Wavelet-Based Total Variation and Nonlocal Similarity Model for Image Denoising

Yan Shen, Qing Liu, Shuqin Lou, and Ya-Li Hou

***Abstract*—To suppress the heavy noise and keep the distinct edges of the images in the low light condition, we propose a denoising model based on the combination of total variation (TV) and nonlocal similarity in the wavelet domain. The TV regularization in the wavelet domain effectively suppresses the heavy noise with the biorthogonal wavelet function; the nonlocal similarity regularization improves the fine image details. Denoising experiments on artificially degraded and low light images show that in the heavy noise condition, the proposed denoising model can suppress the heavy noise effectively and preserve the detail of images than several state-of-the-art methods.**

***Index Terms*—Birothogonal wavelet, heavy noise, nonlocal similarity, split Bregman, total variation (TV).**

I. Introduction

DENOISING aims to recover an image contaminated by noise. This issue is widely discussed in signal and image processing literatures [1]–[3]. A typical way to restore an image from the contaminated image is the total variation (TV) approach [4] based on the gradient of an image. However, the TV model often produces staircase, which degrades the restoration quality of the image [5]. The wavelet shrinkage technique can suppress the staircase caused by the TV model to some extent because the staircase often becomes the high-frequency components in the wavelet domain. Therefore, the combination of wavelet and TV model is an efficient way for image denoising [6]–[12]. However, most denoising methods depend on assume-guarantee low noise level, which is not true in practical applications because the heavy noise often appears with low light (LL) condition [13], low dose computed tomography (CT) [14], sparse reconstruction [15], etc. The low dose CT can alleviate patient's radiation but lead to the heavy noise in the CT image. The sparse reconstruction of image results in the heavy noise in the image. In a word, the heavy noise leads to low quality of image denoising because most of the details of the image are lost or smoothed out in the process of denoising. Therefore, the suppression of the heavy noise is still a challenge for image denoising technique.

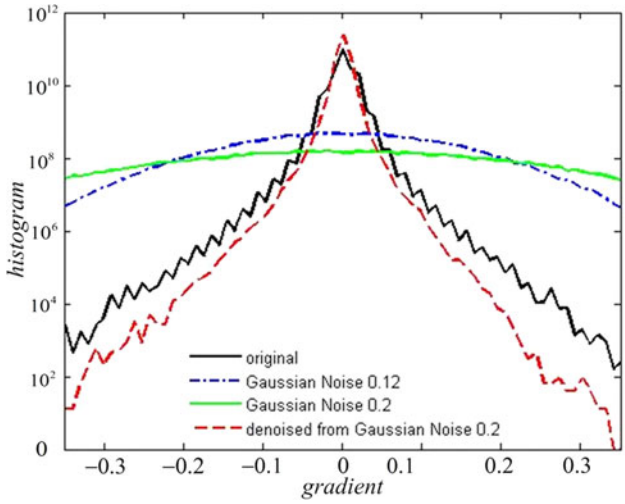

Fig. 1. Curves of the gradient density. Image noise extends the gradients distribution curves and impairs the salient edges through mixing with the high-frequency components of an image. The horizontal axis is the gradient distribution and the vertical axis is the logarithmic density of the gradient.

Fig. 1 shows the curves of logarithmic density (histogram) of a real image's gradients. We can see that with the increase in noise standard deviation, the distribution curves of image gradients along the horizontal axis are extended and the amplitude range (histogram) of image is shrunk along the vertical axis, which degrades the salient edges because the heavy noise deteriorates the edges and thus results in the low peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) of the images.

II. Related Work

In recent years, the denoising techniques have adopted the improved TV models to recover the image. The objective equation is formulated as

$$x = \arg\min_{x} \left\{ \frac{1}{2} \|x - y\|_2^2 + \gamma \cdot \varphi(x) \right\} \tag{1}$$

where y is the noisy image, $\varphi(x)$ is some regularization function of latent image x, and the parameter γ is used to tune $\varphi(x)$. Equation (1) is a TV model when $\varphi(x) = ||\nabla x||_1$. Kamilov *et al.* [6] first transform the image into the gradient domain, and

Manuscript received January 6, 2017; revised February 25, 2017 and March 11, 2017; accepted March 15, 2017. Date of publication March 28, 2017; date of current version May 4, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61475016 and Grant 61301184, in part by Beijing Natural Science Foundation under Grant 4154083, in part by the Fundamental Research Funds for the Central Universities of China under Grant 2015JBM023, and in part by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, under Grant 2015-1098. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Antonio Paiva. *(Corresponding author: Shuqin Lou.)*

The authors are with Beijing Jiaotong University, Beijing 100044, China (e-mail: sheny@bjtu.edu.cn; 15120020@bjtu.edu.cn; shqlou@bjtu.edu.cn; ylhou@bjtu.edu. cn).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2017.2688707

then apply the wavelet transform to the gradients for denoising. However, the wavelet transform will smooth the gradient information of the image and results in the detail loss of the image. Cai *et al.* [7] proposed a variational regularization restoration method with the spline wavelet basis. Durand and Nikolova select $\varphi(x) = [(\nabla x)^2 + a]^{1/2}$ as the regularization term to denoise the image [8]. Wang *et al.* [9] take $\varphi(x)$ as the edge adaptive guiding function to the TV model. Ding and Selesnick [10] propose a unified TV and wavelet denoising method and add the nonconvex penalized function to $\varphi(x)$. Fang *et al.* [11] take $\varphi(x)$ as the sum of TV and wavelet regularization terms. Although some works were extended by using the curvelet, shearlet with the TV regularization [16]–[18], and other block-matching methods [31]–[34], the computation complexity and time are much higher than TV- and wavelet-based methods.

Nonlocal means (NLM) filter is also an effective denoising method for dealing with textures and fine details [19], [20]. NLM takes advantage of the self-similarity of images to restore an unknown pixel. Kindermann *et al.* use the variational method to understand the NLM filter [21]. In [22] and [23], nonlocal operators are introduced to interpret the NLM filter and formalize a systematic and coherent variational framework for nonlocal operators. The nonlocal TV model is the combination of nonlocal similarity and TV model [24], which is solved by the split Bregman iteration.

To improve the denoising performance under the heavy noise, we propose a denoising model combining TV model with nonlocal similarity in the wavelet domain to suppress the heavy noise and enhance the edge details. Then, the split Bregman algorithm is applied to solving this model iteratively. Experimental results show that the proposed model can mitigate noise effectively, keep the edges sharper, and retain higher PSNR and SSIM compared to the previous denoising techniques.

III. Our Model

The heavy noise shrinks the amplitude range of the image, as shown in Fig. 1, and makes the salient edges illegible. Here, we consider the heavy noise with the noise standard deviation of σ above 0.12 because most of the current denoising methods are invalid when σ is above 0.12 [25]. To enhance the edges under the heavy noise, we propose a denoising model combining TV model with nonlocal similarity defined in wavelet domain. Therefore, the classic TV model in (1) can be changed to

$$x = \arg\min_{x} \frac{1}{2} \|Wy - Wx\|_2^2 + \lambda \|\nabla Wx\|_1 + \mu J(Wx) \quad (2)$$

where x is the estimated image with N pixels. W is the wavelet transform, and λ and μ are the parameters for the TV regularization term and the nonlocal filter term defined in the wavelet domain, respectively. The term $||\nabla Wx||_1$ is defined as

$$\|\nabla Wx\|_1 = \sum_{i,j=0}^{N-1} \sqrt{(Wx_{i+1,j} - Wx_{i,j})^2 + (Wx_{i,j+1} - Wx_{i,j})^2}. \quad (3)$$

Equation (2) is an analysis model of x. To simplify (2), we change it into the synthesis model expressed as follows:

$$\hat{s} = \arg\min_{s \in \Omega} \frac{1}{2} \|z - s\|_2^2 + \lambda \|\nabla s\|_1 + \mu J(s) \quad (4)$$

where s and z are the wavelet coefficients, $s = Wx$, $z = Wy$. $s \in \Omega$, where Ω is the support of the wavelet domain of the image. The TV-wavelet regularization term $||\nabla s||_1$ computes the gradient of s. The NLM term $J(s)$ can be written as

$$J(s) = \sum_{p \in \Omega} \sum_{q \in \Omega q} \omega(p, q) \|R_p(s) - R_q(s)\|_2^2 \quad (5)$$

where $R_p(s)$ is the image patch centered at pixel p of size $N_p \times N_p$ ($N_p = 5$), $R_q(s)$ is centered at the neighboring pixel q with the same patch size of $R_p(s)$ and moves in a larger searching window Ω_q of $N_q \times N_q$ ($N_q = 21$) around $R_p(s)$. The weight function ω (p, q) used here is to approximate the similarity between the two adjacent patches $R_p(s)$ and $R_q(s)$.

The TV-wavelet regularization $||\nabla s||_1$ has the property of suppressing the heavy noise by using the gradient information in the wavelet domain. The nonlocal similarity term $J(s)$ defined in wavelet domain has the property of keeping the details of salient edges by adopting the structural similarity of the adjacent patches.

A. Suppress the Heavy Noise With TV-Wavelet Model

The traditional TV-based denoising models cannot suppress the heavy noise effectively because the gradients of images are contaminated by the heavy noise, and thus the edge details are deteriorated. In Fig. 1, the light-colored green curve of the noise standard deviation of 0.2 has the narrowest amplitude (histogram) range, which results in the low quality of denoising. However, the wavelet transform extends the amplitude range of the gradient of wavelet coefficients because the wavelet filters the image and thus enhances the edge information of the image.

The TV-wavelet regularization term $||\nabla s||_1$ of (4) makes use of the gradient information of wavelet coefficients to suppress the heavy noise. Therefore, the TV-wavelet model is represented as

$$\hat{s} = \arg\min_{s} \frac{1}{2} \|z - s\|_2^2 + \lambda \|\nabla s\|_1. \quad (6)$$

For ∇ is not separable, we replaced it by a new variable d. Now, (6) is transformed into

$$(s^{m+1}, d^{m+1}) = \arg\min_{s} \frac{1}{2} \|z - s\|_2^2 + \lambda \|d\|_1 + \frac{\gamma}{2} \|d - \nabla s - b^m\|_2^2 \quad (7)$$

$$b^{m+1} = b^m + \nabla u^{m+1} - d^{m+1} \quad (8)$$

where b is a variable related to the split Bregman iteration, and the penalty parameter γ is a positive constant tuning the TV-wavelet variable s in the low- and high-frequency component, respectively. We minimize either d or s while keeping the other variable fixed by applying the split Bregman iteration to (7). Finally, the denoised image s^{M+1} is obtained after iterated M times or come to the setting error value of ε_1. The detail steps can be referenced in [26].

In Fig. 1, the red dashed curve is the denoised result of our model. It can be seen from Fig. 1 that the red dashed curve is already very approximate the black curve (the original clear image), which means our model is effective in denoising the heavy noise.

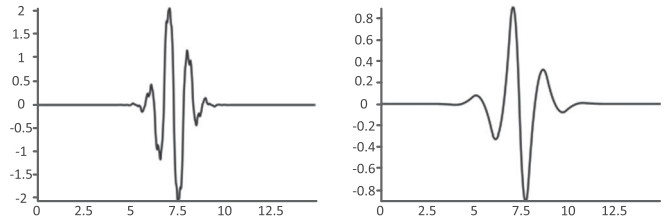

Fig. 2. Coefficients of the (a) decomposing and (b) reconstructing wavelet function of bior 3.7.

Through considerable number of experiments, we find that the biorthogonal wavelet function can obtain better restoration results in both the heavy noise suppression and edge preservation. The reason is that the biorthogonal wavelet has the odd symmetry structures both in decomposing and reconstructing wavelet function, as shown in Fig. 2. Here, we choose the bior 3.7 wavelet in all experiments.

In Fig. 2(a), the difference of the maximum positive and negative peak amplitude can represent the maximal edge variation of the image. On both sides of the two peaks, the amplitudes are small, which contributes less to the edge variation when the wavelet function convolves with the image. Therefore, this odd symmetry structure of the decomposing wavelet function enhances the edges of the image, so as the reconstructing wavelet function in Fig. 2(b).

B. Enhance Edges With the Nonlocal TV-Wavelet Model

In Section III-A, we have revealed that our proposed TV-wavelet model can suppress the heavy noise and obtain salient edges. Even more, we can further improve the edges detail by using the NLM term $J(s)$, the nonlocal similarity defined in the low-frequency band in wavelet domain of (4), to obtain more distinct edges of images. We use the nonlocal similarity in wavelet domain instead of the image domain because the nonlocal similarity in wavelet domain can further preserve the detail information after the heavy noise is suppressed. Here, we apply the nonlocal similarity in the low-frequency part of s^{M+1} and shrink the high-frequency part to distinct the edge detail. Then, the nonlocal TV-wavelet model is represented in

$$\hat{u} = \arg\min_{u} \frac{1}{2} \|f - u\|_2^2 + \mu J(u). \tag{9}$$

We choose the wavelet low frequency of s^{M+1} as the input image f of (9) and u as the wavelet low frequency of the image to be restored. Here, we use the nonlocal similarity in wavelet domain to denoise the image. Therefore, (9) can be expressed as

$$\hat{u} = \operatorname*{argmin}_{u} \frac{1}{2} \|f - u\|_2^2 + \mu \sum_{p} \|\nabla_w u\|_1. \tag{10}$$

Different with the standard TV norm ∇u, here we adopt the nonlocal gradient $\nabla_w u$ [21]

$$\nabla_w u = \sqrt{\sum_{q \in \Omega q} (u(p) - u(q))^2 w(p,q)} \tag{11}$$

$$w(p,q) = \exp\left(-\|R_p(f) - R_q(f)\|^2 / \sigma^2\right). \tag{12}$$

Algorithm 1:

Input: wavelet coefficients f of noisy image.
Output: wavelet coefficients s of desired image.
Initialize: $s^0 = z, d^0 = b^0 = g^0 = c^0 = \mathbf{0}$, $\lambda = [1.2/\sigma + \eta, 4]$, $\gamma = [0.4, 30]$, σ is the standard deviation of noise in the noisy image; $\mu = 0.25$, $\tau = 0.05, \varepsilon_1 = 10^{-3}, \varepsilon_2 = 10^{-4}, m = 0, n = 0$, $M = 90, N = 20$.
For $m = 1$: M
Solve (7) and (8) iteratively to obtain s, d and b.
If $(||s^{m+1} - s^m||_2 / ||s^m||_2 < \varepsilon_1)$; break;
end For
f is the low-frequency band of wavelet coefficients of s^{M+1}, $u^0 = f$.
Shrink the high-frequency part of f.
For $n = 1$: N
Solve (11)–(14) iteratively to obtain w, u, g and c.
If $(||u^{n+1} - u^n||_2 / ||u^n||_2 < \varepsilon_2)$; break;
end For
Combine u^{N+1} and the high-frequency band of s^{M+1} to get final result s.

We solve (10) with the split Bregman iteration process as follows:

$$(u^{n+1}, g^{n+1}) = \arg\min_{u} \mu \|g\|_1 + \frac{1}{2} \|u - f\|_2^2 + \frac{\tau}{2} \|g - \nabla_w u - c^n\|_2^2 \tag{13}$$

$$c^{n+1} = c^n + \nabla_w u^{n+1} - g^{n+1}. \tag{14}$$

The process of solving (13) is detailed in [21]. The overall algorithm is presented in Algorithm 1. In the Initial step, λ and γ include two values, respectively. The first value is the parameter tuning the low frequency in the wavelet domain of z and the second value tuning the high frequency. These parameters are same for all the experiments.

IV. Experimental Results and Analysis

We compare the performance of the proposed method with the previous denoising methods by artificially degraded images and low light (LL) images. We generate artificially degraded images by adding the Gaussian noise to the original clean image with σ of 0.2.

It can be seen from Fig. 3 that when the images are degraded by the heavy noise, the results of the proposed method (the last column) are better than TGV [27], TV [26], IRNTV [28], DTV [29], NLTV [24], and NLmeans [30] methods in noise suppression and detail preservation. All the parameters of the previous methods are tuned to obtain the highest PSNR. For the TGV method, the optimal values of α and β are about 0.19 and 0.15; λ is 5 for TV and 0.14 for IRNTV; for DTV, α and λ are about 2 and 0.16. The parameter η of the proposed method increases 0.01 with different σ, which can be estimated within the feature-free region of the image [3], [13]. From the zoomed areas of Lena images, the hat's details of the proposed method

Fig. 3. Noise suppression of different methods. Images from left to right are noisy images, denoising results of TGV, TV, IRNTV, DTV, NLTV, NLmeans, and the proposed method. The first two rows are Lena and boat images and the last row is LL images. We enlarge parts of these images on right side to see more details of image textural features.

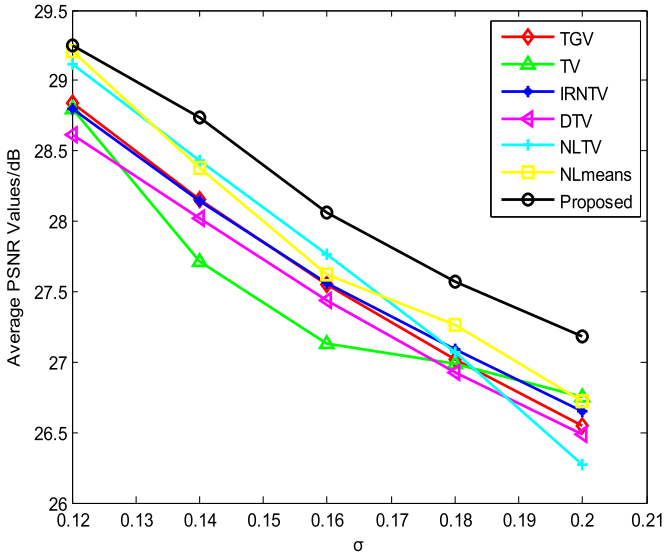

Fig. 4. Comparison of PSNR with the image dataset. Each σ in the horizontal axis corresponds to four classic images.

are more distinct and cleaner than other methods. The images of the third row are the stones in the lake taken under the LL condition. We observe that our method can greatly suppress the heavy noise and keep the edge details of stones; however, other methods failed to preserve them.

Fig. 4 illustrates the comparison of PSNR of the denoised images with different methods. The curves of PSNR are computed by averaging four classic test images shown in the supplementary file. It can be seen that with the increasing of σ along the horizontal axis, the proposed method achieves much higher PSNR than other denoising methods. Especially at σ of 0.2, our method can still reach to 27.18 dB (the input PSNR is 14 dB), higher than other methods. The detail comparison of PSNR and SSIM with different σ is shown in the supplementary file.

Therefore, the proposed method enhances the structures of images and effectively reduces the heavy noise.

The proposed method is also validated by the LL images shown in the supplementary file. Table I shows the comparison of our method with the previous denoising methods by images taken under the LL condition. Images of LL1–LL4 are LL images labeled in the supplementary file. It can be seen from Table I that the proposed method achieves higher PSNR and SSIM than similar methods do and improves PSNR under $\sigma = 0.2$. The detail comparison can be seen in the supplementary file.

TABLE I
COMPARISON OF PSNR AND SSIM OF NATURAL LL IMAGES UNDER THE HEAVY NOISE WITH GAUSSIAN $\sigma = 0.2$. LL MEANS LOW-LIGHT, THE FIRST LINE IS THE PSNR, THE SECOND LINE IS THE SSIM

image	TGV	TV	IRN TV	DTV	NLTV	NL-means	Pro-posed
LL 1	30.48	30.31	31.09	31.86	30.73	32.11	32.72
	0.76	0.75	0.74	0.83	0.73	0.78	0.85
LL 2	25.44	25.14	25.95	26.19	26.40	27.00	26.81
	0.64	0.63	0.67	0.68	0.59	0.65	0.70
LL 3	30.21	30.45	30.52	30.60	28.90	30.12	31.27
	0.78	0.78	0.79	0.78	0.64	0.71	0.81
LL 4	26.15	26.53	26.39	26.80	25.86	26.6	27.03
	0.62	0.68	0.65	0.69	0.67	0.63	0.71

Experimental results show that the proposed model can suppress the heavy noise and enhance details for most of the images. For the image such as LL2, PSNR of the proposed method is lower than the NLmeans method because the structure of the LL2 image is simple. The NLmeans method will smooth the simple structure and obtain higher PSNR. However, the SSIM of the NLmeans is lower than the proposed method because the proposed method has better performance in preserving the structures of the image. The detail comparison can be seen in the supplementary file.

The denoising experiment on Poisson noise is also carried out with the proposed method. The results, given in the supplementary file, validate that the proposed method is also effective.

V. CONCLUSION

In this letter, we propose a denoising model with the combination of TV and nonlocal similarity based on the wavelet domain to suppress the heavy noise and keep the edge details of images. The artificially degraded images and the LL images are used to validate the proposed method. Experimental results show that the proposed method can efficiently suppress the heavy noise and improve the PSNR and SSIM.

REFERENCES

[1] M. Saha, M. K. Naskar, and B. N. Chatterji, "Soft, hard and block thresholding techniques for denoising of mammogram images," *IETE J. Res.*, vol. 61, no. 2, pp. 186–191, Feb. 2015.

[2] S. Allabakash, P. Yasodha, S. V. Reddy, and P. Srinivasulu, "Wavelet transform-based methods for removal of ground clutter and denoising the radar wind profiler data," *IET Signal Process.*, vol. 9, no. 5, pp. 440–448, Jul. 2015.

[3] Y. Shen, S. Q. Lou, and X. Wang, "Novel estimation method of point spread function based on kalman filter for accurately evaluating real optical properties of photonic crystal fibers," *Appl. Opt.*, vol. 53, no. 9, pp. 1838–1845, 2014.

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

[5] T. Y. Zeng, X. L. Li, and M. Ng, "Alternating minimization method for total variation based wavelet shrinkage model," *Commun. Comput. Phys.*, vol. 8, no. 5, pp. 976–994, Nov. 2010.

[6] U. Kamilov, E. Bostan, and M. Unser, "Wavelet shrinkage with consistent cycle spinning generalizes total variation denoising," *IEEE Signal Process. Lett.*, vol. 19, no. 4, pp. 187–190, Apr. 2012.

[7] J. F. Cai, B. Dong, S. Osher, and Z. W. Shen, "Image restoration: Total variation, wavelet frames, and beyond," *J. Amer. Math. Soc.*, vol. 25, pp. 1033–1089, Oct. 2012.

[8] S. Durand and M. Nikolova, "Denoising of frame coefficients using L1 data-fidelity term and edge-preserving regularization," *Multiscale Model. Simul.*, vol. 6, no. 2, pp. 547–576, May 2007.

[9] X. H. Wang, Y. N. Liu, H. W. Zhang, and L. L. Fang, "A total variation model based on edge adaptive guiding function for remote sensing image de-noising," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 34, pp. 89–95, Feb. 2015.

[10] Y. Ding and I. W. Selesnick, "Artifact-free wavelet denoising: Non-convex sparse regularization, convex optimization," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1364–1368, Sep. 2015.

[11] Y. W. Fang, X. M. Huo, and Y. W. Wen, "An algorithm for the proximity operator in hybrid TV-wavelet regularization, with application to MR image reconstruction," *East Asian J. Appl. Math.*, vol. 4, no. 1, pp. 21–34, Feb. 2014.

[12] J. Tian, L. Chen, and L. Ma, "A wavelet-domain non-parametric statistical approach for image denoising," *IEICE Electron. Express*, vol. 7, no. 18, pp. 1409–1415, 2010.

[13] J. Han, J. Yue, Y. Zhang, and L. F. Bai, "Local sparse structure denoising for low-light-level image," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5177–5192, Dec. 2015.

[14] Y. N. Zhu, M. L. Zhao, Y. S. Zhao, H. W. Li, and P. Zhang, "Noise reduction with low dose CT data based on a modified ROF model," *Opt. Express*, vol. 20, no. 16, pp. 17987–18004, Jul. 2012.

[15] C. H. Zheng, Y. F. Hou, and J. Zhang, "Improved sparse representation with low-rank representation for robust face recognition," *Neurocomputing*, vol. 198, pp. 114–124, Jul. 2016.

[16] H. S. Bhadauria and M. L. Dewal, "Medical image denoising using adaptive fusion of curvelet transform and total variation," *Comput. Elect. Eng.*, vol. 39, no. 5, pp. 1451–1460, Jul. 2013.

[17] H. H. Lari and A. Gholami, "Curvelet-TV regularized Bregman iteration for seismic random noise attenuation," *J. Appl. Geophys.*, vol. 109, pp. 233–241, Oct. 2014.

[18] G. R. Easley, D. Labate, and F. Colonna, "Shearlet-based total variation diffusion for denoising," *IEEE Trans. Image Process.*, vol. 18, no. 2, pp. 260–268, Feb. 2009.

[19] B. A. Baudes and J. M. Coll, "A non-local algorithm for image denoising," in *Proc. 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, vol. 2, pp. 60–65.

[20] A. Baudes, B. Coll, and J. M. Morel, "On image denoising method," *SIAM Multiscale Model Simul*., vol. 4, no. 2, pp. 490–530, 2004.

[21] S. Kindermann, S. Osher, and P. W. Jones, "Deblurring and denoising of images by nonlocal functionals," *Multiscale Model. Simul*., vol. 4, no. 4, pp. 1091–1115, Jan. 2005.

[22] G. Gilboa and S. Osher, "Nonlocal linear image regularization and supervised segmentation," *SIAM Multiscale Model. Simul*., vol. 6, no. 2, pp. 595–630, May 2007.

[23] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *SIAM Multiscale Model. Simul*., vol. 7, no. 3, pp. 1005–1028, Jan. 2008.

[24] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 253–276, Jan. 2010.

[25] G. Vaksman, M. Zibulevsky, and M. Elad, "Patch-ordering as a regularization for inverse problems in image processing," *SIAM J. Imag. Sci.*, vol. 9, no. 1, pp. 287–319, 2016.

[26] P. Getreuer, "Total variation deconvolution using split Bregman," *Image Process. Line*, vol. 2, pp. 158–174, Jul. 2012.

[27] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 492–526, 2010.

[28] P. Rodriguez and B. Wohlberg, "Performance comparison of iterative reweighting methods for total variation regularization," in *Proc. 2014. Int. Conf. Image Process.*, Paris, France, Oct. 27–30, 2014, pp. 1758–1762.

[29] I. Bayram and M. E. Kamasak, "Directional total variation," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 781–784, Dec. 2012.

[30] A. Buades, B. Coll, and J. M Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2005, vol. 2, pp. 60–65.

[31] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 479–486.

[32] Y. Lou, T. Zeng, S. Osher, and J. Xin, "A weighted difference of anisotropic and isotropic total variation model for image processing," *SIAM J. Imag. Sci.*, vol. 8, no. 3, pp. 1798–1823, 2015.

[33] J. Li, Q. Yuan, H. Shen, and L. Zhang, "Hyperspectral image recovery employing a multidimensional nonlocal total variation model," *Signal Process.*, vol. 111, pp. 230–248, 2015.

[34] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral–spatial adaptive total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3660–3677, Oct. 2012.