{*} = \arg \min_{q} E_{1}(u^{*}, g).$$
(24)

(ii) Assume that $(u^*, g_1^*, g_2^*) \in X \times X \times X$ is a (local) minimizer of Model 2. Then (u^*, g_1^*, g_2^*) satisfies

$$u^{*} = \arg\min_{u} E_{2}(u, g_{1}^{*}, g_{2}^{*}),$$

$$(g_{1}^{*}, g_{2}^{*}) = \arg\min_{g_{1}, g_{2}} E_{2}(u^{*}, g_{1}, g_{2}).$$
(25)

Proof. (i) Under our assumption, there exists some $\epsilon > 0$ such that

$$E_1(u^*, g^*) \le E_1(u, g) \quad \forall u, g \in X$$

with $||u - u^*||_X \le \epsilon$, $||g - g^*||_X \le \epsilon$. (26)

Then, we have

$$E_1(u^*, g^*) \le E_1(u, g^*) \quad \forall u \in X \text{ with } \|u - u^*\|_X \le \epsilon,$$
(27)

which implies that u^* is a local minimizer of the following minimization problem:

$$\min_{u \in X} E_1(u, g^*) = \frac{1}{2} \|u - f\|_X^2 + \sum_{1 \le i, j \le n} (K(g^* \circ g^*))_{i,j} |Hu|_{i,j}.$$
(28)

According to the nonnegativity of $(K(g^* \circ g^*))_{i,j}$, we can deduce that the objective function of (28) is strictly convex. Then, u^* is the unique global minimizer of (28). Similarly, we can obtain that g^* is the unique global minimizer of the following strictly convex minimization problem:

$$\min_{g \in X} E_1(u^*, g) = \alpha \left\| g - \frac{M}{\alpha} \cdot \mathbf{1} \right\|_X^2 + \sum_{1 \le i, j \le n} g_{i,j}^2(K(|Hu^*|))_{i,j}.$$
(29)

Hence we complete the proof of (i).

(ii) Using the same technique, (ii) can be similarly proved, and we omit the details. $\hfill \Box$

3.2. Discussion on the Proposed Models. In this subsection we discuss the strengths of the proposed models by investigating the numerical behavior of the solution. Proposition 3 implies that there is an interaction between the restored image and the regularization parameter in a solution of our models, which can achieve joint optimality. In other words, both variables benefit from each other. Assume that (u^*, g^*) and (u^*, g_1^*, g_2^*) are solutions of Models 1 and 2, respectively. According to Proposition 3, g^*, g_1^* , and g_2^* can be calculated directly from u^* . Since the minimization subproblems with respect to g^* and g_2^* are actually the same, we just consider g_1^* and g_2^* for the sake of brevity. They are given by

$$(g_1^*)_{i,j} = \frac{M}{\alpha + (K(|\nabla u^*|))_{i,j}},$$

$$(g_2^*)_{i,j} = \frac{M}{\alpha + (K(|Hu^*|))_{i,j}}.$$
(30)

We are interested in the numerical behavior of g_1^* and g_2^* corresponding to regions of u^* with different scales (e.g., texture, flat, and ramp regions). Figure 1(a) shows a 256×256 synthetic image. The corresponding g_1^* and g_2^* are shown in Figures 1(b) and 1(c), respectively, where u^* is set to be the original image. Form Figure 1, numerical behavior of g_1^* and g_2^* is summarized as follows.

- (i) For texture regions of u^* , g_1^* and g_2^* are both small (darker regions indicate smaller value).
- (ii) For flat regions of u^* , g_1^* and g_2^* are both large.
- (iii) For ramp regions of u^* , g_1^* is inversely proportional to the gradient of u^* , whereas g_2^* is large.

Below we pay attention to the restored image u^* . As we have mentioned in the Introduction, values of the regularization parameter control the relative weights of the fidelity and regularization terms. More precisely, small values lead to little regularization, whereas large values result in overregularization. Based on the behavior analysis above, in Model 1, textures of u^* are not compromised due to the corresponding small values of g^* , and the speckle effect caused by the LLT model is removed from flat and ramp regions of u^* due to the corresponding large values of q^* . Therefore, Model 1 can produce higher quality restoration results than the LLT model. Model 2 inherits the advantages of Model 1 and exhibits some new ones. Edges of u^* are well preserved by total variation regularization. At the same time, the staircase effect is suppressed in ramp regions due to the corresponding small values of g_1^* (the fourth-order filter plays the dominate role). Therefore, Model 2 incorporates the strengths of both regularizers while avoiding their drawbacks. Finally, we remark that, compared with the hybrid model (5), Model 2 is more reasonable. Observe that (5) utilizes a fixed weighting function to combine the two regularizers. However, the weighting function is computed from the noisy image, which leads to inaccuracy. On the other hand, the regularization parameter in Model 2 depends on the restored image, which is less affected by noise.

4. Algorithms

In this section, we formulate the numerical scheme for solving our models. Our basic idea is to use the alternating minimization scheme which is described in Algorithms 1 and 2.

Next we would like to show the convergence of our algorithms.

Theorem 4. (i) Let $\{(u^k, g^k)\}$ be the sequence derived from Algorithm 1. Then $\{(u^k, g^k)\}$ converges to $(\hat{u}, \hat{g}) \in X \times X$ (up to a subsequence), and for any $(u, g) \in X \times X$, one has

$$E_1\left(\widehat{u},\widehat{g}\right) \le E_1\left(u,\widehat{g}\right), \qquad E_1\left(\widehat{u},\widehat{g}\right) \le E_1\left(\widehat{u},g\right). \tag{31}$$

(ii) Let $\{(u^k, g_1^k, g_2^k)\}$ be the sequence derived from Algorithm 2. Then $\{(u^k, g_1^k, g_2^k)\}$ converges to $(\hat{u}, \hat{g}_1, \hat{g}_2) \in X \times X \times X$ (up to a subsequence), and for any $(u, g_1, g_2) \in X \times X \times X$, one has

$$E_{2}\left(\widehat{u}, \widehat{g}_{1}, \widehat{g}_{2}\right) \leq E_{2}\left(u, \widehat{g}_{1}, \widehat{g}_{2}\right),$$

$$E_{2}\left(\widehat{u}, \widehat{g}_{1}, \widehat{g}_{2}\right) \leq E_{2}\left(\widehat{u}, g_{1}, g_{2}\right).$$
(32)

Proof. (i) First, we can easily deduce the following inequality from Algorithm 1:

$$E_1\left(u^{k+1}, g^{k+1}\right) \le E_1\left(u^{k+1}, g^k\right) \le E_1\left(u^k, g^k\right).$$
(33)

Then the sequence $\{E_1(u^k, g^k)\}$ is bounded, and there exists a constant $N \ge 0$ such that

$$E_1\left(u^k, g^k\right) \le N, \quad \forall k \in \mathbb{N}.$$
 (34)

The above inequality reads as

$$\frac{1}{2} \left\| u^{k} - f \right\|_{X}^{2} + \alpha \left\| g^{k} - \frac{M}{\alpha} \cdot \mathbf{1} \right\|_{X}^{2} + \sum_{1 \le i, j \le n} \left(g^{k} \right)_{i,j}^{2} \left(K \left(\left| Hu^{k} \right| \right) \right)_{i,j} \le N,$$
(35)

which implies that $\{(u^k, g^k)\}$ is uniformly bounded in $X \times X$. Then we can find a subsequence $\{(u^{n_k}, g^{n_k})\} \in \{(u^k, g^k)\}$ and $(\hat{u}, \hat{g}) \in X \times X$ such that they satisfy

$$(u^{n_k}, g^{n_k}) \xrightarrow{X \times X} (\widehat{u}, \widehat{g}).$$
 (36)

On the other hand, for any $u \in X$, we have

$$E_{1}\left(u^{n_{k+1}}, g^{n_{k+1}}\right) \leq E_{1}\left(u^{n_{k}+1}, g^{n_{k}+1}\right)$$

$$\leq E_{1}\left(u^{n_{k}+1}, g^{n_{k}}\right) \leq E_{1}\left(u, g^{n_{k}}\right).$$
(37)

Recall that the function $E_1: X \times X \to \mathbb{R}$ is continuous; we obtain

$$E_1\left(\widehat{u},\widehat{g}\right) \le E_1\left(u,\widehat{g}\right),\tag{38}$$

by letting *k* tend to infinity. Similarly, for any $g \in X$, we have

$$E_1(\hat{u}, \hat{g}) \le E_1(\hat{u}, g). \tag{39}$$

Hence we complete the proof of (i).

(ii) Using the same technique, (ii) can be similarly proofed, and we omit the details. $\hfill \Box$

It is obvious that (b) in Algorithm 1 and (b) in Algorithm 2 can be easily solved. To solve (a) in Algorithm 1 and (a) in Algorithm 2, we use the alternating direction method (ADM) [25], which is closely related to the augmented Lagrangian method [23], the Douglas-Rachford splitting algorithm [26], and the split Bregman method [27]. For the sake of brevity, we only present the ADM procedure for solving (a) in Algorithm 2. The ADM procedure for solving (a) in Algorithm 1 can be analogously derived. We rewrite (a) in Algorithm 2 by introducing auxiliary variables p and q as follows:

$$\min_{u \in X, p \in Y, q \in Z} \quad \frac{1}{2} \| u - f \|_X^2 + \sum_{1 \le i, j \le n} \left(K \left(g_1^{k-1} \circ g_1^{k-1} \right) \right)_{i,j} |p|_{i,j} \\
+ \sum_{1 \le i, j \le n} \left(K \left(g_2^{k-1} \circ g_2^{k-1} \right) \right)_{i,j} |q|_{i,j} \\
\text{subject to} \quad p = \nabla u, \quad q = Hu.$$
(40)

Then the problem fits the framework of the alternating direction method. By using the Lagrangian multipliers $\lambda_1 \in Y$ and $\lambda_2 \in Z$, the augmented Lagrangian function of (40) is given by

$$\begin{split} L(u, p, q, \lambda_1, \lambda_2) &= \frac{1}{2} \| u - f \|_X^2 \\ &+ \sum_{1 \le i, j \le n} \left(K \left(g_1^{k-1} \circ g_1^{k-1} \right) \right)_{i,j} |p|_{i,j} \\ &+ \sum_{1 \le i, j \le n} \left(K \left(g_2^{k-1} \circ g_2^{k-1} \right) \right)_{i,j} |q|_{i,j} \\ &+ \left\langle \lambda_1, p - \nabla u \right\rangle_Y + \left\langle \lambda_2, q - Hu \right\rangle_Z \\ &+ \frac{\beta}{2} \left(\| p - \nabla u \|_Y^2 + \| q - Hu \|_Z^2 \right), \end{split}$$
(41)

where $\beta > 0$ is the penalty parameter for the linear constraints. Then the minimization of (40) is achieved by an iterative process: in each iteration, we minimize the augmented Lagrangian function (41) with respect to *u*, *p*, and *q*, given the other two updated in previous iteration, and then update λ_1 and λ_2 . The computational procedure is presented in Algorithm 3.

Now we make some remarks on the ADM procedure. First, we observe that every step in Algorithm 3 has a closed-form solution, and then the alternating direction method can be efficiently implemented. Moreover, for a convex objective function with linear constraints, the convergence of the alternating direction method is guaranteed; see, for example, [28, 29]. Second, to obtain an exact solution of (a) in Algorithm 2, one needs to let *L* (the maximum iteration number of the ADM procedure) tend to infinity; that is, $u^k = u^{k,\infty}$. Since the regularization parameters are not optimal, in practice, it is not necessary to compute u^k exactly. For the sake of computational efficiency, only several ADM iterations are implemented. Third, aiming at faster convergence of

Input: $f, \alpha, M, \text{and } r.$ Output: u^k . Initialization: $g^0 = 0.$ while not converged do Step 1. Given g^{k-1} , computing u^k by solving $\min_{u \in X} E_1(u, g^{k-1}) = \frac{1}{2} \|u - f\|_X^2 + \sum_{1 \le i, j \le n} (K(g^{k-1} \circ g^{k-1}))_{i,j} |Hu|_{i,j}.$ (a) Step 2. Given u^k , computing g^k by solving $\min_{g \in X} E_1(u^k, g) = \alpha \|g - \frac{M}{\alpha} \cdot \mathbf{1}\|_X^2 + \sum_{1 \le i, j \le n} g_{i,j}^2 (K(|Hu^k|))_{i,j}.$ (b) end while

ALGORITHM 1: The alternating minimization method for solving Model 1.

```
Input: f, \alpha, M, \text{ and } r.

Output: u^k.

Initialization: g_1^0 = 0 and g_2^0 = 0.

while not converged do

Step 1. Given g_1^{k-1} and g_2^{k-1}, computing u^k by solving

\min_{u \in X} E_2(u, g_1^{k-1}, g_2^{k-1}) = \frac{1}{2} ||u - f||_X^2 + \sum_{1 \le i, j \le n} (K(g_1^{k-1} \circ g_1^{k-1}))_{i,j} |\nabla u|_{i,j} + \sum_{1 \le i, j \le n} (K(g_2^{k-1} \circ g_2^{k-1}))_{i,j} |Hu|_{i,j}. (a)

Step 2. Given u^k, computing g_1^k and g_2^k by solving

\min_{g_{1,g_2 \in X}} E_2(u^k, g_1, g_2) = \alpha (||g_1 - \frac{M}{\alpha} \cdot 1||_X^2 + ||g_2 - \frac{M}{\alpha} \cdot 1||_X^2) + \sum_{1 \le i, j \le n} (g_1)_{i,j}^2 (K(|\nabla u^k|))_{i,j} + (g_2)_{i,j}^2 (K(|Hu^k|))_{i,j}. (b)

end while
```

ALGORITHM 2: The alternating minimization method for solving Model 2.

the ADM procedure, we initialize each iteration with the auxiliary variables and the Lagrangian multipliers updated in the previous iteration. Numerical examples will be given in Section 5.2 to demonstrate the efficiency of our numerical scheme.

5. Numerical Experiments

In this section, we provide numerical results to illustrate the effectiveness of the proposed models. In Section 5.1, we compare the performance of the proposed models with the ROF model (3), the LLT model (4), and the hybrid model (5). In Section 5.2, we give some criterions on choosing the parameters in our algorithms. All the experiments are performed under Windows 7 and MATLAB R2010a (Version 7.10.0.499) running on a desktop with an Intel Core i3-2130 CPU at 3.40 GHz and 4 GB memory.

5.1. Comparative Results. Four test images shown in Figure 1 are considered for the comparative experiment. The intensity range of the original images is scaled to [0, 1]. The degraded images are corrupted by white Gaussian noise with the noise

level 0.1, which are shown in Figures
$$3(a)$$
, $4(a)$, $5(a)$, and $6(a)$. The quality of the restored images is assessed using the relative error (ReErr) and the signal-to-noise ratio (SNR). They are defined as

ReErr =
$$\frac{\|u - u_c\|_{\rm F}}{\|u\|_{\rm F}}$$
, SNR = $10 \log_{10} \frac{\|u - \overline{u}\|_{\rm F}}{\|u - u_c\|_{\rm F}}$, (42)

where u, \overline{u} , and u_c are the original image, the mean intensity value of u, and the restored image, respectively.

Now we present implementation details of the comparative experiment. In our algorithms, parameters M and L are fixed and set to be 0.07 and 10, respectively. For parameters α and r, their values are reported in the corresponding figures. For solving models (3), (4), and (5), we use the alternating direction method (ADM) [22]. For a fair comparison, all of the parameters in these three models are optimized to achieve the best restoration with respect to the ReErr and the SNR values. In all the ADM procedures, we set the penalty parameter β to be 1 as a default value for most of digital images **Input:** $f, g_1^{k-1}, g_2^{k-1}, r, \beta$, and *L*. **Output:** u^k , p^k , q^k , λ_1^k , and λ_2^k . **Initialization:** $p^{k,0} = p^{k-1}$, $q^{k,0} = q^{k-1}$, $\lambda_1^{k,0} = \lambda_1^{k-1}$, and $\lambda_2^{k,0} = \lambda_2^{k-1}$. while not converged and l < L do Step 1. Given $p^{k,l-1}$, $q^{k,l-1}$, $\lambda_1^{k,l-1}$, and $\lambda_2^{k,l-1}$, computing $u^{k,l}$ by $u^{k,l} = \mathscr{F}^{-1}\left(\frac{\mathscr{F}\left(f - \operatorname{div}\left(\lambda_1^{k,l-1} + \beta p^{k,l-1}\right) + H^*\left(\lambda_2^{k,l-1} + \beta q^{k,l-1}\right)\right)}{1 - \beta \mathscr{F}(\Delta) + \beta \mathscr{F}(H^*H)}\right),$ where \mathscr{F} and \mathscr{F}^{-1} denote the discrete Fourier transform and the inverse discrete Fourier (a) transform, respectively, and Fourier transforms of operators Δ and H^*H are regarded as the transforms of their corresponding convolution kernels. $step 2. \text{ Given } u^{k,l}, \lambda_1^{k,l-1}, \text{ and } \lambda_2^{k,l-1}, \text{ update } p^{k,l} \text{ and } q^{k,l} \text{ by the two-dimensional shrinkage} \\ p_{i,j}^{k,l} = \max \left\{ \left| \nabla u^{k,l} - \frac{\lambda_1^{k,l-1}}{\beta} \right|_{i,j} - \frac{\left(K (g_1^{k-1} \circ g_1^{k-1})_{i,j} \right)}{\beta}, 0 \right\} \frac{\left(\nabla u^{k,l} - (\lambda_1^{k,l-1}/\beta) \right)_{i,j}}{\left| \nabla u^{k,l} - (\lambda_1^{k,l-1}/\beta) \right|_{i,j}},$ (b) $q_{i,j}^{k,l} = \max\left\{ \left| Hu^{k,l} - \frac{\lambda_2^{k,l-1}}{\beta} \right|_{i,j} - \frac{\left(K\left(g_2^{k-1} \circ g_2^{k-1} \right) \right)_{i,j}}{\beta}, 0 \right\} \frac{\left(Hu^{k,l} - \left(\lambda_2^{k,l-1} / \beta \right) \right)_{i,j}}{\left| Hu^{k,l} - \left(\lambda_2^{k,l-1} / \beta \right) \right|_{i,j}}, 0 \right\}$ (c) respectively, where $0 \cdot (0/0) = 0$ is assumed *Step 3.* Given $u^{k,l}$, $p^{k,l}$, and $q^{k,l}$, update $\lambda_1^{k,l}$ and $\lambda_2^{k,l}$ by $\lambda_1^{k,l} = \lambda_1^{k,l-1} + \beta \left(p^{k,l} - \nabla u^{k,l} \right),$ (d) $\lambda_2^{k,l} = \lambda_2^{k,l-1} + \beta \left(q^{k,l} - H u^{k,l} \right),$ (e) respectively. end while

ALGORITHM 3: The alternating minimization method for solving (41).

with intensity range in [0, 1]. We terminate all the algorithms by the following stopping criterion:

$$\frac{\|u_{i-1} - u_i\|_{\rm F}}{\|u_{i-1}\|_{\rm F}} \le 10^{-4}.$$
(43)

In Figures 3, 4, 5, and 6, we display the resulting images restored by different models. The corresponding relative errors, SNR values (dB), and computational time (in seconds) are listed in Table 1. From these results, we find that our models restore better images than the ROF model (3), the LLT model (4), and the hybrid model (5), both visually and quantitatively. In the images restored by the ROF model, we observe that sharp edges are preserved, but the staircase effect is clearly present (e.g., the homogeneous regions in Figures 5 and 6). For the restored results of the LLT model, the staircase effect is suppressed, but some speckle effect is introduced in homogeneous regions (e.g., the background of Figures 3, 4, 5, and 6). We also note that both models (3) and (4) compromise details and textures due to using the global constant regularization parameter. For the images restored by the hybrid model, visual artifacts are almost suppressed, but details are not recovered well (e.g., the ramp region in Figure 3). This is due to the inaccurate estimation of their weighting function from the observed image. Our methods, on the other hand, restore the homogeneous regions significantly better while preserving more details. To better understand the behavior of our methods, some zoomedin local results and contour plots are shown in Figures 7,

8, and 9. We can find that the homogeneous regions are almost smooth (e.g., the background of Figure 3, the flat and ramp regions in Figure 7, and the face of Barbara in Figure 9). The contour plots given in Figure 8 also visually illustrate the above observation. At the same time, details and textures are preserved clearly without being overregularized (e.g., the ramp region in Figure 3, the textures in Figure 7, and the textures on the scarf in Figure 9). Furthermore, we find that Model 2 performs better than Model 1 with respect to preserving edges (e.g., the contours around the lips and nose of Lena in Figure 8). In Figure 10, we show the final values of the regularization parameters in our models compared with the weighting functions in the hybrid model. It is clear that Figures 10(a) and 10(e) are noisy and inaccurate, whereas Figures 10(b)-10(d) and 10(f)-10(h) are less affected by noise. Moreover, we observe that the values of the regularization parameters are small in detail and texture regions, and they are large in homogeneous regions. Therefore, our methods are superior with respect to removing noise while preserving details.

5.2. Parameter Study. In this section, we present some criterions on choosing the parameters which are necessary to start up our algorithms. In our experiments, the parameter settings are the same as those in Section 5.1, except the testing parameters.

The window size r is used to control the smoothness of the spatially dependent regularization parameter. We test

FIGURE 11: SNR for results restored by our models with different window size *r*: (a) Model 1 and (b) Model 2.

our algorithms for different values of r varying from 1 to 19. Figure 11 shows the plots of the SNR values of the restored images. We see from the plots that a small value of r (leads to a sharp regularization parameter) is suitable for images with sharp features (e.g., r = 1 for Figure 2(a) and r = 3 for Figure 2(c)), whereas a moderate value of r (leads to a relatively smooth regularization parameter) is needed for images containing textures (e.g., r = 7 for Figures 2(b) and 2(d)). For large values of r ($r \ge 9$), we find that the changes of the SNR values are not significant. However, if r becomes too large, then the regularization parameter is oversmooth, which compromises image details.

The parameters M and α jointly measure the values of the spatially dependent regularization parameter. The difficulty of tuning these parameters is that they depend not only on the noise level but also on the images. For this reason, trial by error for the two parameters is used. More precisely, we simplify the tuning process by fixing M and searching for the best α of each image.

We also study the behavior of the proposed numerical scheme with respect to the setting of *L* and our initialization. For the setting of Figure 2(c), Table 2 shows the summarized results of our models for different values of L with our initialization and zero initialization (the auxiliary variables and the Lagrangian multipliers are initialized to zero). We see from Table 2 that when using zero initialization, only sufficiently large L (e.g., L = 30 and 40) can provide good restoration result with increasing computational time. Our initialization, on the other hand, produces results as effectively as zero initialization with slightly less computation time. In addition, there is no considerable difference between the relative errors and the SNR values for different values of L. Since a too small or too large L yields comparatively high computational cost, a moderate L (e.g., L = 10) is recommended in our algorithms.

6. Concluding Remarks

We have proposed two regularization models for image denoising, in which both the restored image and the spatially dependent regularization parameter are estimated simultaneously. By the construction of our models, both variables mutually benefit from each other during the denoising process, which can achieve joint optimality. The numerical behavior of the regularization parameter and the strengths of our models have been discussed in detail. The proposed models, which are nonconvex, can be asymptotically solved by the proposed alternating minimization scheme. Finally, the numerical results indicated that our methods outperform several popular methods with respect to both noise removal and detail preservation.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is supported by 973 Program (2013CB329404), NSFC (61170311), Chinese Universities Specialized Research Fund for the Doctoral Program (20110185110020), and Sichuan Province Sci. and Tech. Research Project (2012GZX0080).

References

[1] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.



- [2] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*,, Birkhäuser, Boston, Mass, USA, 1984.
- [3] D. Strong and T. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," *Inverse Problems*, vol. 19, no. 6, pp. S165–S187, 2003.
- [4] W. Ring, "Structural properties of solutions to total variation regularization problems," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 34, no. 4, pp. 799–810, 2000.
- [5] J. Weickert, Anisotropic Diffusion in Image Processing, B. G. Teubner, Stuttgart, Germany, 1998.
- [6] A. Buades, B. Coll, and J. M. Morel, "The staircasing effect in neighborhood filters and its solution," *IEEE Transactions on Image Processing*, vol. 15, pp. 1499–1505, 2006.
- [7] T. Chan, A. Marquina, and P. Mulet, "High-order total variation-based image restoration," *SIAM Journal on Scientific Computing*, vol. 22, no. 2, pp. 503–516, 2000.
- [8] M. Lysaker, A. Lundervold, and X.-C. Tai, "Noise removal using fourth-order partial differential equations with applications to medical magnetic resonance images in space and time," *IEEE Transactions on Image Processing*, vol. 12, pp. 1579–1590, 2003.
- [9] Y.-L. You and M. Kaveh, "Fourth-order partial differential equations for noise removal," *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1723–1730, 2000.
- [10] W. Hinterberger and O. Scherzer, "Variational methods on the space of functions of bounded Hessian for convexification and denoising," *Computing*, vol. 76, no. 1-2, pp. 109–133, 2006.
- [11] M. Lysaker, S. Osher, and X.-C. Tai, "Noise removal using smoothed normals and surface fitting," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1345–1357, 2004.
- [12] J. B. Greer and A. L. Bertozzi, "Traveling wave solutions of fourth order PDEs for image processing," SIAM Journal on Mathematical Analysis, vol. 36, no. 1, pp. 38–68, 2004.
- [13] M. Lysaker and X.-C. Tai, "Iterative image restoration combining total variation minimization and a second-order functional," *International Journal of Computer Vision*, vol. 66, pp. 5– 18, 2006.
- [14] F. Li, C.-M. Shen, J.-S. Fan, and C.-L. Shen, "Image restoration combining a total variational filter and a fourth-order filter," *Journal of Visual Communication and Image Representation*, vol. 18, pp. 322–330, 2007.
- [15] F.-C. Yang, K. Chen, and B. Yu, "Efficient homotopy solution and a convex combination of ROF and LLT models for image restoration," *International Journal of Numerical Analysis and Modeling*, vol. 9, no. 4, pp. 907–927, 2012.
- [16] Q.-S. Chang, X.-C. Tai, and L. Xing, "A compound algorithm of denoising using second-order and fourth-order partial differential equations," *Numerical Mathematics Theory Methods and Applications*, vol. 2, no. 4, pp. 353–376, 2009.
- [17] Z.-F. Pang and Y.-F. Yang, "A projected gradient algorithm based on the augmented Lagrangian strategy for image restoration and texture extraction," *Image and Vision Computing*, vol. 29, pp. 117–126, 2011.
- [18] G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Variational denoising of partly textured images by spatially varying constraints," *IEEE Transactions on Image Processing*, vol. 15, pp. 2281–2289, 2006.
- [19] M. Bertalmio, V. Caselles, B. Rougé, and A. Solé, "TV based image restoration with local constraints," *Journal of Scientific Computing*, vol. 19, pp. 95–122, 2003.
- [20] A. Almansa, C. Ballester, V. Caselles, and G. Haro, "A TV based restoration model with local constraints," *Journal of Scientific Computing*, vol. 34, no. 3, pp. 209–236, 2008.

- [21] Y.-Q. Dong, M. Hintermüller, and M. M. Rincon-Camacho, "Automated regularization parameter selection in multi-scale total variation models for image restoration," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 82–104, 2011.
- [22] K. Bredies, Y.-Q. Dong, and M. Hintermüller, "Spatially dependent regularization parameter selection in total generalized variation models for image restoration," *International Journal of Computer Mathematics*, vol. 90, no. 1, pp. 109–123, 2013.
- [23] C.-L. Wu and X.-C. Tai, "Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 300–339, 2010.
- [24] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientic, Boston, Mass, USA, 2003.
- [25] M. K. Ng, P. Weiss, and X. Yuan, "Solving constrained totalvariation image restoration and reconstruction problems via alternating direction methods," *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2710–2736, 2010.
- [26] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 3, pp. 293–318, 1992.
- [27] T. Goldstein and S. Osher, "The split Bregman method for L1regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.
- [28] M. Ng, F. Wang, and X. Yuan, "Inexact alternating direction methods for image recovery," *SIAM Journal on Scientific Computing*, vol. 33, no. 4, pp. 1643–1668, 2011.
- [29] B. He and X. Yuan, "On the O(1/n) convergence rate of the Douglas-Rachford alternating direction method," *SIAM Journal* on Numerical Analysis, vol. 50, no. 2, pp. 700–709, 2012.