

SAR Image Speckle Reduction Based on Nonconvex Hybrid Total Variation Model

Yuli Sun¹, Lin Lei¹, Dongdong Guan¹, Xiao Li¹, and Gangyao Kuang, *Senior Member, IEEE*

Abstract—Speckle noise inherent in synthetic aperture radar (SAR) images seriously affects the visual effect and brings great difficulties to the postprocessing of the SAR image. Due to the edge-preserving feature, total variation (TV) regularization-based techniques have been extensively utilized to reduce the speckle. However, the strong scatters in SAR image with radiometry several orders of magnitude larger than their surrounding regions limit the effectiveness of TV regularization. Meanwhile, the ℓ_1 -norm first-order TV regularization sometimes causes staircase artifacts as it favors solutions that are piecewise constant, and it usually underestimates high-amplitude components of image gradient as the ℓ_1 -norm uniformly penalizes the amplitude. To overcome these shortcomings, a new hybrid variation model, called Fisher–Tippett (FT) distribution- ℓ_p -norm first- and second-order hybrid TVs (HTpVs), is proposed to reduce the speckle after removing the strong scatters. Especially, the FT-HTpV inherits the advantages of the distribution based data fidelity term, the nonconvex regularization, and the higher order TV regularization. Therefore, it can effectively remove the speckle while preserving point scatters and edges and reducing staircase artifacts well. To efficiently solve the nonconvex minimization problem, an iterative framework with a nonmonotone-accelerated proximal gradient (nmAPG) method and a matrix-vector acceleration strategy are used. Extensive experiments on both the simulated and real SAR images demonstrate the effectiveness of the proposed method.

Index Terms—High-order total variation (TV), ℓ_p -norm, speckle, synthetic aperture radar (SAR) image, TV.

I. INTRODUCTION

A. Background

SYNTHETIC aperture radar (SAR) is an advanced active sensor system that collects data of the target day or night and regardless of weather conditions. It can overcome some shortcomings of the optical and infrared systems, such as being easily hindered by adverse atmospheric conditions [1]. This enables SAR to be widely used in ocean monitoring, disaster relief, military development, and many other fields.

Manuscript received December 13, 2019; revised March 23, 2020 and June 8, 2020; accepted June 10, 2020. Date of publication July 8, 2020; date of current version January 21, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61701508. (Yuli Sun and Lin Lei contributed equally to this work.) (Corresponding author: Lin Lei.)

Yuli Sun, Lin Lei, Xiao Li, and Gangyao Kuang are with the College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China (e-mail: sunyuli@mail.ustc.edu.cn; alaleilin@163.com; lxcherishm@126.com; kuangyeats@hotmail.com).

Dongdong Guan is with the High-Tech Institute of Xi'an, Xi'an 710025, China (e-mail: guandd@hotmail.com).

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2020.3002561

Unfortunately, due to the nature of coherent imaging system, SAR is inherently affected by speckle noise [2], which strongly degrades the appearance of images visually and diminishes the performance of subsequent automated tasks, such as target detection [3], change detection [4], and segmentation and classification [5]. Therefore, it is very important to reduce the speckle noise as a preprocessing step in the SAR image applications.

Over the past few decades, researchers have made great efforts on SAR despeckling, and a large number of methods have been proposed, such as the Lee filter [6], the Kuan filter [7], the Frost filter [8], and the maximum *a posteriori* (MAP) filter [9]. However, these traditional spatial domain methods often degrade the spatial resolution of the images and tend to oversmooth the features, such as edges and textures. Then, transform domain filters have been developed and have achieved remarkable performance in recent years, such as wavelet transform [10], [11], curvelet transform [12], [13], and shearlet transform [14], [15]. The procedure of transform domain filtering can be generally divided into three steps: transforming the original images, estimating the noise-free coefficients, and obtaining the denoised images by inverse-transforming from the processed coefficients. Although the transform domain methods can effectively suppress the speckle, they still have weaknesses on the backscatter preservation in homogeneous regions, details preservation in heterogeneous regions, and may cause pixel distortion and generate artificial artifacts. This is mainly due to the inherent disadvantage of the transform domain and regardless of some useful local or global characteristics of the image.

In order to overcome the abovementioned issues, several methods have been proposed by using the nonlocal (NL) means approach, which exploits the self-similarity of the image by selecting similar pixels in an extended search window based on patchwise similarity. Deledalle *et al.* [16] proposed the probabilistic patch-based (PPB) algorithm that derives the similarity between noisy patches from the noise distribution and introduces an iterative method to refine the weights. The block-matching 3-D for SAR images (SAR-BM3D) [17] modifies the major processing steps of the BM3D algorithm [18] according to the peculiarities of SAR image, which chooses the similarity measure by taking into account the probabilistic noise distribution of speckle and adopts the wavelet shrinkage in the 3-D domain by using a linear minimum mean-square error (MMSE)

estimation approach. Recently, the nonlocal low-rank model (NLRM)-based despeckling algorithms have been proposed by researchers [19], [20]. Guan *et al.* [21] proposed a low-rank minimization model based on the Fisher–Tippett (FT) distribution and the weighted nuclear norm to recover the underlying low-rank component from the patch group matrix. The multiscale NLRM is further developed by simultaneously exploring the NLRM and multiscale prior by selecting similar patches from different scales of the SAR image [22]. Despite their good performances, the NL frameworks require the computation of a large number of patch similarity measures, which makes the computational complexity very high.

Recently, a deep learning approach has shown the high capability of data denoising, and its application to SAR despeckling has also been exploited. Because of the outstanding capabilities in efficiently learning task-specific filters, the convolutional-neural-networks (CNNs) have been recently used in SAR image despeckling and gain good performance with respect to state-of-the-art methods, both in terms of the signal-to-noise ratio (SNR) and feature preservation, and in terms of visual effect [23]–[25]. However, there are two major flaws in these deep learning-based methods: the process of training networks is time-consuming, and it requires a lot of pairs of speckle-free/speckled SAR images to construct training set, which yields a high cost of manual operation in practice. In particular, even for experts, it is also very difficult to prepare a speckle free version of an SAR image.

At the same time, there are some algorithms using other additional information to remove speckles. For example, the ratio-based multitemporal SAR image denoising algorithm uses the ratio image provided by the ratio between an image and the temporal mean of stack [26]. With multitemporal images, spatial and temporal information can be combined in the process of denoising, which makes it possible to improve the speckle suppression effects by comparing it with processing a single image. The optical-guided NL despeckling algorithm leverages information drawn from coregistered optical imagery, which can provide additional information to support the speckle removal process as they share some important structure information [27], [28].

Another popular approach for SAR image despeckling is based on variational models, which minimizes some appropriate energy functions, consisting of a regularization term (also called image prior), and a data fidelity term. Among them, the Markov random field (MRF) is often used to integrate a prior term. For a range of degradation mechanisms, including blurring, nonlinear deformations, and multiplicative or additive noise, the posterior distribution is an MRF with a structure akin to the image model [29]. As a powerful approach of stochastic models, MRF has shown good performance in SAR image processing, such as contextual-based segmentation [30] and SAR image despeckling [31]–[33]. Xie *et al.* [31] incorporated the wavelet Bayesian estimation technique and MRF-based image regularization for SAR speckle reduction, where the MRF is used to characterize the intrascale contextual dependence of wavelet. The Gaussian MRF (GMRF)-based contextual analysis and the Bayesian-framework-based pixelwise analysis are combined in [32], which is further extended to the

multipolarized SAR despeckling and shows effective performance [33]. The higher order MRF-based variational model has also been investigated by researchers for multiplicative noise reduction [34], which is defined by a set of linear filters.

The total variation (TV) model is another widely used variation model for its good performance in smoothing homogeneous regions while preserving edges, which uses the ℓ_1 -norm of the gradient-magnitude images as the regularization term and minimizes the following energy function:

$$\min_{\tilde{\mathbf{X}}} \{E(\tilde{\mathbf{X}}) = D(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) + \lambda \psi_{\text{TV}}(\tilde{\mathbf{X}})\} \quad (1)$$

where $\tilde{\mathbf{Y}} \in \mathbb{R}^S$ is the vectorized observation SAR image with S pixels, $\tilde{\mathbf{X}} \in \mathbb{R}^S$ is the image to be restored, and $\lambda > 0$ is the tradeoff parameter. The data fidelity term $D(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ represents the information of the observed image, which makes the processed image close to the observed image, and the regularization term $\psi_{\text{TV}}(\tilde{\mathbf{X}})$ represents the smoothness of the image, which makes the gradient-magnitude image to be sparse. The link between the TV regularization and sparse representation is given in [35]. By adopting different fidelity and regularization strategies, different TV models are evolved, for example, the ℓ_2 -norm data fidelity term plus the ℓ_1 -norm TV regularization [36], the nonquadratic data fidelity term (the log-likelihood) plus the ℓ_1 -norm TV regularization [37], and the ℓ_2 -norm data fidelity term plus the ℓ_p -norm TV [38]. Although the TV regularization admits many desirable properties, such as the most notable sharp edge-preserving feature, it still has some shortcomings as it favors solutions that are piecewise constant, which will cause some undesired artifacts, such as the so-called staircasing artifact. To overcome the staircase effect, some researchers replace the TV by total generalized variation (TGV), which has been proven to be able to eliminate the staircase effect by being aware of higher order smoothness [39]–[41]. This higher order TV regularization has been also introduced to SAR despeckling and achieves good performance [42], [43].

As mentioned in [44], an excellent SAR despeckling algorithm must satisfy the following characteristics: 1) reducing speckle in homogeneous areas; 2) preserving details of SAR image such as edges, textures, and point targets; 3) radiometric preservation; and 4) artifact-free. In pursuit of this goal, this article presents a classification-based SAR image despeckling method with hybrid TV regularization (called cFT-HTpV for short), which is a combination of first-order TV and high-order TV.

With the assumption that the gradient-magnitude of the image is sparse, the TV regularization can effectively reduce the noise and preserve the sharp edges. However, this advantage of TV regularization in the SAR image is limited. Due to the inherent scattering mechanism of the SAR image, there are always strong scattering points in the SAR image, especially in urban areas, which having radiometry several orders of magnitude larger than their neighborhood pixels. These strong scatters limit the effectiveness of TV regularization in two aspects: first, they break the assumption that an image is made of regions of constant radiometry; second, they bring much higher gradients than other pixels and occupy a large

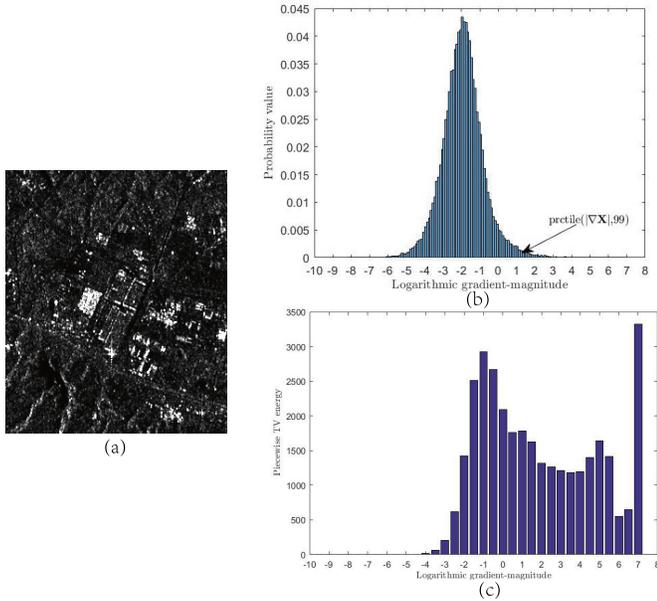


Fig. 1. Disequilibrium of TV functions in the SAR image. (a) SAR image. (b) Gradient-magnitude histogram. (c) Energy distribution of TV function.

part of TV function value, which will reduce the denoising effect on other regions. Fig. 1(a)¹ shows a single-look complex (SLC) SAR image acquired by Sentinel-1, whose gradient-magnitude histogram is shown in Fig. 1(b). The biggest gradient-magnitude is 956.36, and only 1% gradient amplitude values are larger than 3.84, whereas they contribute almost half the energy of TV function, as shown in Fig. 1(c). This extreme imbalance makes it difficult to smooth the homogeneous areas and greatly reduces the denoising effect of TV regularization. To address this challenge, we preliminarily screen out the strong scattering points in the SAR image by using the ratio detectors [45], then despeckle other nonstrong scattering points, and finally add the strong scattering points back to the despeckled image. This idea of step-by-step processing is also used in the recent despeckling method in [46], which is based on complex wavelet shrinkage and NL principal component analysis (PCA) filtering. Meanwhile, in order to reduce the influence of strong scatters, another type of SAR image despeckling method is proposed in [47] by decomposing the radar scene into two components: a background (with TV regularization) and strong scatterers (with low ℓ_0 -pseudonorm).

After obtaining the classification result, we need to use the energy minimization model to remove the speckle, where the data fidelity term is based on the FT distribution and the regularization term is based on the ℓ_p -norm of the hybrid TV (HTpV). On the one hand, because it is based on a noise distribution model that is closer to the real situation, it can make better use of the original image information than the ℓ_2 -norm data fidelity term, which corresponds to the Gaussian distribution. On the other hand, the first-order-based TV model will yield a serious staircase effect, but the higher order

TV regularization can alleviate this staircase effect; however, it is not as good at preserving the edge as the first-order-based TV, as shown in Fig. 2. Motivated by the advantages of first-order and higher order TV regularization, HTpV uses the first- and second-order hybrid TV regularizations to achieve the complementary effect. Meanwhile, because the ℓ_1 -norm regularization function tends to underestimate high-amplitude components of image gradient as it uniformly penalizes the amplitude, HTpV uses the nonconvex ℓ_p -norm function instead of the ℓ_1 -norm regularization to address this issue, which has a superior performance in noise-removing and edge-preserving than the ℓ_1 -norm regularization model.

B. Contribution

The main contributions of this article are summarized as follows.

- 1) We use a step-by-step despeckling framework. First, we divide the image into two classes of strong scattering points and others, and then, the despeckling operations are performed separately.
- 2) We propose an FT-HTpV despeckling model, which can inherit the advantages of noise distribution model, the nonconvex regularization, and the first- and second-order hybrid TV regularizations.
- 3) In order to solve the FT-HTpV minimization, an iterative framework with a nonmonotone-accelerated proximal gradient nmAPG) method [48] and a matrix-vector acceleration strategy is used. Extensive experiments on both the simulated and real SAR images demonstrate the effectiveness of the proposed method.

C. Outline and Notation

The overall structure of this article is given as follows. Section II gives the new cFT-HTpV-based despeckling algorithm. Section III presents the numerical results. In the end, we provide our conclusion in Section IV.

Here, we define our notation. We define the ℓ_p -norm of the vector $\mathbf{x} \in \mathbb{R}^N$ as $\|\mathbf{x}\|_p = (\sum_n |x_n|^p)^{(1/p)}$. Especially, we define ℓ_1 -, ℓ_2 -, and ℓ_∞ -norms of \mathbf{x} as $\|\mathbf{x}\|_1 = \sum_n |x_n|$, $\|\mathbf{x}\|_2 = (\sum_n |x_n|^2)^{1/2}$, and $\|\mathbf{x}\|_\infty = \max_n |x_n|$, respectively. Given a matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$, \mathbf{A}^T is defined as the transpose of \mathbf{A} , and $(\mathbf{A}\mathbf{x})_m$ is defined as the component m of $\mathbf{A}\mathbf{x}$. \mathbf{I}_N represents an $N \times N$ identity matrix, and $\mathbf{1}_{M \times N}$ represents $M \times N$ -dimensional matrix with all elements being 1. $\langle \cdot, \cdot \rangle$ denotes the inner product, \odot denotes the Hadamard product, and \otimes denotes the Kronecker product.

II. PROPOSED DESPECKLING METHOD

Let $\tilde{\mathbf{X}} = \{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_S\} \in \mathbb{R}^S$ be the theoretical reflectivity or the 2-D $M \times N$ noise-free SAR intensity image with $S = M \times N$. If $\tilde{\mathbf{N}} = \{\tilde{n}_1, \tilde{n}_2, \dots, \tilde{n}_S\}$ is the fully developed speckle, then the observed SAR image $\tilde{\mathbf{Y}} = \{\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_S\}$ is related to $\tilde{\mathbf{X}}$ by the well accepted multiplicative model as $\tilde{\mathbf{Y}} = \tilde{\mathbf{X}} \odot \tilde{\mathbf{N}}$. To obtain the additive noise model, we apply the logarithmic transform on both sides of this multiplicative model, and then, we have

$$\mathbf{Y} = \mathbf{X} + \mathbf{N} \quad (2)$$

¹Data processed by ESA, <http://www.copernicus.eu/>

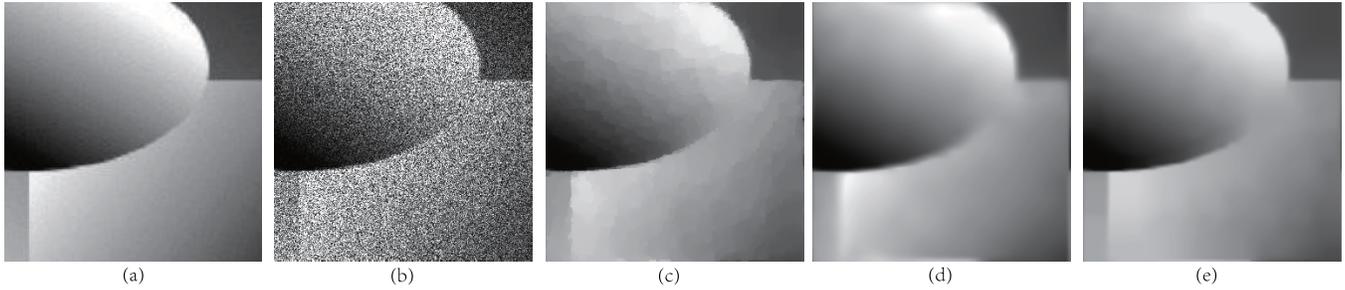


Fig. 2. Image despeckling results of different TV regularizations. (a) Original image. (b) Noisy image with speckle noise. (c) Result of first-order-based TV. (d) Result of second-order-based TV. (e) Result of first- and second-order hybrid TVs.

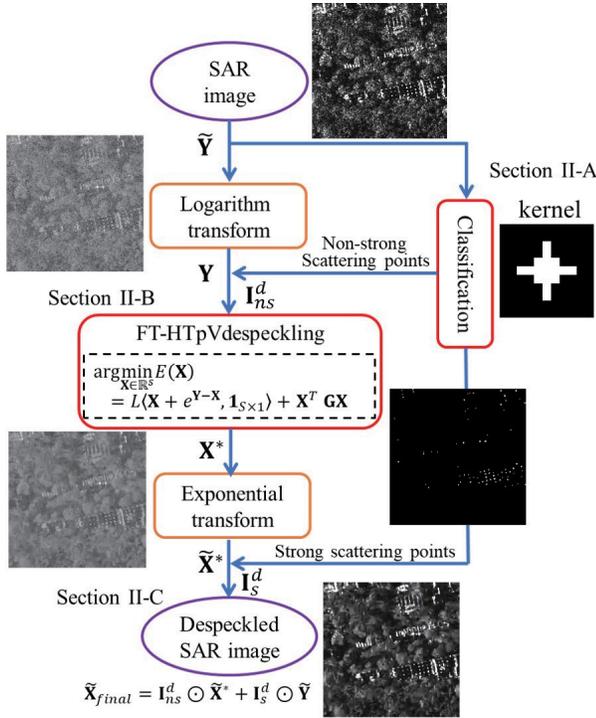


Fig. 3. Framework of the proposed cFT-HTpV method.

where \mathbf{Y} , \mathbf{X} , and \mathbf{N} are the logarithmic transform of $\tilde{\mathbf{Y}}$, $\tilde{\mathbf{X}}$, and $\tilde{\mathbf{N}}$, respectively. Then, the despeckling problem can be transformed into how to recover \mathbf{X} from \mathbf{Y} . The proposed method consists of three steps: 1) strong scattering points' classification; 2) FT-HTpV despeckling; and 3) image fusion of strong scattering points. The framework of cFT-HTpV is illustrated in Fig. 3.

A. Classification Strategy

Due to the inherent strong scattering point of the SAR image, the denoising effect of TV regularization will be reduced as the strong scattering point does not satisfy the assumption of piecewise smoothness, and it will dominate the energy of TV function. Thus, we need to screen out these strong scattering points first. There are many methods that can be used to detect the strong scattering points, such as the common ratio detectors [45] and the constant false alarm rate (CFAR)-based detection method [49]. In this article, we use the ratio detector for simplicity and efficiency.

The 11×11 kernel shown in Fig. 3 is used to identify the strong scattering points with the dark pixels being 0 and white

pixels being 1. For each pixel in the image $\tilde{\mathbf{Y}}$, we multiply its local neighborhood with the kernel point by point and then add the results to get the ratio detector R_p . Intuitively, R_p can be computed as $R_p = R_w/R_d$, where R_w and R_d represent the sum of values in white and dark pixels in the defined 11×11 kernel, respectively. It can easily be found that for general pixels (nonstrong scattering points), R_p is very small as there are only 17 white pixels in the 11×11 kernel. However, for the strong scattering points, because they are several orders of magnitude higher than the surrounding pixels, R_p will still be large. Then, we can classify the SAR image into two classes with a threshold R_T : if R_p is larger (or equal) than R_T , the pixel will be assigned as strong scattering point; otherwise, the pixel will be considered as nonstrong scattering point. Then, we can obtain the strong scattering point index matrix $\mathbf{I}_s^d \in R^{M \times N}$ with the eight neighborhoods of (i, j) to be 1 if the (i, j) th pixel is classified as strong scattering point, and other elements to be 0. Similarly, we can obtain the nonstrong scattering point index matrix $\mathbf{I}_{ns}^d \in R^{M \times N}$ with $\mathbf{I}_s^d + \mathbf{I}_{ns}^d = \mathbf{1}_{M \times N}$. After this classification, the strong scattering point will be excluded from the next FT-HTpV despeckling process and, finally, added back to the despeckled image.

B. Proposed FT Distribution-Based ℓ_p -Norm Hybrid TV Model

1) *FT Distribution-Based Data Fidelity Term*: From the view point of the Bayesian MAP, restoring the clean image $\tilde{\mathbf{X}}$ from the observed noisy image $\tilde{\mathbf{Y}}$ is equivalent to solving optimization problems as

$$\begin{aligned} \hat{\mathbf{X}} &= \arg \max_{\tilde{\mathbf{X}}} P(\tilde{\mathbf{X}}|\tilde{\mathbf{Y}}) = \arg \max_{\tilde{\mathbf{X}}} \frac{P(\tilde{\mathbf{Y}}|\tilde{\mathbf{X}})P(\tilde{\mathbf{X}})}{P(\tilde{\mathbf{Y}})} \\ &= \arg \max_{\tilde{\mathbf{X}}} P(\tilde{\mathbf{Y}}|\tilde{\mathbf{X}})P(\tilde{\mathbf{X}}). \end{aligned} \quad (3)$$

Taking the negative logarithm, (3) can be rewritten as

$$\begin{aligned} \hat{\mathbf{X}} &= \arg \min_{\tilde{\mathbf{X}}} \{-\log P(\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}) - \log P(\tilde{\mathbf{X}})\} \\ &= \arg \min_{\tilde{\mathbf{X}}} \{D(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) + G(\tilde{\mathbf{X}})\} \end{aligned} \quad (4)$$

where the first term $D(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ is the data fidelity term and the second term $G(\tilde{\mathbf{X}})$ is the prior term [50]. From (4), we can find that the form of $D(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ depends on the noise distribution. Considering the classical additive white Gaussian noise (AWGN) model as $\tilde{\mathbf{Y}} = \tilde{\mathbf{X}} + \tilde{\mathbf{N}}$, where $\tilde{\mathbf{N}}$ is the zero-mean AWGN with standard deviation $\sigma_{\tilde{\mathbf{N}}}$. Then, minimizing the

log-likelihood data fidelity term $D(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ amounts to

$$\min_{\tilde{\mathbf{X}}} \left\{ -\log p_{\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}}(\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}) = \frac{1}{\sigma_{\tilde{\mathbf{N}}}^2} \|\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}\|_2^2 + C_1 \right\} \quad (5)$$

where $p_{\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}}(\tilde{\mathbf{Y}}|\tilde{\mathbf{X}})$ denotes the conditional probability density function (pdf) and C_1 is a constant (irrelevant for estimation purposes).

From (5), we can find that the ℓ_2 -norm type of data fidelity term $D(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ is suitable for the Gaussian distribution. However, the SAR image generally follows a heavy-tail non-Gaussian distribution [44], so we need to change $D(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ to fit the noise distribution in SAR image. The statistical properties of speckle have been widely studied, and they are the topic of a large body of the literature. Here, we use the widely recognized and used the Gamma distribution with unit mean [44]. Reconsidering the multiplicative model $\tilde{\mathbf{Y}} = \tilde{\mathbf{X}} \odot \tilde{\mathbf{N}}$, the pdf of $\tilde{\mathbf{N}}$ under the assumption that pixels are statistically independent is given by

$$p_{\tilde{\mathbf{N}}}(\tilde{\mathbf{N}}) = \prod_{i=1}^S \frac{L^L \tilde{n}_i^{L-1} e^{-L\tilde{n}_i}}{\Gamma(L)}, \quad \tilde{n}_i \geq 0 \quad (6)$$

where $\Gamma(\cdot)$ denotes the Gamma function and L is the equivalent number of looks (ENL).

For the additive noise model (2) that obtained by the logarithmic transformation, the pdf of the random \mathbf{N} is given by [51]

$$p_{\mathbf{N}}(\mathbf{N}) = \prod_{i=1}^S p_{n_i}(e^{n_i}) e^{n_i} = \prod_{i=1}^S \frac{L^L}{\Gamma(L)} e^{Ln_i} e^{-Le^{n_i}} \quad (7)$$

which is known as the FT distribution. Then, we can obtain

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X}) = \prod_{i=1}^S \frac{L^L}{\Gamma(L)} e^{L(y_i - x_i)} e^{-Le^{(y_i - x_i)}}. \quad (8)$$

Similar to (5), we can obtain the data fidelity term $D(\mathbf{X}, \mathbf{Y})$, which is based on the FT distribution by taking the negative logarithm of (8)

$$D(\mathbf{X}, \mathbf{Y}) = -\log p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X}) = L \sum_{i=1}^S (x_i + e^{y_i - x_i}) + C_2 \quad (9)$$

where C_2 is a constant. Regardless of this irrelevant constant for estimation purposes, we have the following data fidelity term (10), which is strictly convex:

$$D(\mathbf{X}, \mathbf{Y}) = L \sum_{i=1}^S (x_i + e^{y_i - x_i}). \quad (10)$$

2) ℓ_p -Norm Hybrid TV Model: Since we deal with the discrete formulation of the image, we consider the discrete forms of the first-order TV and higher-order TV. The standard TV (first-order TV) regularizer can be expressed as

$$\begin{aligned} \psi_{TV}(\mathbf{X}) &= \|\|\nabla\mathbf{X}\|_1 = \sum_{i=1}^M \sum_{j=1}^N \|\nabla\mathbf{X}\|_{i,j} \\ &= \sum_{i=1}^M \sum_{j=1}^N \left(\left| (\mathbf{D}_h^+\mathbf{X})_{i,j} \right|^p + \left| (\mathbf{D}_v^+\mathbf{X})_{i,j} \right|^p \right)^{\frac{1}{p}} \end{aligned} \quad (11)$$

where \mathbf{D}_h^+ and \mathbf{D}_v^+ denote the forward horizontal and vertical difference operators with periodic boundary condition (assuming that \mathbf{X} is periodically extended), respectively. $(\nabla\mathbf{X})_{i,j} := \left((\mathbf{D}_h^+\mathbf{X})_{i,j}, (\mathbf{D}_v^+\mathbf{X})_{i,j} \right)$ is the first-order difference vector at pixel (i, j) , and $|\nabla\mathbf{X}|$ is the gradient-magnitude image with $|\nabla\mathbf{X}|_{i,j} := \|\nabla\mathbf{X}\|_{i,j}$, where the ℓ_p -norm could either be $p = 2$ corresponding to the isotropic TV or $p = 1$ corresponding to the anisotropic TV.

The discrete second-order TV regularizer can be expressed as

$$\begin{aligned} \psi_{s-TV}(\mathbf{X}) &= \|\|\nabla^2\mathbf{X}\|_1 = \sum_{i=1}^M \sum_{j=1}^N \|\nabla^2\mathbf{X}\|_{i,j} \\ &= \sum_{i=1}^M \sum_{j=1}^N \left(\left| (\mathbf{D}_{hh}^{++}\mathbf{X})_{i,j} \right|^2 + \left| (\mathbf{D}_{hv}^{++}\mathbf{X})_{i,j} \right|^2 \right. \\ &\quad \left. + \left| (\mathbf{D}_{vh}^{++}\mathbf{X})_{i,j} \right|^2 + \left| (\mathbf{D}_{vv}^{++}\mathbf{X})_{i,j} \right|^2 \right)^{\frac{1}{2}} \end{aligned} \quad (12)$$

where $|\nabla^2\mathbf{X}|$ is the second-order derivative magnitude image with $|\nabla^2\mathbf{X}|_{i,j} := \|\nabla^2\mathbf{X}\|_{i,j}$, and $(\nabla^2\mathbf{X})_{i,j} := \left((\mathbf{D}_{hh}^{++}\mathbf{X})_{i,j}, (\mathbf{D}_{hv}^{++}\mathbf{X})_{i,j}, (\mathbf{D}_{vh}^{++}\mathbf{X})_{i,j}, (\mathbf{D}_{vv}^{++}\mathbf{X})_{i,j} \right)$ is the second-order difference matrix at pixel (i, j) with

$$\begin{aligned} (\mathbf{D}_{hh}^{--}\mathbf{X})_{i,j} &:= (\mathbf{D}_h^-(\mathbf{D}_h^+\mathbf{X}))_{i,j} \\ (\mathbf{D}_{hv}^{++}\mathbf{X})_{i,j} &:= (\mathbf{D}_h^+(\mathbf{D}_v^+\mathbf{X}))_{i,j} \\ (\mathbf{D}_{vh}^{++}\mathbf{X})_{i,j} &:= (\mathbf{D}_v^+(\mathbf{D}_h^+\mathbf{X}))_{i,j} \\ (\mathbf{D}_{vv}^{++}\mathbf{X})_{i,j} &:= (\mathbf{D}_v^+(\mathbf{D}_v^+\mathbf{X}))_{i,j} \end{aligned} \quad (13)$$

where \mathbf{D}_h^- and \mathbf{D}_v^- denote the backward horizontal and vertical difference operators with periodic boundary condition, respectively.

The ℓ_1 -norm TV penalizes the image gradient uniformly as its potential function is defined as a linear function $\psi(t) = |t|$, which may lead to underestimation of high-amplitude components. In order to address this issue, the ℓ_p -norm TV (TpV) is proposed as

$$\psi_{TpV}(\mathbf{X}) = \|\|\nabla\mathbf{X}\|_p^p = \sum_{i=1}^M \sum_{j=1}^N (|\nabla\mathbf{X}|_{i,j})^p. \quad (14)$$

From (14), we can find that the TpV encourages penalty of minor changes in \mathbf{X} while decreases the penalty of large variations in \mathbf{X} by comparing with the ℓ_1 -norm TV.

In order to combine the advantages of the first- and second-order TV and leverage the advantage of the nonconvex ℓ_p -norm regularization model, we propose the hybrid regularization model HTpV (15), as shown at the bottom of the next page. In (15), $0 < \beta < 1$ is a parameter to balance the first-order variation regularization and the second-order variation regularization. Meanwhile, it can be seen that ψ_{TpV} and ψ_{s-TV} in ψ_{HTpV} [see (15)] are approximate but not the real ℓ_p -norm of $|\nabla\mathbf{X}|$ in (11) and $|\nabla^2\mathbf{X}|$ in (12), respectively. This is because the horizontal and vertical finite differences are

separated in ψ_{HTpV} . However, this type of ψ_{HTpV} (15) can be treated as the ℓ_p -norm of anisotropic hybrid TV regularization.

By combining the FT distribution-based data fidelity term $D(\mathbf{X}, \mathbf{Y})$ in (10), the ℓ_p -norm hybrid TV model ψ_{HTpV} , and the nonstrong scattering point index \mathbf{I}_{ns}^d , we can obtain the FT-HTpV model (16), as shown at the bottom of the page.

3) *Algorithm for the FT-HTpV*: In the Appendix, we simplify the model (16) by defining two $M \times M$ banded circulant matrices $\mathbf{D}_{1,M}$ and $\mathbf{D}_{2,M}$, as in (A.1), using an approximation strategy. Therefore, the model (16) can be rewritten in a matrix-vector form as

$$\min_{\mathbf{X}} \left\{ E(\mathbf{X}) = L \underbrace{\langle \mathbf{X} + e^{\mathbf{Y}-\mathbf{X}}, \mathbf{1}_{S \times 1} \rangle}_{f(\mathbf{X})} + \underbrace{\mathbf{X}^T \mathbf{G} \mathbf{X}}_{g(\mathbf{X})} \right\}. \quad (17)$$

Next, we choose the proximal gradient (PG) framework to solve this minimization problem. Given a symmetric positive semidefinite matrix $\mathbf{H} \in \mathbb{R}^{S \times S}$, we define the quadratic approximation of $E(\mathbf{X})$ at a given point \mathbf{Z} as

$$Q_{\mathbf{H}}(\mathbf{X}, \mathbf{Z}) = g(\mathbf{X}) + f(\mathbf{Z}) + \langle \nabla f(\mathbf{Z}), \mathbf{X} - \mathbf{Z} \rangle + \frac{1}{2} \langle \mathbf{X} - \mathbf{Z}, \mathbf{H}(\mathbf{X} - \mathbf{Z}) \rangle. \quad (18)$$

As the gradient of $f(\mathbf{X})$ does not satisfy the Lipschitz continuity, there is no such \mathbf{H} that satisfies $f(\mathbf{X}) + g(\mathbf{X}) \leq Q_{\mathbf{H}}(\mathbf{X}, \mathbf{Z})$ for any $\mathbf{X}, \mathbf{Z} \in \mathbb{R}^S$. To reduce the gap with this requirement, we treat the abovementioned approximation (18) as the form of the Taylor expansion and ignore the higher order terms, and then, we can set \mathbf{H} to be

$$\mathbf{H} = \text{dexp}^{\mathbf{Y}-\mathbf{Z}} \quad (19)$$

where $\text{dexp}^{\mathbf{Y}-\mathbf{Z}} \in \mathbb{R}^{S \times S}$ is a diagonal matrix with the diagonal elements $(\text{dexp}^{\mathbf{Y}-\mathbf{Z}})_{i,i} = e^{y_i - z_i}$. Then, for any $\mathbf{X}^0 \in \mathbb{R}^S$, the k th iteration of the PG framework [43] for solving (17) is

$$\begin{aligned} \mathbf{X}^{k+1} &= \arg \min_{\mathbf{X}} Q_{\mathbf{H}}(\mathbf{X}, \mathbf{X}^k) \\ &= \arg \min_{\mathbf{X}} \left\{ g(\mathbf{X}) + f(\mathbf{X}^k) + \langle \nabla f(\mathbf{X}^k), \mathbf{X} - \mathbf{X}^k \rangle \right. \\ &\quad \left. + \frac{1}{2} \langle \mathbf{X} - \mathbf{X}^k, \text{dexp}^{\mathbf{Y}-\mathbf{X}^k}(\mathbf{X} - \mathbf{X}^k) \rangle \right\}. \quad (20) \end{aligned}$$

The first-order optimality condition of minimizing $Q_{\mathbf{H}}(\mathbf{X}, \mathbf{X}^k)$ is that

$$\partial g(\mathbf{X}) + L \mathbf{1}_{S \times 1} - L \text{dexp}^{\mathbf{Y}-\mathbf{X}^k} \mathbf{1}_{S \times 1} + \text{dexp}^{\mathbf{Y}-\mathbf{X}^k} (\mathbf{X} - \mathbf{X}^k) = \mathbf{0}. \quad (21)$$

However, as the matrix \mathbf{G} also contains the variable \mathbf{X} , it is not easy to calculate the subgradient of $g(\mathbf{X})$. Here, we use a two-step iterative strategy to calculate the optimal \mathbf{X}^{k+1} of (20).

Step 1: With the n th iteration \mathbf{X}^n , we update the matrix \mathbf{A}^n as

$$\mathbf{A}^n = 2\mathbf{G}^n + \text{dexp}^{\mathbf{Y}-\mathbf{X}^k} \quad (22)$$

where \mathbf{G}^n can be updated by using \mathbf{X}^n in (A.6).

Step 2: Fix the matrix $\mathbf{G} = \mathbf{G}^n$ in $g(\mathbf{X})$, and then, \mathbf{X}^{n+1} can be obtained by solving the linear system

$$\mathbf{A}^n \mathbf{X}^{n+1} = \mathbf{b}^k \quad (23)$$

where the vector \mathbf{b}^k is

$$\mathbf{b}^k = L \text{dexp}^{\mathbf{Y}-\mathbf{X}^k} \mathbf{1}_{S \times 1} - L \mathbf{1}_{S \times 1} + \text{dexp}^{\mathbf{Y}-\mathbf{X}^k} \mathbf{X}^k. \quad (24)$$

As the matrix \mathbf{A}^n in (22) is a nine-point Laplacian matrix that is sparse, real, symmetric, and positive definite, the linear system (23) can be solved efficiently by using iterative solvers, such as the conjugate gradient (CG) method and the minimum residual (MinRes) method. Here, we apply the preconditioned conjugate gradient (PCG) method, which can accelerate CG method by using some preconditioners, such as incomplete Cholesky, Jacobi, and successive overrelaxation. In the experiment, we use the incomplete Cholesky preconditioner.

The proposed FT-HTpV is summarized in Table I, which consists of an external PG iteration for updating $\nabla f(\mathbf{X}^k)$ and $\text{dexp}^{\mathbf{Y}-\mathbf{X}^k}$ in \mathbf{b}^k and an internal iteration for solving subproblem (20). The parameters k_{\max} and n_{\max} are the max numbers of external and internal iterations, respectively. A reasonable termination criterion for the FT-HTpV can be set as $(\|\mathbf{X}^{k+1} - \mathbf{X}^k\|_2) / (\max\{\|\mathbf{X}^k\|_2, 1\}) < \zeta$ for a small positive parameter ζ , which means that there is no longer any appreciable changes in the iteration and the algorithm runs into convergence. Meanwhile, because FT-HTpV uses the PG framework, it can be accelerated by using the nmAPG method [48], which

$$\begin{aligned} &\psi_{HTpV}(\mathbf{X}) \\ &= \underbrace{\beta \sum_{i=1}^M \sum_{j=1}^N \left(|(\mathbf{D}_h^+ \mathbf{X})_{i,j}|^p + |(\mathbf{D}_v^+ \mathbf{X})_{i,j}|^p \right)}_{\psi_{TpV}(\mathbf{X})} + (1 - \beta) \underbrace{\sum_{i=1}^M \sum_{j=1}^N \left(|(\mathbf{D}_{hh}^+ \mathbf{X})_{i,j}|^p + |(\mathbf{D}_{hv}^+ \mathbf{X})_{i,j}|^p + |(\mathbf{D}_{vh}^+ \mathbf{X})_{i,j}|^p + |(\mathbf{D}_{vv}^+ \mathbf{X})_{i,j}|^p \right)}_{\psi_{s-TpV}(\mathbf{X})} \end{aligned} \quad (15)$$

$$\begin{aligned} \min_{\mathbf{X}} \left\{ E(\mathbf{X}) = L \sum_{i=1}^S (x_i + e^{y_i - x_i}) + \lambda \beta \sum_{i=1}^M \sum_{j=1}^N \left(|(\mathbf{I}_{ns}^d \odot \mathbf{D}_h^+ \mathbf{X})_{i,j}|^p + |(\mathbf{I}_{ns}^d \odot \mathbf{D}_v^+ \mathbf{X})_{i,j}|^p \right) \right. \\ \left. + \lambda (1 - \beta) \sum_{i=1}^M \sum_{j=1}^N \left(|(\mathbf{I}_{ns}^d \odot \mathbf{D}_{hh}^+ \mathbf{X})_{i,j}|^p + |(\mathbf{I}_{ns}^d \odot \mathbf{D}_{hv}^+ \mathbf{X})_{i,j}|^p + |(\mathbf{I}_{ns}^d \odot \mathbf{D}_{vh}^+ \mathbf{X})_{i,j}|^p + |(\mathbf{I}_{ns}^d \odot \mathbf{D}_{vv}^+ \mathbf{X})_{i,j}|^p \right) \right\} \quad (16) \end{aligned}$$

TABLE I

MAIN IMPLEMENTATION STEPS OF FT-HTpV FOR SAR DESPECKLING

Algorithm 1. FT-HTpV

Initialization:
 Given \mathbf{Y}_1 , and select $p, \lambda, \beta, \varepsilon$. Initialize \mathbf{X}^0 .

Main iteration loop:
 for $k = 1, 2, \dots, k_{max}$ do
 b-updating: compute \mathbf{b}^k according to equation (24).
 Internal iteration loop:
 for $n = 1, 2, \dots, n_{max}$ do
 G-updating: compute \mathbf{G}^n according to equation (A.6).
 Compute \mathbf{A}^n by $\mathbf{A}^n = 2\mathbf{G}^n + \text{dexp}^{\mathbf{Y} - \mathbf{X}^k}$.
 X-updating: compute \mathbf{X}^{n+1} according to (23) by PCG method.
 End for
 X-updating: $\mathbf{X}^{k+1} = \mathbf{X}^{n+1}$.
 End for

TABLE II

MAIN IMPLEMENTATION STEPS OF FT-HTpV USING nmAPG

Algorithm 2. nmAPG based FT-HTpV

Given \mathbf{Y}_1 , and select $p, \lambda, \beta, \varepsilon$. Initialize $\mathbf{Z}^1 = \mathbf{X}^1 = \mathbf{X}^0$, and $t^1 = t^0 = 1, c^1 = E(\mathbf{X}^1), q^1 = 1, \eta \in [0, 1), \delta > 0$.

for $k = 1, 2, \dots, k_{max}$ do
 $\mathbf{u}^k = \mathbf{X}^k + \frac{t^{k-1}}{t^k}(\mathbf{Z}^k - \mathbf{X}^k) + \frac{t^{k-1}-1}{t^k}(\mathbf{X}^k - \mathbf{X}^{k-1})$
 Calculate $\mathbf{Z}^{k+1} = \arg \min_{\mathbf{X}} Q_{\mathbf{H}}(\mathbf{X}, \mathbf{u}^k)$ in (20) by using the two-step iterative strategy (Internal iteration loop in Algorithm 1).
 if $E(\mathbf{Z}^{k+1}) \leq c^k - \delta \|\mathbf{Z}^{k+1} - \mathbf{u}^k\|_2^2$ then
 $\mathbf{X}^{k+1} = \mathbf{Z}^{k+1}$
 else
 Calculate $\mathbf{v}^{k+1} = \arg \min_{\mathbf{X}} Q_{\mathbf{H}}(\mathbf{X}, \mathbf{X}^k)$ in (20) by using the two-step iterative strategy (Internal iteration loop in Algorithm 1).
 $\mathbf{X}^{k+1} = \begin{cases} \mathbf{Z}^{k+1}, & \text{if } E(\mathbf{Z}^{k+1}) \leq E(\mathbf{v}^{k+1}) \\ \mathbf{v}^{k+1}, & \text{otherwise} \end{cases}$
 End if
 $t^{k+1} = \frac{\sqrt{4(t^k)^2 + 1} + 1}{2}$
 $q^{k+1} = \eta q^k + 1$
 $c^{k+1} = \frac{\eta q^k c^k + E(\mathbf{X}^{k+1})}{q^{k+1}}$
 End for

is an extension of the accelerated PG (APG) method [52] for nonconvex problem. The nmAPG-based FT-HTpV is summarized in Table II, which uses another PG framework as the monitor.

C. Image Fusion of Strong Scattering Points

After obtaining the FT-HTpV despeckling result \mathbf{X}^* , we first transform it into $\tilde{\mathbf{X}}^* = e^{\mathbf{X}^*}$ and then add the classified strong scattering points back to the $\tilde{\mathbf{X}}^*$ to get the final despeckling image as $\tilde{\mathbf{X}}_{final} = \mathbf{I}_{ns}^d \odot \tilde{\mathbf{X}}^* + \mathbf{I}_s^d \odot \tilde{\mathbf{Y}}$.

III. NUMERICAL EXPERIMENTS

In this section, simulations are performed to demonstrate the proposed conclusions and evaluate the performance of the proposed classification-based cFT-HTpV. We apply some state-of-the-art despeckling methods in comparison: SDD with ℓ_1 -norm TV [38], MIDAL [37], building smearing (BS)-TV [47], MAP-S [53], POTDF [54], PPB [16], SAR-BM3D [17], NL-SAR [55], FANS [56], FT-WNNM [21], and SAR-CNN [57]. Among these methods, SDD and MIDAL are TV-based despeckling methods with a quadratic and

log-likelihood data fidelity term, respectively; BS-TV is a TV-based method with the joint estimation of the background and of the strong scatters; MAP-S is a segmentation-based MAP despeckling method in the undecimated wavelet domain; POTDF is a sparse representation-based despeckling method; PPB, SAR-BM3D, NL-SAR, and FANS are NL mean-based despeckling methods; FT-WNNM is an NL rank-based despeckling method; and SAR-CNN is a deep learning-based method based on a CNN. The executable codes of the compared methods are downloaded from the authors' websites, except for BS-TV, MAP-S, and SAR-CNN that are implemented by ourselves. All experiments are performed in MATLAB 2015b running on ASUS laptop with Intel Core i7-8550U CPU and 8 GB of RAM, except for the SAR-CNN implemented in Tensorflow.

A. Parameter Analysis

The main parameters in this algorithm are the classification ratio threshold R_T , the ENL parameter L , the penalty parameter λ , and the balance parameter β .

As shown in Section II-A, the threshold R_T controls the accuracy of strong scattering points detection. From the perspective of detection alone, if R_T is set too large, there will be many false negatives; on the contrary, if this parameter is set very small, there will be many false positives. Generally, this parameter should be determined according to the image structure and noise level. For example, if the imaging region is the forest, grassland, water area, and so on, this R_T can be set relatively large; however, when the imaging region is in the urban area or the target of interest is a strong scattering target, such as vehicles and airplanes, this R_T cannot be set too large. However, from the perspective of the cFT-HTpV process, the value range of this parameter is relatively large. In other words, the algorithm is not very sensitive to this parameter. This is mainly due to that the purpose of the classification is to reduce the proportion of these strong scattering points in the energy of TV function. When there are many missed detections caused by large R_T , the remaining strong scattering points can also be well preserved in the subsequent FT-HTpV process. When there are some false detections caused by small R_T , since its number only accounts for a very small part of the total number of pixels, it has little effect on the whole despeckling process. We recommend setting the value range of this threshold R_T to [0.5, 2]-based on empirical evidence.

The ENL parameter L can be obtained according to prior knowledge or can be estimated from an homogeneous region of $\tilde{\mathbf{Y}}$ as [58]

$$\text{ENL} = S_C \left(\frac{\mu_{\text{HR}}}{\sigma_{\text{HR}}} \right)^2$$

$$S_C = \begin{cases} 1, & \text{intensity image} \\ 4/\pi - 1, & \text{amplitude image} \end{cases} \quad (25)$$

where μ_{HR} is the mean value of the selected homogeneous region and σ_{HR}^2 is the corresponding variance.

The penalty parameter λ controls the strength of regularization and data fidelity. In general, when the input image is contaminated by high-level speckle, we need to choose a

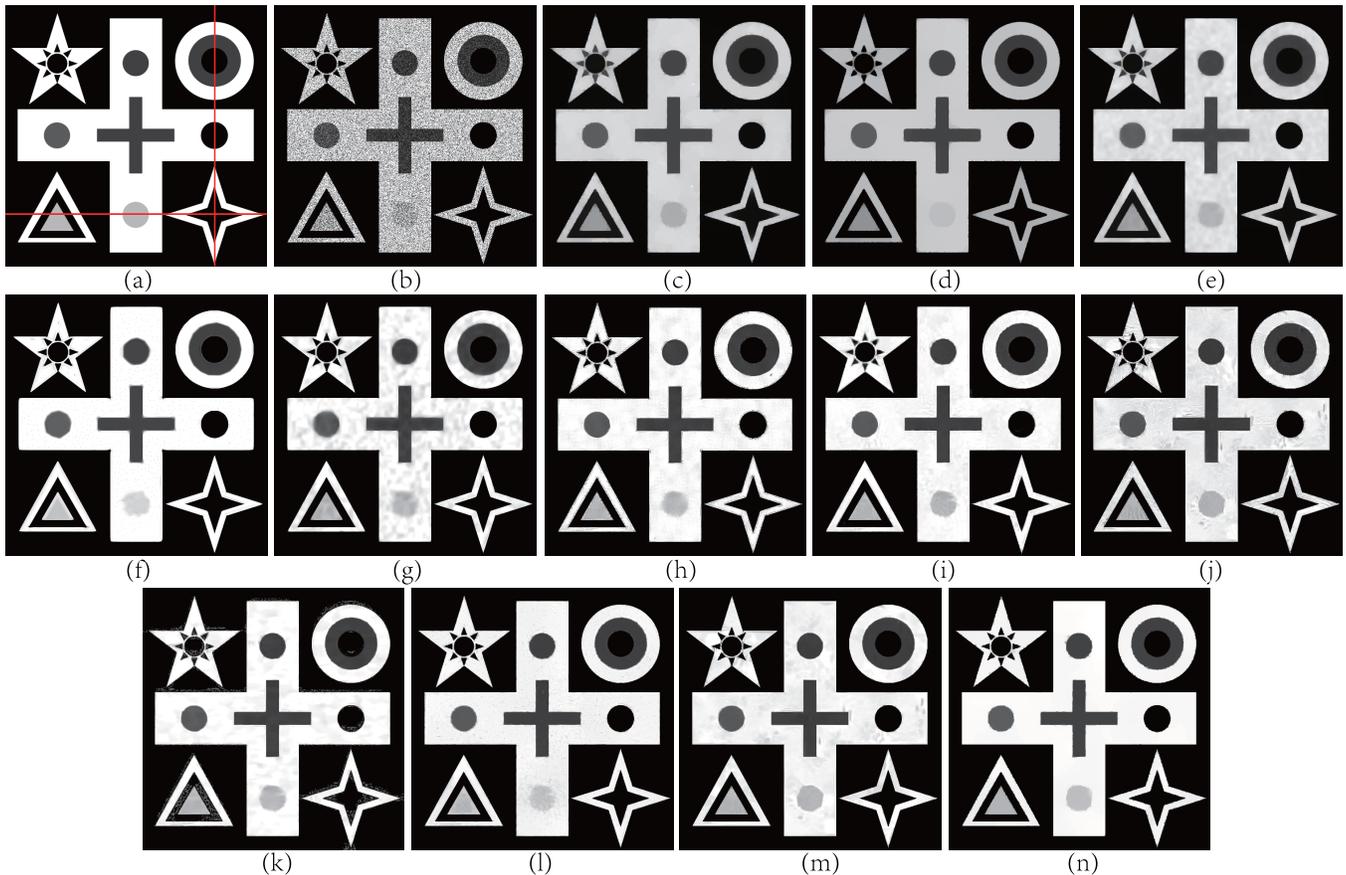


Fig. 4. Despeckling results on Geometric image with size 593×593 . (a) Clean Geometric image. (b) Contaminated image by the speckle noise with $L = 1$. (c) SDD. (d) MIDAL. (e) BS-TV. (f) MAP-S. (g) POTDF. (h) PPB. (i) SAR-BM3D. (j) NL-SAR. (k) FANS. (l) FT-WNNM. (m) SAR-CNN. (n) cFT-HTpV.

large λ to encourage the regularization while relaxing data fidelity. Conversely, when the noise level in the image is low, we need to choose a small λ to penalize the regularization while encouraging fidelity. In light of this, some researchers suggest using adaptive methods to select this parameter during the iterations [59], which has been proved to be effective. In this article, we use the same adaptive method to select λ during the iterations on the basis of the behavior of the objective function with the initialization $\lambda^0 = 0.5$ for the single-look data sets and $\lambda^0 = 0.1$ for the multilook data sets.

The balance parameter $0 < \beta < 1$ is a weighting parameter that balances the first-order variation regularization and the second-order variation regularization. Actually, the model (16) with $\beta = 1$ and $\beta = 0$ is corresponding to the nonconvex first-order TpV and nonconvex second-order TV, respectively. Inspired by the adaptive strategy in [60] and [61], we set β as an edges detection function as $\beta = (\gamma + \|\nabla(G_\sigma * \mathbf{X})\|_2^2) / (1 + \gamma + \|\nabla(G_\sigma * \mathbf{X})\|_2^2)$, where $\gamma > 0$ is a very small positive number added to ensure $\beta < 1$, G_σ is the Gaussian kernel, and σ denotes the standard deviation. When $\|\nabla(G_\sigma * \mathbf{X})\|_2$ is large (large $\|\nabla(G_\sigma * \mathbf{X})\|_2$ corresponds to the edges), the first-order TpV regularization is predominant; thus, the image edges and contours can be well preserved. Conversely, when $\|\nabla(G_\sigma * \mathbf{X})\|_2$ is small (small $\|\nabla(G_\sigma * \mathbf{X})\|_2$ corresponds to smooth areas), the nonconvex

second-order TV regularization is predominant; thus, it can reduce the staircase effects in the smooth regions.

For comparing the methods, we use the default parameters for POTDF, PPB, SAR-BM3D, NL-SAR, FANS, and FT-WNNM in their codes; we vary the regularization parameter λ from 10^{-3} to 20 (with 30 logarithmically equally spaced) for the TV-based methods (SDD, MIDAL, and BS-TV) and then select the best one as the result; we use the four-taps Daubechies' (db4) orthogonal wavelet decomposition with four levels for the MAP-S; and we construct the training set of 300×144 patches (40×40 pixels) for SAR-CNN, which is cut from 250 nature images and 50 SAR images with size 512×512 , and train the SAR-CNN following the steps and parameters in [57].

B. Experiments on Simulated Images

We first test all the methods on simulated images that are generated by adding simulated speckle to noise-free images. Because we already have the comparative noiseless images, we can evaluate the despeckled results with quality measures: the structural similarity index (SSIM) [62] and the peak SNR (PSNR) defined as $\text{PSNR} = 10\log_{10}(255^2/\text{MSE})$ with MSE representing the mean square error between the despeckled image and noise-free image. Fig. 4 presents the comparison between different methods on the piecewise affine Geometric image contaminated by the speckle noise with $L = 1$.

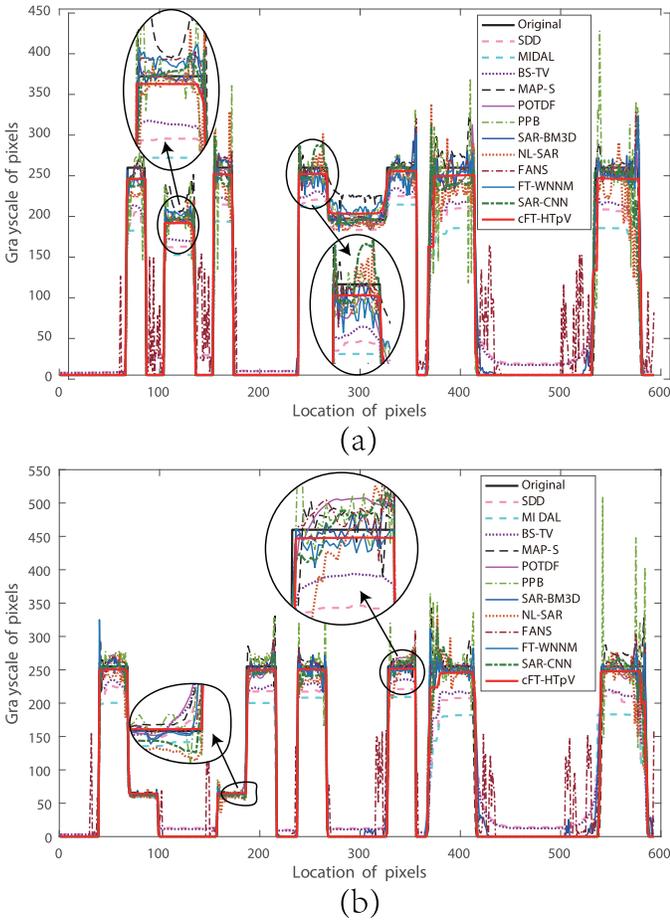


Fig. 5. Image profiles of comparing methods along the horizontal rows and vertical columns on the Geometric image. (a) Horizontal profiles. (b) Vertical profiles.

Fig. 5 displays the comparing despeckling images generated by different methods along the horizontal rows and vertical columns that are marked with red lines in Fig. 4(a). From Fig. 4, we can find that all the methods can remove noise while maintaining sharp edges efficiently. Since the original image is composed of piecewise constant geometric patterns, it is not surprising that the TV-based SDD, MIDAL, and cFT-HTpV can obtain smoother results than other methods, such as POTDF and FANS. By comparing Fig. 4(c)–(e) with Fig. 4(n), it can be shown that the ℓ_p -norm-based cFT-HTpV can effectively alleviate the ℓ_1 -norm TV’s shortcoming of underestimate high-amplitude components. This can also be easily confirmed in Fig. 5, where the error between cFT-HTpV result and true value is much smaller than that of other TV-based methods, such as MIDAL, SDD, and BS-TV. As can be seen from the enlarged curves in Fig. 5, the cFT-HTpV obtains the best speckle removal image, followed by SAR-BM3D. From Figs. 4 and 5, it can be seen that the cFT-HTpV also has advantages in preserving edges when comparing with PPB and FANS.

Figs. 6 and 7 present the comparisons between different methods on Lena and Cameraman, respectively. Both images are contaminated by the speckle noise with $L = 3$. The PSNR and SSIM values of the despeckled images of Geometric,

TABLE III
QUANTITATIVE MEASURES ON THE DESPECKLED IMAGES OF DIFFERENT METHODS

Methods	<i>Geometric</i> $L = 1$		<i>Lena</i> $L = 3$		<i>Cameraman</i> $L = 3$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Noisy	10.606	0.663	18.184	0.407	16.508	0.380
SDD	19.485	0.685	25.112	0.725	24.338	0.698
MIDAL	17.555	0.950	25.982	0.805	24.875	0.788
BS-TV	21.528	0.703	26.863	0.784	25.850	0.743
MAP-S	24.831	0.939	27.943	0.825	26.662	0.781
POTDF	26.906	0.907	27.914	0.851	26.567	0.809
PPB	22.867	0.922	27.374	0.819	25.784	0.755
SAR-BM3D	28.109	0.951	29.135	0.860	27.402	0.817
NL-SAR	23.572	0.912	29.226	0.855	27.293	0.811
FANS	23.446	0.909	29.412	0.865	27.638	0.823
FT-WNNM	23.072	0.952	29.277	0.867	27.689	0.815
SAR-CNN	27.598	0.956	29.364	0.862	27.937	0.819
cFT-HTpV	28.359	0.989	28.603	0.846	27.301	0.818

Lena, and Cameraman generated by all the methods are listed in Table III with the optimal two values written in bold for each criterion. As shown in Figs. 6 and 7, the first-order TV-based SDD, MIDAL, and BS-TV methods bring obvious staircase artifact, which is effectively overcome by the proposed hybrid first- and second-order-based cFT-HTpVs. The MAP-S also brings serious undesirable artifacts, as shown in Figs. 6(f) and 7(f). In order to fully compare these despeckled images, we mark some details with the red regions in Figs. 6 and 7. It can be seen that the NL mean-based PPB, SAR-BM3D, NL-SAR, and FANS can bring impressive despeckling results, but they also introduce some undesirable artifacts, for example, located on the face of Lena. However, the other despeckling methods, such as cFT-HTpV, POTDF, FT-WNNM, and SAR-CNN, oversmooth the images and cause the loss of some important details, for example, the details between nose and mouth of Lena. From this point of view, preserving detail and smoothing noise need to be balanced in the despeckling process. From the indicators of PSNR and SSIM listed in Table III, we can find that the cFT-HTpV performs best in the Geometric, which obtains the highest PSNR and SSIM. In the Lena, the methods to obtain the highest PSNR and SSIM are FANS and FT-WNNM, respectively. In the Cameraman, the SAR-CNN gets the highest PSNR, and the FANS gets the best SSIM. The proposed cFT-HTpV is inferior to FANS, SAR-BM3D, FT-WNNM, and SAR-CNN and better than POTDF, PPB, SDD, MIDAL, BS-TV, and MAP-S on PSNR and SSIM in Lena and Cameraman.

C. Experiments on Synthetic Data Sets

In this section, we test these methods on the benchmarking framework proposed in [63], where five suitable canonical scenes (homogeneous, digital elevation model (DEM), squares, corner, and building) and corresponding objective measures are selected to consider the abilities of different methods on speckle suppression and feature preservation. The mean and variance of ratio (MoR and VoR) measure the degree of radiometric preservation; the ENL, refined ENL (referred to as ENL*), and the despeckling gain (DG) evaluate the speckle suppression in homogeneous areas; the coefficient of variation (C_x) measures the degree of texture preservation

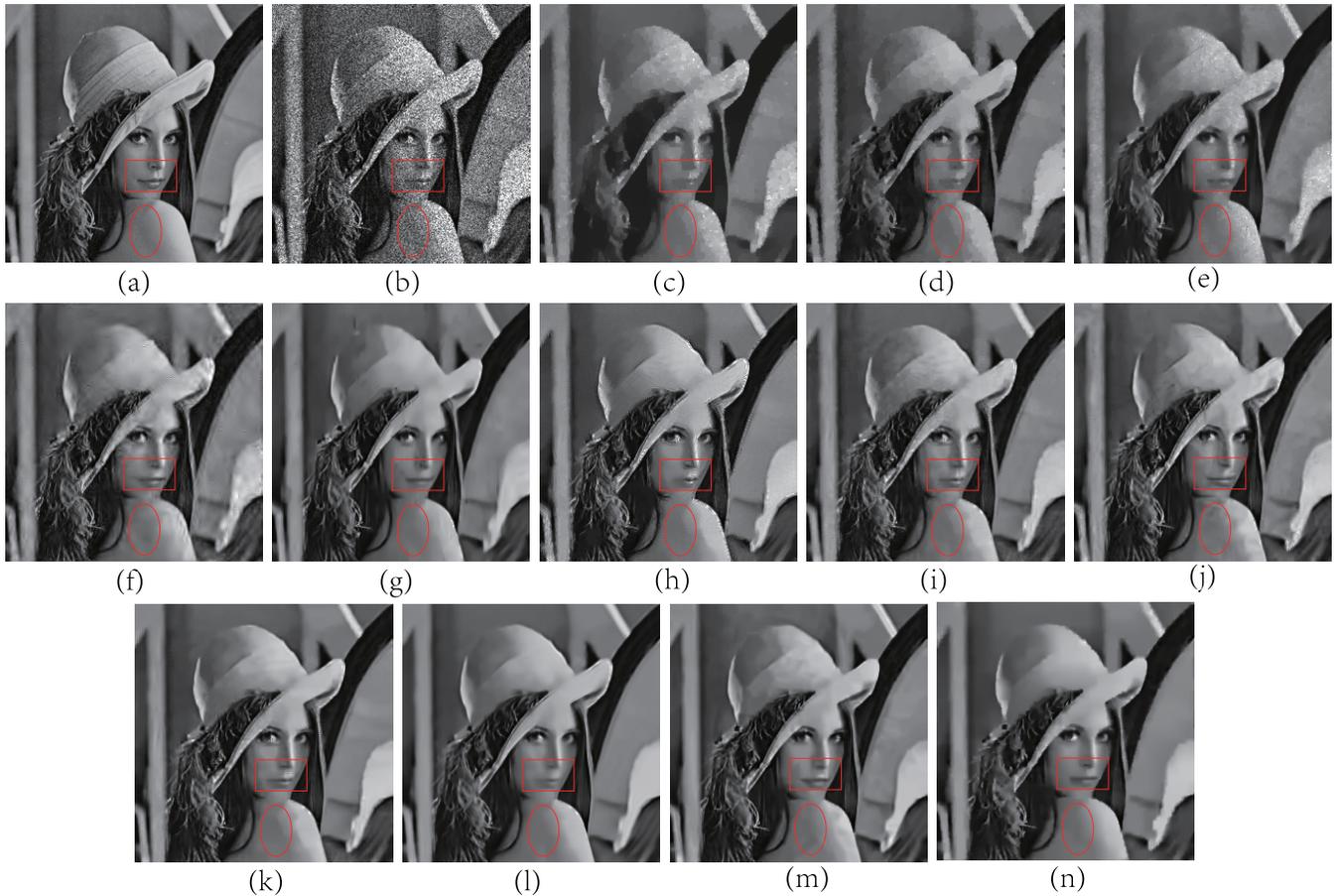


Fig. 6. Despeckling results on the Lena image with size 256×256 . (a) Clean Lena image. (b) Contaminated image by the speckle noise with $L = 3$. (c) SDD. (d) MIDAL. (e) BS-TV. (f) MAP-S. (g) POTDF. (h) PPB. (i) SAR-BM3D. (j) NL-SAR. (k) FANS. (l) FT-WNNM. (m) SAR-CNN. (n) cFT-HTpV.

for DEM image; the edge smearing (ES) and Pratt's figure of merit (FOM) measure edge preservation for Squares image; the intensity contrast values C_{NN} and C_{BG} evaluate the degree of radiometric preservation through the filtering process in Corner image; and the double reflection (C_{DR}) measures the preservation degree of the position and the radiometric characteristics of the building radar return and the BS measure the distortion of the radiometric building profile in the range direction for the Building image. For more details of how these data sets are generated, why these parameters are selected, and how to compute these parameters, we refer the readers to [63]. Therefore, for a good method, the computed values for MoI, MoR, C_x , C_{NN} , C_{BG} , and C_{DR} should be close to their clean values; the computed values for ENL, ENL*, DG, and FOM should be large values; and the computed values for ES and BS should be small values.

Table IV reports the measure values of different methods on the benchmarking framework, where the best and second-best results are shown in bold. All the results reported in Table IV are generated by averaging the despeckling results over eight independent single-look simulated images of the same scene. Fig. 8 shows the clean, noisy, and cFT-HTpV despeckling results of each scene, except the Building data set, where the ratio images for the reference and cFT-HTpV are shown.

Meanwhile, since the boundary value of the clean squares data set is very different, as shown in Fig. 8, we slightly modify the forms of the banded circulant difference matrices $\mathbf{D}_{1,M}$ and $\mathbf{D}_{2,M}$ to fit this data set. As pointed out in [63], this benchmarking framework can provide insight into the abilities and limitations of SAR despeckling approaches. From Table IV and Fig. 8, we can find that the proposed cFT-HTpV shows a strong speckle removing ability in homogeneous scene, with an ENL of 385.87 and DG of 24.25. In the DEM scene, cFT-HTpV gains the nearest value to the clean C_x and the relatively large DG of 5.29, which means that cFT-HTpV can not only remove speckle but also preserve the texture well. In the squares scene, the cFT-HTpV shows its good characteristics in edge-preserving with the second-smallest ES that is only higher than SAR-BM3D and the second-largest FOM that is only smaller than FT-WNNM. In the case of corner scene, the best C_{NN} and C_{BG} that used to measure the radiometric preservation through the filtering process are obtained by cFT-HTpV. This is mainly due to the fact that the strong scattering points are detected first in the cFT-HTpV, and then, the image is despeckled. For the last building scene, the cFT-HTpV can also provide acceptable results (C_{DR} and BS) that are slightly inferior to SAR-BM3D and FT-WNNM, which also shows good performance on building feature preservation.

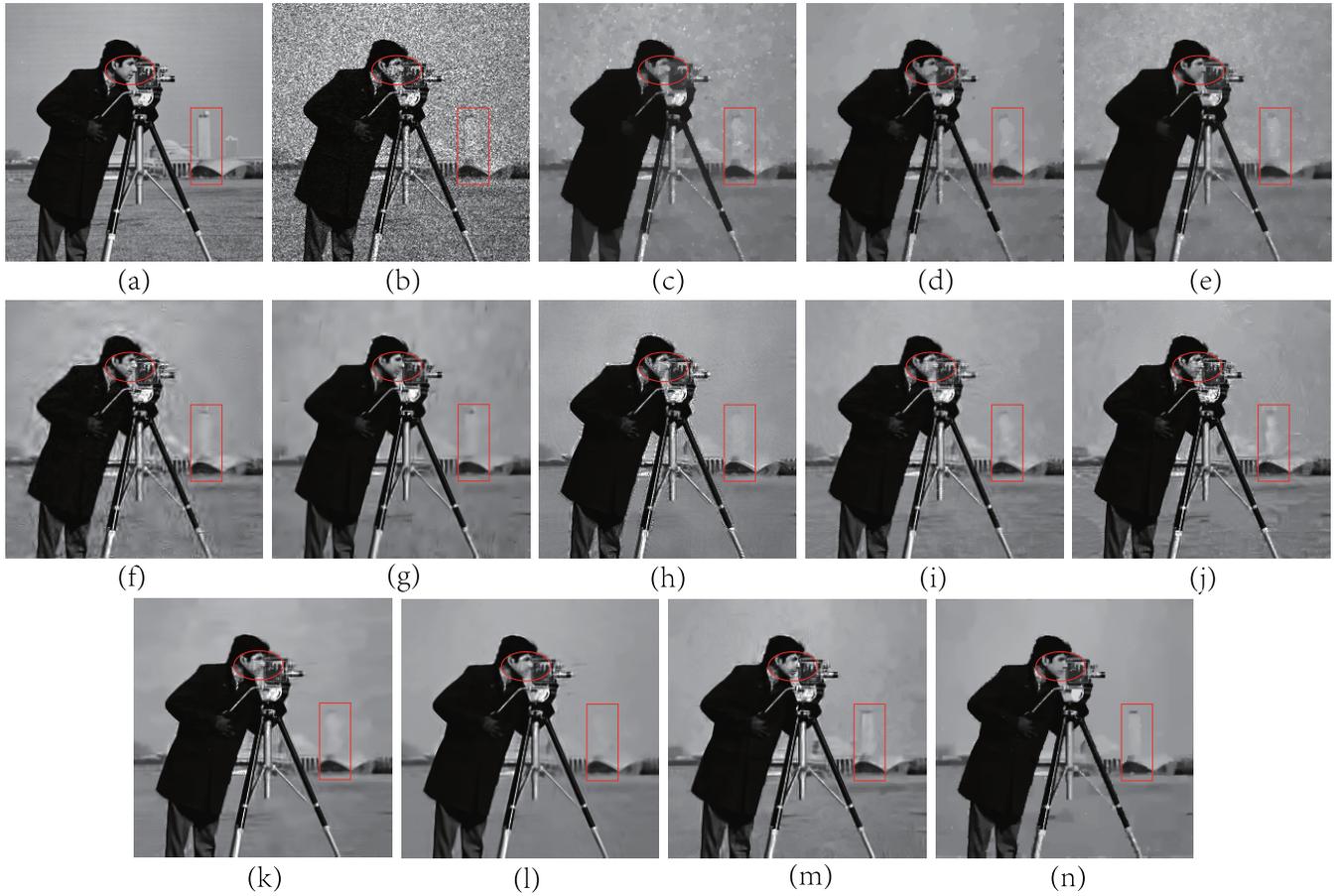


Fig. 7. Despeckling results on Cameraman image with size 256×256 . (a) Clean Cameraman image. (b) Contaminated image by the speckle noise with $L = 3$. (c) SDD. (d) MIDAL. (e) BS-TV. (f) MAP-S. (g) POTDF. (h) PPB. (i) SAR-BM3D. (j) NL-SAR. (k) FANS. (l) FT-WNNM. (m) SAR-CNN. (n) cFT-HTpV.

TABLE IV
QUANTITATIVE MEASURES ON THE BENCHMARKING FRAMEWORK OF DIFFERENT METHODS

Scenes	Measures	Noisy	Clean	SDD	MIDAL	POTDF	PPB	SAR-BM3D	FANS	FT-WNNM	cFT-HTpV
Homogeneous	MoI	0.998	1.000	0.885	0.979	1.008	0.997	0.978	1.008	0.997	0.998
	MoR	/	0.998	1.104	1.001	0.982	0.960	0.979	0.983	0.998	0.999
	ENL	1.00	436.97	216.16	178.88	79.29	127.68	102.44	147.70	363.23	385.87
	ENL*	1.02	510.36	250.39	199.33	86.26	140.89	111.91	161.15	504.57	479.38
	DG	0	/	12.91	21.32	18.35	20.29	19.40	20.48	23.73	24.25
DEM	MoI	1.003	1.000	0.767	0.677	0.782	0.998	0.968	0.962	0.973	0.959
	MoR	/	1.001	1.070	1.000	1.065	0.911	0.833	0.843	0.838	1.003
	C_x	3.54	2.40	2.33	2.42	2.19	2.71	2.43	2.55	2.41	2.39
	DG	0	/	3.91	4.27	5.27	3.68	5.32	4.99	5.68	5.29
Squares	ES(up)	0.010	/	0.138	0.038	0.095	0.070	0.036	0.050	0.079	0.065
	ES(down)	0.029	/	0.354	0.157	0.252	0.209	0.113	0.155	0.247	0.141
	FOM	0.792	0.993	0.747	0.841	0.763	0.837	0.847	0.823	0.874	0.856
Corner	C_{NN}	7.77	7.75	7.85	7.04	6.77	3.75	7.39	7.09	7.40	7.81
	C_{BG}	36.50	36.56	36.98	31.79	31.79	32.69	35.45	35.50	35.21	36.54
Building	C_{DR}	65.90	65.90	66.97	63.55	65.88	64.90	65.91	65.61	65.90	65.71
	BS	0.09	/	2.39	7.07	7.33	3.13	1.46	3.51	1.41	1.776

In order to fully assess the radiometric preservation through the despeckling process, we also plot the range profiles and average range profiles obtained for different methods on the corner scene and building scene, respectively. In Fig. 9(a), the logarithmic values of 20 pixels above and below the center of the corner are plotted; in Fig. 9(b), the average

range profiles obtained by averaging over the range lines interested by the presence of the building are plotted in the logarithmic scale. By comparing these range profiles with the clean one, we can find that cFT-HTpV, FT-WNNM, and SAR-BM3D perform very well, while MIDAL strongly affects the radiometric properties of the corner and building, giving

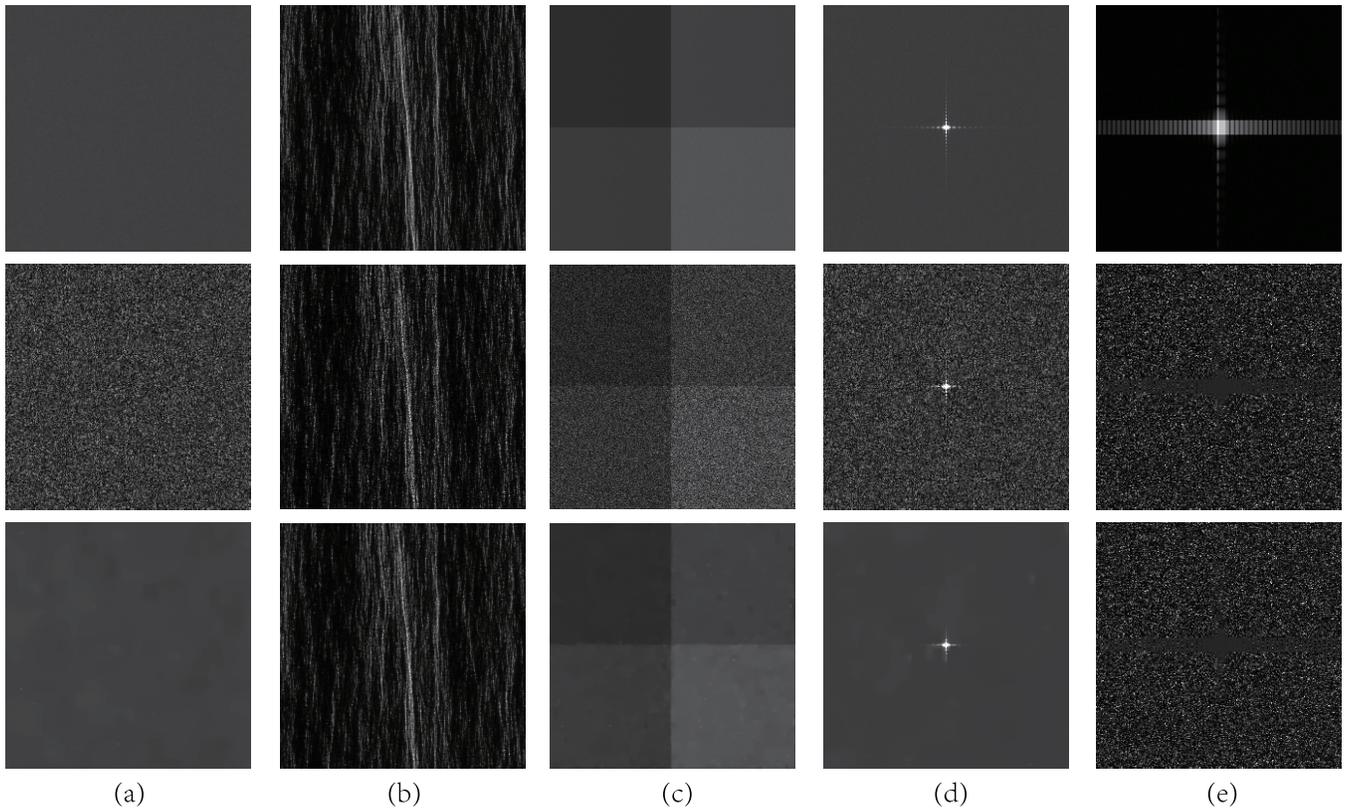


Fig. 8. Testing cFT-HTpV on the SAR benchmarking framework. The first four columns (from top to bottom) are the clean scenes, noisy scenes, and cFT-HTpV despeckling results, respectively; the last column (from top to bottom) are the clean scene, ratio image for the reference, and ratio image for cFT-HTpV, respectively. Left to Right: Different scenes. (a) Homogeneous. (b) DEM. (c) Squares. (d) Corner. (e) Building. The display window is $[0, 2]$.

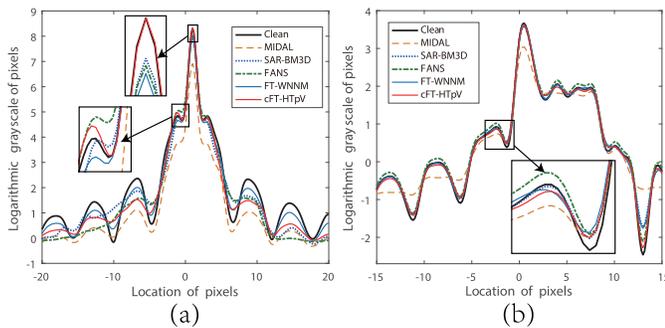


Fig. 9. Image profiles obtained by different methods compared with the clean one. (a) Range profiles on the corner scene. (b) Average range profiles on building scene.

rise to a serious smoothing effect. The quantitative measures of corner scene and building scene in Table IV also confirm the above conclusions.

D. Experiments on Real SAR Images

In this section, we conduct experiments on three real SAR images to observe the performance of the proposed cFT-HTpV, where the testing images contain different scenes, different noise levels (single-look/multilooks), and different sensor characteristics as: 1) a three-look SAR image, a small building in Washington area, sensed by the Ku-band MiniSAR

with a 0.3-m resolution, named Building (with size 256×256 , as shown in Fig. 10(a)²); 2) a five-look SAR image, a small part of the coast, sensed by C-band Sentinel-1 with HH polarization and 20-m resolution, named Coast [with size 256×256 , as shown in Fig. 11(a)]; and 3) a single-look SAR image, a small part of town in Shanxi, China, sensed by C-band Sentinel-1 with VH polarization and 5-m resolution, named Town [with size 286×241 , as shown in Fig. 12(a)]. In Figs. 10 and 11, all the methods remove the speckle in homogeneous regions, such as the land area in Fig. 10 and the sea surface in Fig. 11. By comparing the SDD, MIDAL, BS-TV, and cFT-HTpV results in Figs. 10 and 11, we can find that the hybrid TV-based cFT-HTpV can efficiently reduce the staircase effect, which is brought in by the first-order-based TV methods. The smooth performance of MAP-S on the homogeneous regions is not as good as NL-based methods, as shown in Fig. 10(e). By observation, the POTDF introduces some pointwise artifacts in the despeckled results, which can be seen in Fig. 10(f). Besides, the POTDF may also oversmooth the image and cause some details to be lost, as shown in Fig. 11(f), where some weak objects, especially the road details, are neglected. Besides, we can also find that the SAR-BM3D, NL-SAR, and FANS have less ability to smooth image in homogeneous regions than other methods, which can be clearly seen in Fig. 11(h)–(j).

²Data processed by Sandia National Laboratories, <http://www.sandia.gov>

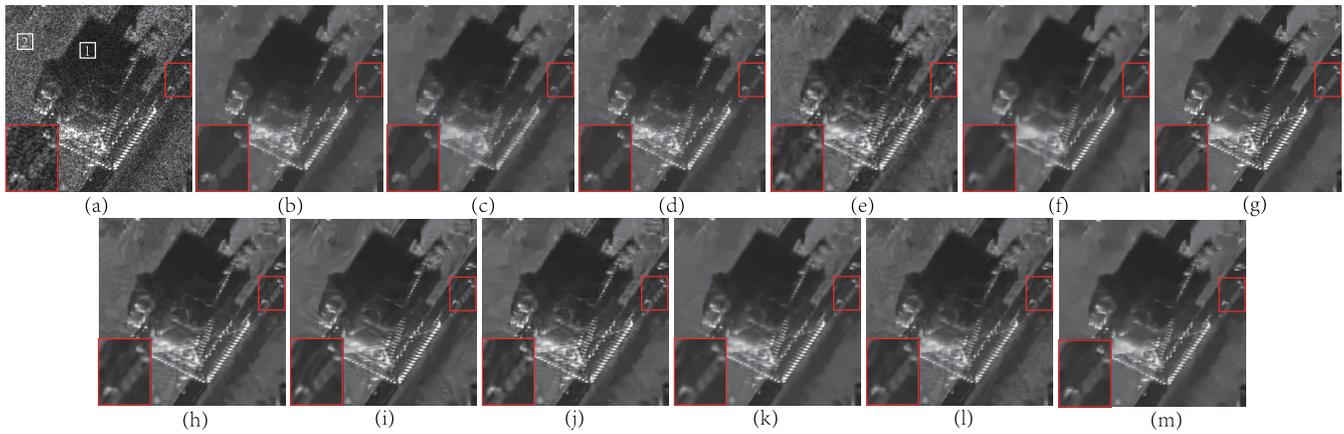


Fig. 10. Despeckling results on Building data set: a three-look SAR image sensed by MiniSAR with image size 256×256 . (a) Original speckled image. (b) SDD. (c) MIDAL. (d) BS-TV. (e) MAP-S. (f) POTDF. (g) PPB. (h) SAR-BM3D. (i) NL-SAR. (j) FANS. (k) FT-WNNM. (l) SAR-CNN. (m) cFT-HTpV. The white boxes in (a) are the selected homogeneous regions for computing ENL.

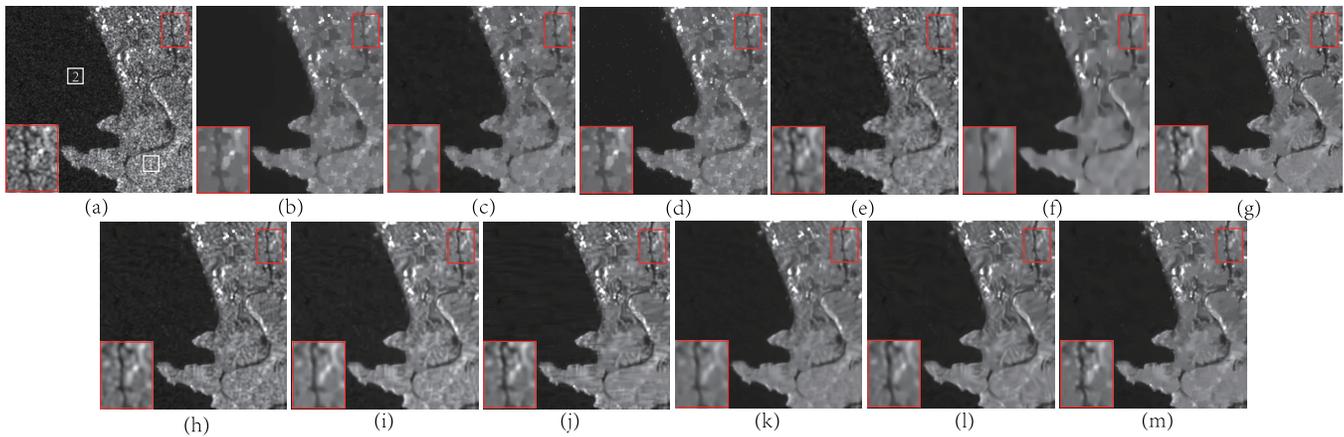


Fig. 11. Despeckling results on Coast data set: a five-look SAR image sensed by Sentinel-1 with image size 256×256 . (a) Original speckled image. (b) SDD. (c) MIDAL. (d) BS-TV. (e) MAP-S. (f) POTDF. (g) PPB. (h) SAR-BM3D. (i) NL-SAR. (j) FANS. (k) FT-WNNM. (l) SAR-CNN. (m) cFT-HTpV. The white boxes in (a) are the selected homogeneous regions for computing ENL.

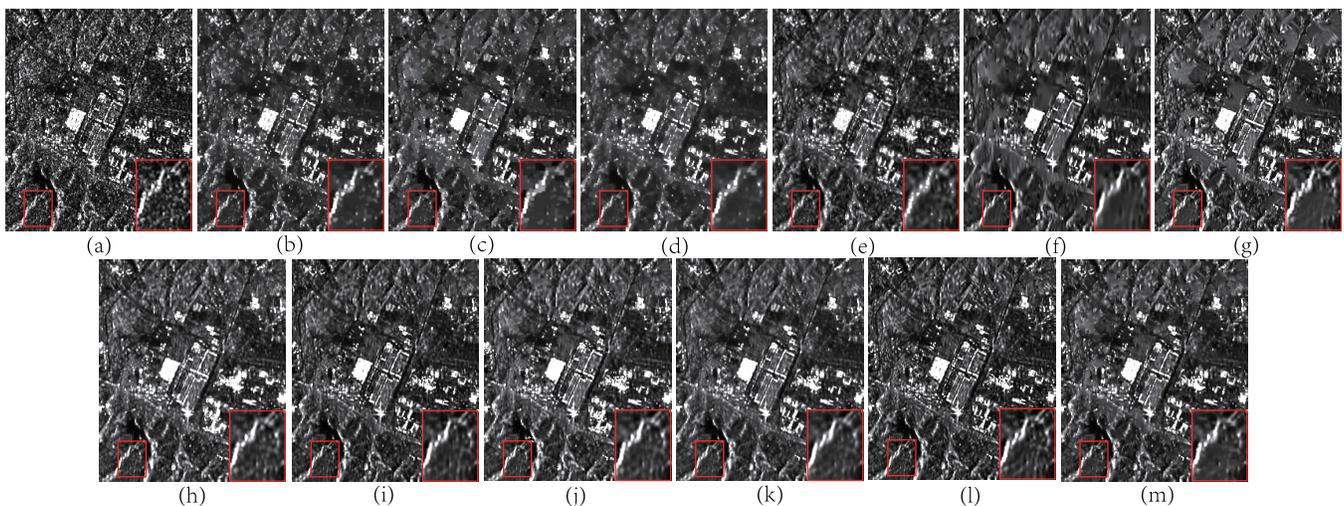


Fig. 12. Despeckling results on Town data set: a single-look SAR image sensed by Sentinel-1 with image size 286×241 . (a) Original speckled image. (b) SDD. (c) MIDAL. (d) BS-TV. (e) MAP-S. (f) POTDF. (g) PPB. (h) SAR-BM3D. (i) NL-SAR. (j) FANS. (k) FT-WNNM. (l) SAR-CNN. (m) cFT-HTpV.

Visually, the PPB, FT-WNNM, SAR-CNN, and cFT-HTpV can obtain better results in these two data sets with better speckle reduction in homogeneous regions and better detail

preservation. To further test the abilities of these methods, we use a more challenging single-look data set, which includes roads, buildings, land, mountains, and lakes, as shown

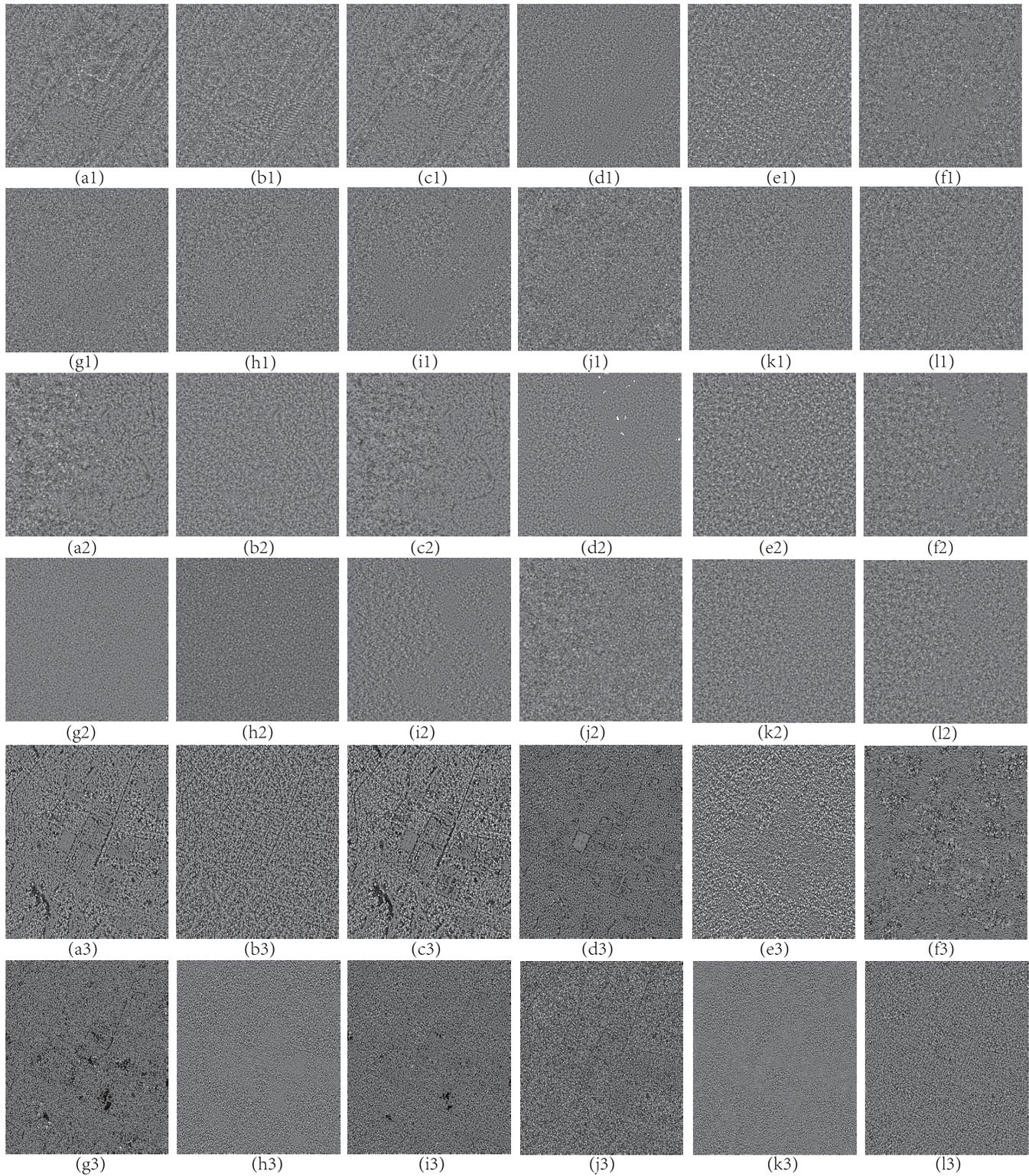


Fig. 13. Ratio images of different methods on three data sets. The top two lines, the middle two lines, and the bottom two lines correspond to the Building, Coast, and Town data sets, respectively. (a) SDD. (b) MIDAL. (c) BS-TV. (d) MAP-S. (e) POTDF. (f) PPB. (g) SAR-BM3D. (h) NL-SAR. (i) FANS. (j) FT-WNNM. (k) SAR-CNN. (l) cFT-HTpV.

in Fig. 12. By comparing the despeckled images, it can be found that the SDD, MIDAL, and POTDF are not as good as other methods in edge preserving, and the SAR-BM3D and FANS cannot suppress the speckle well in this

complicated scene as other methods, such as FT-WNNM, SAR-CNN, and cFT-HTpV. In addition, the PPB despeckled image in Fig. 12(g) shows some undesirable artifacts and presents some blocky patterns. On the whole, the cFT-HTpV

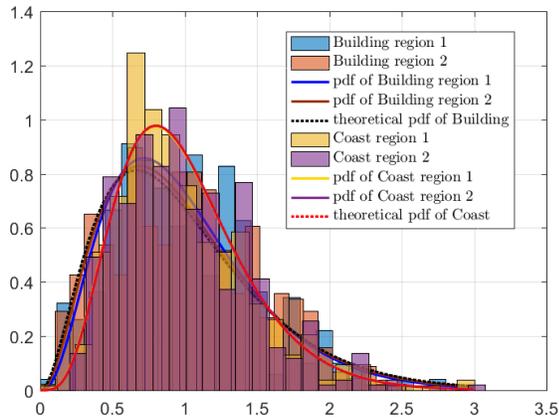


Fig. 14. Empirical versus theoretical speckle pdfs in the selected homogeneous regions. The fitted Gamma distribution parameters of the two regions in Building are $\alpha_1 = \beta_1 = 3.5316$ and $\alpha_2 = \beta_2 = 3.1950$, respectively, where α is the shape parameter and β is the rate parameter of the Gamma distribution. The theoretical Gamma distribution parameter of these two regions in Building is $\alpha = \beta = 3$. For the two homogeneous regions selected in Coast, the fitted Gamma distribution parameters are $\alpha_1 = \beta_1 = 4.9531$ and $\alpha_2 = \beta_2 = 5.0030$, respectively, where the theoretical parameters are $\alpha = \beta = 5$.

has satisfactory performance in speckle removal and detailed maintenance. Meanwhile, in the Town data set, the ratio detector in cFT-HTpV can accurately detect strong scattering points, especially the “cross bright spots” in the image, which has very high radiation reflectivity. However, for the rectangular white area with high radiometric reflectivity around the center of the image, the ratio detector fails to detect it. In this case, a simple improved method is to filter out the edges with large gradient value by thresholding method and then put them into the index set \mathbf{I}_s^d .

Fig. 13 presents the ratio images for these data sets. As can be seen from Fig. 13(a)–(e), the SDD, MIDAL, BS-TV, MAP-S, and POTDF show obvious edges and structures in the ratio images, which means that some edges or details are oversmoothed. In the ratio images of NL-SAR, SAR-CNN, and cFT-HTpV, there are no obvious structures and patterns as other methods, which shows good edge-preservation ability.

In order to further compare the effects of these methods, the quantitative comparison in terms of ENL and MoR image is reported in Table V. The homogeneous area used to calculate the ENL is indicated by rectangular boxes in the Figs. 10(a) and 11(a). They are 25×25 pixel squares. In Fig. 14, we plot the histogram of each selected homogeneous region [normalized by $\tilde{\mathbf{X}}/\text{mean}(\tilde{\mathbf{X}})$] and the corresponding fitted pdfs that obey the Gamma distribution. From this, we can see that the empirical speckle distribution fits very well with the theoretical one with some small deviations probably due to the insufficient sample (625 pixels) and the original SAR data geocoding. A high value of ENL indicates a strong ability to remove speckles in homogeneous areas. MoR can measure the degree of radiometric preservation [44]. If MoR is close to the ideal value 1, it shows that the method has good radiometric preservation ability. Combining Table V and Fig. 13, we can find that the MAP-S and SAR-BM3D can preserve the radiometric well with MoR relatively closing to the

TABLE V
QUANTITATIVE MEASURES ON THE BUILDING AND COAST DATA SETS OF DIFFERENT METHODS

Methods	Building dataset			Coast dataset		
	ENL(1)	ENL(2)	MoR	ENL(1)	ENL(2)	MoR
Original	4.01	3.89	/	4.58	4.95	/
SDD	899.87	178.71	1.041	63.43	1328.58	1.023
MIDAL	459.54	500.76	1.073	83.65	147.40	1.040
BS-TV	854.05	104.43	1.038	37.42	40.09	1.018
MAP-S	45.71	91.91	0.958	39.70	145.43	1.029
POTDF	206.38	748.42	1.059	155.83	344.70	1.031
PPB	235.01	281.21	0.972	150.39	377.18	0.974
SAR-BM3D	173.69	122.46	0.978	20.08	35.23	0.972
NL-SAR	400.65	167.51	1.028	39.31	63.97	0.823
FANS	180.25	385.45	0.966	62.27	101.47	0.966
FT-WNNM	511.21	833.62	0.990	117.37	448.85	0.990
SAR-CNN	132.61	112.18	1.024	86.10	497.79	1.017
cFT-HTpV	463.56	911.09	0.983	145.76	702.61	0.986

TABLE VI
QUANTITATIVE MEASURES OF CFT-HTpV WITH DIFFERENT NORMS ON THE NIMES DATA SETS

Norm	ENL(1)	ENL(2)	MoR	PSNR	SSIM
$p = 1$	35.01	31.71	0.938	24.549	0.836
$p = 0.7$	118.80	103.56	0.937	25.016	0.849
$p = 0.5$	144.56	174.75	0.932	24.912	0.834
$p = 0.2$	199.20	262.77	0.913	24.117	0.794

ideal value. However, the ENL of MAP-S and SAR-BM3D are smaller than other methods, which means that the MAP-S and SAR-BM3D have poor speckle reduction performance in the homogeneous areas. The MoR values of SDD, MIDAL, and POTDF deviate more from the ideal value than other methods, which means that there exists heavier radiometric distortion. Among all these comparing methods, the FT-WNNM and cFT-HTpV achieve better results in terms of removing speckle in homogeneous areas while preserving the radiometric well. However, the cFT-HTpV is more efficient than FT-WNNM, as shown in Section III-F. Based on the abovementioned analysis, the proposed cFT-HTpV can achieve quite comparable or better performance in terms of speckle suppression, radiometric preservation, detail preservation, and visual effect at the same time.

E. Test of Different Norms p

In order to investigate the performance of the nonconvex HTpV model under different norms, we test the proposed cFT-HTpV with $p = 1, 0.7, 0.5, 0.2$ on the Nimes data set. Fig. 15 shows the original image, the noisy image with $L = 4$, and the despeckled images with different p 's. Table VI reports the quantitative measures for comparing. From Fig. 15 and Table VI, we can find that when p is small (e.g., $p = 0.2$), the model has strong ability of noise removing. Thus, it has a large ENL in the homogeneous regions. However, some details are oversmoothed and lost. This is mainly due to the fact that a smaller p can induce a sparser solution in the first- and second-order TV domains. Such strong sparse constraint can smooth the noise very well in the homogeneous regions, but, at the same time, it will remove some important details. With the increase in p , the ability of HTpV to preserve the details

TABLE VII
COMPUTATIONAL TIME (SECONDS) OF DIFFERENT METHODS

Image size	SDD	MIDAL	BS-TV	MAP-S	POTDF	PPB	SAR-BM3D	NL-SAR	FANS	FT-WNNM	cFT-HTpV
256×256	2.47	8.04	265.55	17.86	9.41	13.78	15.51	3.46	1.83	191.35	3.23
512×512	16.48	54.75	1468.24	66.43	16.67	47.45	61.98	12.14	6.88	1235.85	22.43

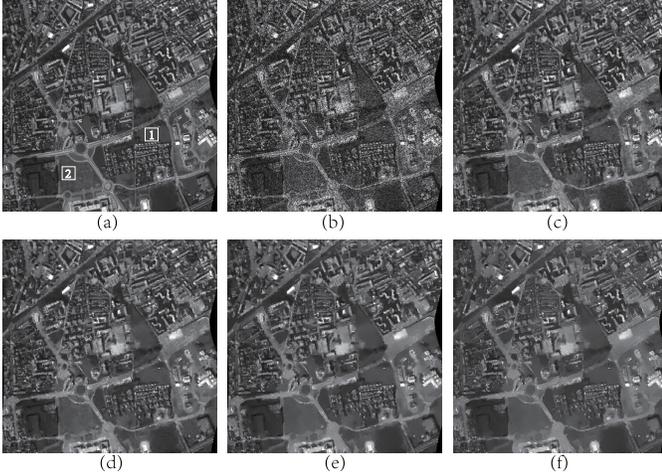


Fig. 15. Despeckling results of cFT-HTpV with different norms on the Nimes data set. (a) Original image with size 512×512 . (b) Speckled image with $L = 4$. (c) cFT-HTpV with $p = 1$. (d) cFT-HTpV with $p = 0.7$. (e) cFT-HTpV with $p = 0.5$. (f) cFT-HTpV with $p = 0.2$. The white boxes in (a) are the selected homogeneous regions for computing ENL.

is stronger, but the ability to reduce the speckle is weaker. Therefore, we can set $p = 0.7$ as a compromise choice.

F. Comparison of Computational Time

The main space and time complexity of the cFT-HTpV is concentrating on the process of solving FT-HTpV (see Algorithms 1 and 2). For the space complexity, storing the difference operators \mathbf{C}_h and \mathbf{C}_v requires $\mathcal{O}(2S)$ each, storing \mathbf{C}_{hh} and \mathbf{C}_{vv} requires $\mathcal{O}(3S)$ each, storing \mathbf{C}_{hv} or \mathbf{C}_{vh} requires $\mathcal{O}(4S)$, and storing the matrix \mathbf{A} requires $\mathcal{O}(13S)$. We can find that the space complexity of the proposed algorithm is very low. For the time complexity, we assume that the numbers of external and internal iterations are k and n , respectively. Taking Algorithm 1 for example, the following holds.

- 1) *Updating \mathbf{b}* : As the matrix $\text{dexp}^{\mathbf{Y}-\mathbf{X}^k}$ is a diagonal matrix, updating \mathbf{b} requires $\mathcal{O}(kS)$.
- 2) *Updating \mathbf{A}* : As \mathbf{W} in (A.7) is a diagonal matrix, each update of matrix \mathbf{W} requires $\mathcal{O}(S^2)$ for calculating $\mathbf{C}\mathbf{X}$ in (A.2), and each update of matrix \mathbf{A} requires $\mathcal{O}(S^3)$ for calculating \mathbf{G} in (A.6). Then, updating \mathbf{A} requires $\mathcal{O}(knS^3)$.
- 3) *Calculating \mathbf{X}* : As we applying the efficient PCG method to solve the linear system (23), therefore, updating \mathbf{X} requires $\mathcal{O}(kni_{\text{pcg}}S^2)$, where i_{pcg} is the iteration number of the PCG method (its upper bound is set to 100 in the experiment).

Although the complexity of the algorithm is very high in the abovementioned theoretical analysis, which requires $\mathcal{O}(knS^3)$, in practice, it does not need a long computational time. The main reason is that the matrices

of \mathbf{W} , \mathbf{G} , and \mathbf{A} are all sparse and symmetric, which can greatly reduce the calculation time. For example, since there are only $2S$ nonzero elements in matrix \mathbf{C}_h , it only requires $\mathcal{O}(S)$ for calculating the multiplication $\mathbf{C}_h\mathbf{X}$ in (A.2). Meanwhile, we adopt the nmAPG framework to accelerate the algorithm, which can effectively reduce the number of iterations, thus further reducing the computational time.

Table VII reports the CPU times of different methods. The FANS outperforms all others in computational time. The two most time-consuming methods are the FT-WNNM and BS-TV mainly because of their high complexities on the NL rank minimization process and the graph-cuts method, respectively. The proposed cFT-HTpV is in the middle level of time consumption by using the efficient matrix-vector approach and the nmAPG framework. It takes less time than BS-TV, FT-WNNM, MIDAL, and SAR-BM3D but more time than FANS and SDD.

IV. CONCLUSION

In this article, we propose a new SAR image despeckling method by using a step-by-step framework. It first divides the image into strong scattering points and nonstrong scattering points and then uses the FT-HTpV model to remove the speckle. Because it is based on the FT distribution characteristics of noise, it has better data fidelity than the least-squares form data fidelity term. More importantly, this model also uses the ℓ_p -norm of hybrid TV as the regularization, which inherits the advantages of the nonconvex regularization and the higher order TV regularization. Therefore, this new model effectively removes the speckle while preserving edges and reducing staircase artifacts well. To solve the FT-HTpV algorithm efficiently, we choose the nmAPG framework and use a matrix-vector strategy to accelerate the algorithm. Extensive experiments on both the simulated and real SAR images demonstrate the effective performance of the proposed cFT-HTpV, which can achieve quite comparable or better performance in terms of speckle suppression, radiometric preservation, detail preservation, visual effect, and solving efficiency at the same time by comparing with some state-of-the-art methods. Meanwhile, this article mainly focuses on single-polarization SAR image despeckling, but the framework proposed in this article can also be extended to polarimetric SAR (PolSAR) data. For example, we can replace the ordinary TV regularization term with the proposed classification-based HTpV in the Wishart-TV (WisTV)-based model [64] for PolSAR data speckle reduction, or we can embed the classification-based HTpV into the plug-and-play MuLoG scheme [65], which includes the Gaussian denoisers within a multichannel SAR speckle reduction technique. As the PolSAR has attracted more and more attention, exploring the application of the proposed model in PolSAR is an important work for us in the future.

APPENDIX

SIMPLIFYING FT-HTPV MODEL

In order to transform the model (16) into matrix operation form, we define two $M \times M$ banded circulant matrices $\mathbf{D}_{1,M}$ and $\mathbf{D}_{2,M}$ as

$$\mathbf{D}_{1,M} = \begin{bmatrix} -1 & 1 & & & 0 \\ 0 & -1 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & 0 & -1 & 1 \\ 1 & & & & 0 & -1 \end{bmatrix}$$

$$\mathbf{D}_{2,M} = \begin{bmatrix} 1 & 0 & & & -1 \\ -1 & 1 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & & -1 & 1 & 0 \\ 0 & & & & -1 & 1 \end{bmatrix}. \quad (\text{A.1})$$

Then, for $\mathbf{X} \in \mathbb{R}^{M \times N}$, we have $\mathbf{D}_v^+ \mathbf{X} = \mathbf{D}_{1,M} \mathbf{X}$, $\mathbf{D}_h^+ \mathbf{X} = \mathbf{X} \mathbf{D}_{1,N}^T$, $\mathbf{D}_v^- \mathbf{X} = \mathbf{D}_{2,M} \mathbf{X}$, and $\mathbf{D}_h^- \mathbf{X} = \mathbf{X} \mathbf{D}_{2,N}^T$. As $\mathbf{X} \in \mathbb{R}^S$ also represents the vectorized image, by defining the vectorize operator as $\text{vec}(\cdot)$, $\mathbf{C}_h = \tilde{\mathbf{I}}_{\text{ns}}^d (\mathbf{D}_{1,N} \otimes \mathbf{I}_M)$, $\mathbf{C}_v = \tilde{\mathbf{I}}_{\text{ns}}^d (\mathbf{I}_N \otimes \mathbf{D}_{1,M})$, $\mathbf{C}_{hh} = \tilde{\mathbf{I}}_{\text{ns}}^d ((\mathbf{D}_{2,N} \mathbf{D}_{1,N}) \otimes \mathbf{I}_M)$, $\mathbf{C}_{hv} = \tilde{\mathbf{I}}_{\text{ns}}^d (\mathbf{D}_{1,N} \otimes \mathbf{D}_{1,M})$, $\mathbf{C}_{vh} = \tilde{\mathbf{I}}_{\text{ns}}^d (\mathbf{D}_{1,N} \otimes \mathbf{D}_{1,M})$, and $\mathbf{C}_{vv} = \tilde{\mathbf{I}}_{\text{ns}}^d (\mathbf{I}_N \otimes (\mathbf{D}_{2,M} \mathbf{D}_{1,M}))$, where $\tilde{\mathbf{I}}_{\text{ns}}^d \in \mathbb{R}^{S \times S}$ is a diagonal matrix with the diagonal elements $(\tilde{\mathbf{I}}_{\text{ns}}^d)_{i,i} = (\text{vec}(\mathbf{I}_{\text{ns}}^d))_i$, we have

$$\begin{aligned} \text{vec}(\mathbf{I}_{\text{ns}}^d \odot (\mathbf{D}_h^+ \mathbf{X})) &= \mathbf{C}_h \mathbf{X} \\ \text{vec}(\mathbf{I}_{\text{ns}}^d \odot (\mathbf{D}_v^+ \mathbf{X})) &= \mathbf{C}_v \mathbf{X} \\ \text{vec}(\mathbf{I}_{\text{ns}}^d \odot (\mathbf{D}_{hh}^+ \mathbf{X})) &= \mathbf{C}_{hh} \mathbf{X} \\ \text{vec}(\mathbf{I}_{\text{ns}}^d \odot (\mathbf{D}_{hv}^+ \mathbf{X})) &= \mathbf{C}_{hv} \mathbf{X} \\ \text{vec}(\mathbf{I}_{\text{ns}}^d \odot (\mathbf{D}_{vh}^+ \mathbf{X})) &= \mathbf{C}_{vh} \mathbf{X} \\ \text{vec}(\mathbf{I}_{\text{ns}}^d \odot (\mathbf{D}_{vv}^+ \mathbf{X})) &= \mathbf{C}_{vv} \mathbf{X}. \end{aligned} \quad (\text{A.2})$$

Substituting (A.2) into (16), we have

$$\begin{aligned} E(\mathbf{X}) &= \sum_{i=1}^S L(x_i + e^{y_i - x_i}) + \lambda \beta (\|\mathbf{C}_h \mathbf{X}\|_p^p + \|\mathbf{C}_v \mathbf{X}\|_p^p) \\ &\quad + \lambda (1 - \beta) (\|\mathbf{C}_{hh} \mathbf{X}\|_p^p + \|\mathbf{C}_{hv} \mathbf{X}\|_p^p \\ &\quad \quad + \|\mathbf{C}_{vh} \mathbf{X}\|_p^p + \|\mathbf{C}_{vv} \mathbf{X}\|_p^p) \\ &= \sum_{i=1}^S L(x_i + e^{y_i - x_i}) + \lambda \beta (|(\mathbf{C}_h \mathbf{X})_i|^p + |(\mathbf{C}_v \mathbf{X})_i|^p) \\ &\quad + \lambda (1 - \beta) (|(\mathbf{C}_{hh} \mathbf{X})_i|^p + |(\mathbf{C}_{hv} \mathbf{X})_i|^p \\ &\quad \quad + |(\mathbf{C}_{vh} \mathbf{X})_i|^p + |(\mathbf{C}_{vv} \mathbf{X})_i|^p). \end{aligned} \quad (\text{A.3})$$

Inspired by the method in [38], we use an approximation strategy to estimate the p -power function $|t|^p$ as $t^2 (|t|^{2-p} + \varepsilon)^{-1}$, where $\varepsilon > 0$ is a small positive constant to control the approximation accuracy. It is easy to find that

$$\begin{aligned} 0 &\leq \frac{t^2}{|t|^{2-p} + \varepsilon} < |t|^p \\ \lim_{\varepsilon \rightarrow 0} \frac{t^2}{|t|^{2-p} + \varepsilon} &= |t|^p. \end{aligned} \quad (\text{A.4})$$

Applying this approximation strategy to (A.3), model (A.3) can be rewritten in a matrix-vector form as

$$E(\mathbf{X}) = L \langle \mathbf{X} + e^{\mathbf{Y} - \mathbf{X}}, \mathbf{1}_{S \times 1} \rangle + \mathbf{X}^T \mathbf{G} \mathbf{X} \quad (\text{A.5})$$

where the matrix \mathbf{G} is

$$\begin{aligned} \mathbf{G} &= \lambda \beta (\mathbf{C}_h^T \mathbf{W}_h \mathbf{C}_h + \mathbf{C}_v^T \mathbf{W}_v \mathbf{C}_v) \\ &\quad + \lambda (1 - \beta) (\mathbf{C}_{hh}^T \mathbf{W}_{hh} \mathbf{C}_{hh} + \mathbf{C}_{hv}^T \mathbf{W}_{hv} \mathbf{C}_{hv} \\ &\quad \quad + \mathbf{C}_{vh}^T \mathbf{W}_{vh} \mathbf{C}_{vh} + \mathbf{C}_{vv}^T \mathbf{W}_{vv} \mathbf{C}_{vv}) \end{aligned} \quad (\text{A.6})$$

and $\mathbf{W}_h, \mathbf{W}_v, \mathbf{W}_{hh}, \mathbf{W}_{hv}, \mathbf{W}_{vh},$ and \mathbf{W}_{vv} are all $S \times S$ diagonal matrices with the diagonal elements defined as

$$\begin{aligned} (\mathbf{W}_h)_{i,i} &= (|(\mathbf{C}_h \mathbf{X})_i|^{2-p} + \varepsilon)^{-1} \\ (\mathbf{W}_v)_{i,i} &= (|(\mathbf{C}_v \mathbf{X})_i|^{2-p} + \varepsilon)^{-1} \\ (\mathbf{W}_{hh})_{i,i} &= (|(\mathbf{C}_{hh} \mathbf{X})_i|^{2-p} + \varepsilon)^{-1} \\ (\mathbf{W}_{hv})_{i,i} &= (|(\mathbf{C}_{hv} \mathbf{X})_i|^{2-p} + \varepsilon)^{-1} \\ (\mathbf{W}_{vh})_{i,i} &= (|(\mathbf{C}_{vh} \mathbf{X})_i|^{2-p} + \varepsilon)^{-1} \\ (\mathbf{W}_{vv})_{i,i} &= (|(\mathbf{C}_{vv} \mathbf{X})_i|^{2-p} + \varepsilon)^{-1}. \end{aligned} \quad (\text{A.7})$$

It should be noted that although $\mathbf{C}_{hv} = \mathbf{C}_{vh}$ and $\mathbf{W}_{hv} = \mathbf{W}_{vh}$, we still express them separately for the sake of the integrity. Therefore, we obtain a very simple expression of $E(\mathbf{X})$ as in (A.5)

ACKNOWLEDGMENT

The authors would like to thank Prof. Fatih Nar of Konya Food and Agriculture University for providing the codes of SDD [38]. They would also like to thank the editors and the anonymous reviewers for their careful reading of an earlier version of this article and constructive suggestions that improved the presentation of this article.

REFERENCES

- [1] J. R. Jensen, *Remote Sensing of the Environment: An Earth Resource Perspective*. Upper Saddle River, NJ, USA: Prentice-Hall, 2009.
- [2] J. W. Goodman, "Some fundamental properties of speckle," *J. Opt. Soc. Amer.*, vol. 66, no. 11, pp. 1145–1150, Nov. 1976.
- [3] G. Dong and G. Kuang, "Classification on the monogenic scale space: Application to target recognition in SAR image," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2527–2539, Aug. 2015.
- [4] H.-C. Li, T. Celik, N. Longbotham, and W. J. Emery, "Gabor feature based unsupervised change detection of multitemporal SAR images based on two-level clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2458–2462, Dec. 2015.
- [5] D. Xiang, T. Tang, Y. Ban, Y. Su, and G. Kuang, "Unsupervised polarimetric SAR urban area classification based on model-based decomposition with cross scattering," *ISPRS J. Photogramm. Remote Sens.*, vol. 116, pp. 86–100, Jun. 2016.
- [6] J.-S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 2, pp. 165–168, Mar. 1980.
- [7] V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman, "A model for radar images and its application to adaptive digital filtering of multiplicative noise," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, no. 2, pp. 157–166, Mar. 1982.
- [8] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel, "Adaptive noise smoothing filter for images with signal-dependent noise," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-7, no. 2, pp. 165–177, Mar. 1985.
- [9] F. Argenti, T. Bianchi, and L. Alparone, "Multiresolution MAP despeckling of SAR images based on locally adaptive generalized Gaussian pdf modeling," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3385–3399, Nov. 2006.

- [10] B.-H. Wang, C.-Y. Zhao, and Y.-Y. Liu, "An improved SAR interferogram denoising method based on principal component analysis and the goldstein filter," *Remote Sens. Lett.*, vol. 9, no. 1, pp. 81–90, Jan. 2018.
- [11] A. Achim, P. Tsakalides, and A. Bezerianos, "SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 8, pp. 1773–1784, Aug. 2003.
- [12] J. J. Ranjani and S. J. Thiruvengadam, "Dual-tree complex wavelet transform based SAR despeckling using interscale dependence," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2723–2731, Jun. 2010.
- [13] D. Gleich, M. Kseneman, and M. Datcu, "Despeckling of TerraSAR-X data using second-generation wavelets," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 68–72, Jan. 2010.
- [14] B. Hou, X. Zhang, X. Bu, and H. Feng, "SAR image despeckling based on nonsubsampling shearlet transform," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 3, pp. 809–823, Jun. 2012.
- [15] S. Liu, M. Shi, S. Hu, and Y. Xiao, "Synthetic aperture radar image de-noising based on shearlet transform using the context-based model," *Phys. Commun.*, vol. 13, pp. 221–229, Dec. 2014.
- [16] C.-A. Deledalle, L. Denis, and F. Tupin, "Iterative weighted maximum likelihood denoising with probabilistic patch-based weights," *IEEE Trans. Image Process.*, vol. 18, no. 12, pp. 2661–2672, Dec. 2009.
- [17] S. Parrilli, M. Poderico, C. V. Angelino, and L. Verdoliva, "A nonlocal SAR image denoising algorithm based on LLMMSE wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 606–616, Feb. 2012.
- [18] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [19] J. Geng, J. Fan, X. Ma, H. Wang, and K. Cao, "An iterative low-rank representation for SAR image despeckling," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 72–75.
- [20] G. Chen, G. Li, Y. Liu, X.-P. Zhang, and L. Zhang, "SAR image despeckling by combination of fractional-order total variation and nonlocal low rank regularization," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 3210–3214.
- [21] D. Guan, D. Xiang, X. Tang, and G. Kuang, "SAR image despeckling based on nonlocal low-rank regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3472–3489, Jun. 2019.
- [22] D. Guan, D. Xiang, X. Tang, and G. Kuang, "A SAR image despeckling method using multi-scale nonlocal low-rank model," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 421–425, Mar. 2020.
- [23] P. Wang, H. Zhang, and V. M. Patel, "SAR image despeckling using a convolutional neural network," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1763–1767, Dec. 2017.
- [24] F. Guo, G. Zhang, Q. Zhang, R. Zhao, M. Deng, and K. Xu, "Speckle suppression by weighted Euclidean distance anisotropic diffusion," *Remote Sens.*, vol. 10, no. 5, p. 722, May 2018.
- [25] F. Lattari, B. Gonzalez Leon, F. Asaro, A. Rucci, C. Prati, and M. Matteucci, "Deep learning for SAR image despeckling," *Remote Sens.*, vol. 11, no. 13, p. 1532, Jun. 2019.
- [26] W. Zhao, C.-A. Deledalle, L. Denis, H. Maitre, J.-M. Nicolas, and F. Tupin, "Ratio-based multitemporal SAR images denoising: RABASAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3552–3565, Jun. 2019.
- [27] S. Vitale, D. Cozzolino, G. Scarpa, L. Verdoliva, and G. Poggi, "Guided patchwise nonlocal SAR despeckling," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6484–6498, Sep. 2019.
- [28] L. Verdoliva, R. Gaetano, G. Ruello, and G. Poggi, "Optical-driven nonlocal SAR despeckling," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 314–318, Feb. 2015.
- [29] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
- [30] J. Nazarinezhad and M. Dehghani, "A contextual-based segmentation of compact PolSAR images using Markov random field (MRF) model," *Int. J. Remote Sens.*, vol. 40, no. 3, pp. 985–1010, Feb. 2019.
- [31] H. Xie, L. E. Pierce, and F. T. Ulaby, "SAR speckle reduction using wavelet denoising and Markov random field modeling," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2196–2212, Jan. 2002.
- [32] M. Mahdianpari, M. Motagh, and V. Akbari, "Speckle reduction and restoration of synthetic aperture radar data with an adoptive Markov random field model," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2012, pp. 276–279.
- [33] M. Mahdianpari, M. Motagh, V. Akbari, F. Mohammadimanes, and B. Salehi, "A Gaussian random field model for de-speckling of multi-polarized synthetic aperture radar data," *Adv. Space Res.*, vol. 64, no. 1, pp. 64–78, Jul. 2019.
- [34] Y. Chen, W. Feng, R. Ranftl, H. Qiao, and T. Pock, "A higher-order MRF based variational model for multiplicative noise reduction," *IEEE Signal Process. Lett.*, vol. 21, no. 11, pp. 1370–1374, Nov. 2014.
- [35] J.-F. Cai, B. Dong, S. Osher, and Z. Shen, "Image restoration: Total variation, wavelet frames, and beyond," *J. Amer. Math. Soc.*, vol. 25, no. 4, pp. 1033–1089, 2012.
- [36] Y. Zhao, J. G. Liu, B. Zhang, W. Hong, and Y.-R. Wu, "Adaptive total variation regularization based SAR image despeckling and despeckling evaluation index," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2765–2774, May 2015.
- [37] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Multiplicative noise removal using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1720–1730, Jul. 2010.
- [38] C. Ozcan, B. Sen, and F. Nar, "Sparsity-driven despeckling for SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 115–119, Jan. 2016.
- [39] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 492–526, 2010.
- [40] J.-J. Mei, T.-Z. Huang, S. Wang, and X.-L. Zhao, "Second order total generalized variation for speckle reduction in ultrasound images," *J. Franklin Inst.*, vol. 355, no. 1, pp. 574–595, Jan. 2018.
- [41] S. Wang, T.-Z. Huang, X.-L. Zhao, J.-J. Mei, and J. Huang, "Speckle noise removal in ultrasound images by first- and second-order total variation," *Numer. Algorithms*, vol. 78, no. 2, pp. 513–533, Jun. 2018.
- [42] W. Feng, H. Lei, and Y. Gao, "Speckle reduction via higher order total variation approach," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1831–1843, Apr. 2014.
- [43] H. Markarian and S. Ghofrani, "High-TV based CS framework using MAP estimator for SAR image enhancement," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 9, pp. 4059–4073, Sep. 2017.
- [44] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*. Rijeka, Croatia: SciTech, 2004.
- [45] A. Lopes, E. Nezry, R. Touzi, and H. Laur, "Structure detection and statistical adaptive speckle filtering in SAR images," *Int. J. Remote Sens.*, vol. 14, no. 9, pp. 1735–1758, Jun. 1993.
- [46] R. Farhadiani, S. Homayouni, and A. Safari, "Hybrid SAR speckle reduction using complex wavelet shrinkage and non-local PCA-based filtering," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 5, pp. 1489–1496, May 2019.
- [47] S. Lobry, L. Denis, and F. Tupin, "Multitemporal SAR image decomposition into strong scatterers, background, and speckle," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 8, pp. 3419–3429, Aug. 2016.
- [48] H. Li and Z. Lin, "Accelerated proximal gradient methods for nonconvex programming," in *Proc. NIPS*, Dec. 2015, pp. 379–387.
- [49] M. di Bisceglie and C. Galdi, "CFAR detection of extended objects in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 833–843, Apr. 2005.
- [50] G. Aubert and J.-F. Aujol, "A variational approach to removing multiplicative noise," *SIAM J. Appl. Math.*, vol. 68, no. 4, pp. 925–946, Jan. 2008.
- [51] H. Xie, L. E. Pierce, and F. T. Ulaby, "Statistical properties of logarithmically transformed speckle," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 3, pp. 721–727, Mar. 2002.
- [52] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, Jan. 2009.
- [53] T. Bianchi, F. Argenti, and L. Alparone, "Segmentation-based MAP despeckling of SAR images in the undecimated wavelet domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 9, pp. 2728–2742, Sep. 2008.
- [54] B. Xu, Y. Cui, Z. Li, B. Zuo, J. Yang, and J. Song, "Patch ordering-based SAR image despeckling via transform-domain filtering," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 4, pp. 1682–1695, Apr. 2015.
- [55] C.-A. Deledalle, L. Denis, F. Tupin, A. Reigber, and M. Jager, "NL-SAR: A unified nonlocal framework for resolution-preserving (Pol)(In)SAR denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2021–2038, Apr. 2015.
- [56] D. Cozzolino, S. Parrilli, G. Scarpa, G. Poggi, and L. Verdoliva, "Fast adaptive nonlocal SAR despeckling," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 2, pp. 524–528, Feb. 2014.

- [57] G. Chierchia, D. Cozzolino, G. Poggi, and L. Verdoliva, "SAR image despeckling through convolutional neural networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 5438–5441.
- [58] H. Feng, B. Hou, and M. Gong, "SAR image despeckling based on local homogeneous-region segmentation by using pixel-relativity measurement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 7, pp. 2724–2737, Jul. 2011.
- [59] D. Lazzaro, E. L. Piccolomini, and F. Zama, "A nonconvex penalization algorithm with automatic choice of the regularization parameter in sparse imaging," *Inverse Problems*, vol. 35, no. 8, Aug. 2019, Art. no. 084002.
- [60] F. Li, C. Shen, J. Fan, and C. Shen, "Image restoration combining a total variational filter and a fourth-order filter," *J. Vis. Commun. Image Represent.*, vol. 18, no. 4, pp. 322–330, Aug. 2007.
- [61] L. Tang, Y. Ren, Z. Fang, and C. He, "A generalized hybrid nonconvex variational regularization model for staircase reduction in image restoration," *Neurocomputing*, vol. 359, pp. 15–31, Sep. 2019.
- [62] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [63] G. Di Martino, M. Poderico, G. Poggi, D. Riccio, and L. Verdoliva, "Benchmarking framework for SAR despeckling," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1596–1615, Mar. 2014.
- [64] X. Nie, H. Qiao, and B. Zhang, "A variational model for PolSAR data speckle reduction based on the wishart distribution," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1209–1222, Apr. 2015.
- [65] C.-A. Deledalle, L. Denis, S. Tabti, and F. Tupin, "MuLoG, or how to apply Gaussian denoisers to multi-channel SAR speckle reduction?" *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4389–4403, Sep. 2017.



Yuli Sun received the B.S. degree from Xiangtan University, Xiangtan, China, in 2009, and the M.S. degree from the University of Science and Technology of China, Hefei, China, in 2014. He is pursuing the Ph.D. degree in information and communication engineering with the College of Electronic Science, National University of Defense Technology, Changsha, China.

His research interests cover machine learning and remote sensing image processing.



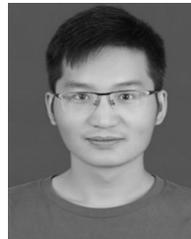
Lin Lei received the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2008.

She is an Associate Professor with the School of Electronic Science, National University of Defense Technology. Her research interests include computer vision, remote sensing image interpretation, and data fusion.



Dongdong Guan received the B.S. degree in surveying and mapping engineering from Wuhan University, Wuhan, China, in 2013, and the M.S. degree in photogrammetry and remote sensing and the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2015 and 2019, respectively.

He works in the High-Tech Institute of Xi'an. His research interests include synthetic aperture radar image processing, machine learning, and computation vision in remote sensing applications.



Xiao Li received the B.S. degree in electrical engineering and automation from the University of Jinan, Jinan, China, in 2015, and the M.S. degree in control science and engineering from Xiangtan University, Xiangtan, China, in 2018. He is pursuing the Ph.D. degree in information and communication engineering with the National University of Defense Technology, Changsha, China.

His research interests include image processing and pattern recognition, representation and dictionary learning, and computational pathology applications.



Gangyao Kuang (Senior Member, IEEE) received the B.S. and M.S. degrees in geophysics from the Central South University of Technology, Changsha, China, in 1988 and 1991, respectively, and the Ph.D. degree in communication and information from the National University of Defense Technology, Changsha, in 1995.

He is a Professor with the School of Electronic Science, National University of Defense Technology. His research interests include remote sensing, SAR image processing, change detection, SAR ground moving target indication, and classification with polarimetric SAR images.