

# Multitemporal SAR Image Despeckling Based on Block-Matching and Collaborative Filtering

Giovanni Chierchia, *Member, IEEE*, Mireille El Gheche, Giuseppe Scarpa, *Member, IEEE*,  
and Luisa Verdoliva, *Member, IEEE*

**Abstract**—We propose a despeckling algorithm for multitemporal synthetic aperture radar (SAR) images based on the concepts of block-matching and collaborative filtering. It relies on the nonlocal approach, and it is the extension of SAR-BM3D for dealing with multitemporal data. The technique comprises two passes, each one performing grouping, collaborative filtering, and aggregation. In particular, the first pass performs both the spatial and temporal filtering, while the second pass only the spatial one. To avoid increasing the computational cost of the technique, we resort to lookup tables for the distance computation in the block-matching phases. The experiments show that the proposed algorithm compares favorably with respect to state-of-the-art reference techniques, with better results both on simulated speckled images and on real multitemporal SAR images.

**Index Terms**—Despeckling, multitemporal, nonlocal filter, synthetic aperture radar (SAR).

## I. INTRODUCTION

EXTRACTING useful information from remote sensing data is much easier nowadays, thanks to the availability of advanced sensors that produce high-resolution images of the earth. This is especially true for synthetic aperture radar (SAR) multitemporal systems, which are able to provide images of the same area in different temporal intervals, allowing for a continuous monitoring of the earth surface. The availability of this data is of great importance for both land-cover classification [2]–[4] and for the analysis of environmental changes [5]–[7].

However, a major problem of SAR images is the presence of speckle, which strongly impairs the performance of the aforementioned tasks. This called for an intense research activity on SAR despeckling in the past decade [8]. The major difficulties of despeckling are the strong noise intensity (especially in single-look images) and the high variability of

physical properties and statistics, which prevents using a single model for the whole image. In order to face these issues, the first filters proposed for SAR images are spatial estimators based on the linear MMSE approach with some form of adaptivity to the local image content [9]–[11]. In this way, the level of smoothing is modulated according to the region heterogeneity.

Nevertheless, it is not easy to achieve a good smoothing capability without losing spatial details and fine structures, due to the nonstationarity of SAR data. In this regard, the main merit of the above spatial filters is the low computational complexity. Different strategies can be pursued to handle the nonstationarity: choosing large windows while using different geometrical ratio detectors [12] or selecting only statistically homogeneous data inside the analysis window [13].

The nonstationarity issue is even more significant when filtering multitemporal data, since in this case it is important to preserve both spatial and temporal resolutions. If it is acceptable that in some areas, like buildings and homogeneous fields, the spatial response remains unchanged over short period, this is not true anymore for more heterogeneous regions or when observing long period time series. To avoid radiometric degradation in filtered images, several approaches have been proposed in the selection of the processing window.

In the context of polarimetric (Pol-SAR) and interferometric (InSAR) filtering, a region growing technique was proposed in [14]. In this paper, an adaptive neighborhood is built by testing each pixel so as to ensure the validity of the stationarity assumption within the window, and the similarity test is carried out exclusively on the amplitude distance. A similar approach is also used in [15] to filter multitemporal data, where a 3-D adaptive neighborhood is formed by taking into account both spatial and temporal information.

A simple, yet effective, denoising filter for multitemporal SAR data was proposed by De Grandi *et al.* [16]. Specifically, in order to avoid filtering data with different statistics caused by temporal change variations, large windows are used for homogeneous areas, while for nonhomogeneous areas a series of detectors with progressively smaller windows are used to locate edges, fine structures, or point targets. Filtering is then carried out with a basic linear MMSE estimator [17] or minimum-variance unbiased estimator [18]. The latter estimator was also adopted in [19] and [20], leading to an unbiased temporal averaging tailored to the speckle noise.

In more recent approaches, the adaptive spatio-temporal neighborhood is derived by analyzing the temporal evolution of the data directly from the change detection

Manuscript received May 23, 2016; revised October 19, 2016, February 15, 2017, and April 18, 2017; accepted May 10, 2017. This work was supported in part by the Italian Ministry of Education, University, and Research, within the framework of the PRIN 2012 project “Very High Spatial and Spectral Resolution Remote Sensing: A Novel Integrated Data Analysis System,” and in part by the CNRS MASTODONS project under Grant 2016TABASCO. (Corresponding author: Luisa Verdoliva.)

G. Chierchia is with the Université Paris Est, LIGM (UMR 8049), CNRS, ENPC, ESIEE Paris, UPEM, F-93162, Noisy-le-Grand, France (e-mail: giovanni.chierchia@esiee.fr).

M. El Gheche is with the Université de Bordeaux, F-33400 Talence, France (e-mail: mireille.el-gheche@u-bordeaux.fr).

G. Scarpa and L. Verdoliva are with the DIETI, Università Federico II di Naples, 80138 Napoli, Italy (e-mail: giscarpa@unina.it; verdoliv@unina.it).

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

Digital Object Identifier 10.1109/TGRS.2017.2707806

matrices [21], [22]. These matrices are composed of similarity cross-test responses based on the coefficient of variation, and they contain information on changed and unchanged pixels.

In the last few years, the nonlocal paradigm has been used in several techniques for different SAR imaging modalities [23], such as amplitude SAR [1], [24], [25], Pol-SAR [26], InSAR [27], Pol-InSAR [28], and differential InSAR [29]. The main reason of their success lies in the ability to group together and jointly filter only the pixels that are considered “similar” through an appropriate patch-based measure [30]. While this approach is conceptually equivalent to searching for the best adaptive neighborhood, it turns out to be more effective, due to the specific characteristics of SAR images.

In the nonlocal context, Su *et al.* [31] recently proposed a multitemporal extension of the single-image SAR despeckling algorithm named Probabilistic Patch-Based denoising algorithm (PPB) [24]. The extended method, called 2S-PPB, is a two-step procedure that takes advantage of a simple change detection method in order to properly filter the data along the temporal direction with classical PPB.

In this paper, we present a new multitemporal SAR despeckling algorithm. First, we propose a nonlocal version of the minimum-variance unbiased estimator for multicomponent images degraded by multiplicative noise [18], which leads to a remarkable improvement of the single-image SAR despeckling technique called SAR-BM3D [1]. Based on this observation, we then propose a multitemporal oriented version of SAR-BM3D, which includes a block-matching phase tailored to multitemporal SAR images and accelerated through the use of lookup tables, as well as a 4-D collaborative filtering that exploits the proposed nonlocal temporal filter.

This paper is organized as follows. Section II describes Nonlocal Temporal Filter (NLTF), the nonlocal filter that we are going to adopt along the temporal direction. Section III details multitemporal SAR-BM3D (MSAR-BM3D), the multitemporal algorithm based on SAR-BM3D. Section IV illustrates results of experiments carried out on both synthetic and real multitemporal SAR images. Finally, Section V draws the conclusion.

## II. TEMPORAL FILTERING

A multitemporal image consists of several SAR images (referred to as “components” in the following) acquired at  $M$  different instants. Under the hypothesis of fully developed speckle, these components can be expressed as

$$z_i(s) = x_i(s) u_i(s) \quad i = 1, \dots, M \quad (1)$$

where for each position  $s$  in the image lattice,  $z_i(s)$  is the backscattered signal,  $x_i(s)$  is the noise-free reflectance, and  $u_i(s)$  is the speckle in intensity format, characterized by the independence from  $x_i(s)$ , a unitary mean, and a standard deviation  $\sigma = 1/\sqrt{L}$ , with  $L$  being the number of looks.

Furthermore, it is assumed that the same resolution element on the ground is illuminated by the radar beam in the same way, so that it corresponds to the same position in the image lattice for all the components. This can be reasonably satisfied

through the use of overlapping paths in the SAR acquisition phase, and a proper spatial coregistration of the components.

In the literature, several techniques have been proposed in order to filter along the temporal direction. We will first review the basic approaches (Section II-A) and then describe the proposed nonlocal solution (Section II-B).

### A. Related Work

A principled way to reduce the speckle in a multitemporal image amounts to averaging the pixels  $z_1(s), \dots, z_M(s)$ . This is optimal when the underlying reflectances  $x_1(s), \dots, x_M(s)$  are identical, as the speckle standard deviation is reduced by a factor  $M$ , similar to a multilook processing. However, in a more realistic situation, the reflectance can vary from one component to another, due to a change in the dielectric and geometrical properties of the elementary scatterers. Hence, adjustments need to be made.

An effective manner to take into account the temporal changes can be achieved by introducing appropriate weights in the average process, so as to balance the difference in reflectivity at different times. This leads to the linear estimator

$$\hat{x}_i(s) = \mu_i \sum_{k=1}^M \frac{\alpha_k}{\mu_k} z_k(s) \quad (2)$$

where  $\mu_i = \mathbb{E}\{z_i\}$  denotes the expected mean of  $z_i$  (under the assumption of stationarity) and  $\alpha = [\alpha_1, \dots, \alpha_M]^T$  is a vector of weights to be determined. Note that the above estimator is unbiased when  $\sum_{k=1}^M \alpha_k = 1$ , since  $\mathbb{E}\{z_i\} = \mathbb{E}\{x_i\}$ . Therefore, one can simply select constant weights in order to obtain an unbiased temporal averaging, that is

$$\alpha_k = \frac{1}{M} \quad (3)$$

as done in [19] and [20].

From a statistical standpoint, a possible approach to find the “optimal” weights consists of requiring (2) to be a minimum-variance unbiased estimator as proposed in [18]. Assuming a spatial homogeneity for the reflectivity, namely,  $x_k(s) = \mathbb{E}\{x_k\}$  (which is certainly true inhomogeneous areas), the optimal weights can be expressed as the solution of the linear system

$$\alpha = A^{-1}c \quad (4)$$

where  $c = [10 \dots 0]^T$  is a (column) vector of length  $M$ , and  $A = [A_{ik}]_{1 \leq i \leq M, 1 \leq k \leq M}$  is a nonsingular matrix defined as

$$A_{ik} = \begin{cases} 1, & \text{if } i = 1 \\ \rho_{1k} - \rho_{ik}, & \text{if } i = 2, \dots, M. \end{cases} \quad (5)$$

Hereabove,  $\rho_{ik}$  denotes the correlation coefficient between  $z_i$  and  $z_k$ , namely

$$\rho_{ik} = \frac{\mathbb{E}\{(z_i - \mu_i)(z_k - \mu_k)\}}{\sqrt{\mathbb{E}\{(z_i - \mu_i)^2\}\mathbb{E}\{(z_k - \mu_k)^2\}}} \quad (6)$$

The performance of the weighted average in (2) strongly depends on the ability of accurately estimating the expected means  $\mu_i$  and correlation coefficients  $\rho_{ik}$ . Of course, the stationarity hypothesis advocated earlier is not very realistic, as these statistics can vary significantly from one position to

another of the image lattice. It is therefore crucial to introduce some form of adaptivity in the statistic estimation process.

A possible way to adaptively estimate the statistics consists of using a window centered around the pixel of interest [19]. However, a typical problem of local estimators is that the window should be fairly large in order to obtain reliable estimates. This inevitably leads to errors in correspondence of edges, fine structures, and point targets, since the data around these areas are unlikely to be homogeneous. A different approach is followed in [31] where change detection guides the averaging process. Since only stable pixels should be involved in the temporal denoising, this method employs binary weights that are set to zero whenever a temporal change is detected.

### B. Proposed Nonlocal Temporal Filter

A powerful approach to perform temporal filtering is given by the nonlocal paradigm. The idea amounts to collecting a number of blocks with the same reflectivity and differing only in the noise realization. This is feasible because the images exhibit self-similarities, in the sense that most blocks repeat almost identically over and over in the image. Once these similar blocks are identified, one can rely on a large amount of homogeneous data (even with small blocks), which can significantly improve the estimation of statistics with respect to the sliding-window approach.

The procedure works as follows. For a given position  $s$ , a 3-D patch is built by extracting the  $N_1 \times N_1$  blocks around the position  $s$  of every component. The latter is then compared with the 3-D patches located in an  $N_w \times N_w$  window centered in  $s$ , using an appropriate measure (see Section III-A). Denoting by  $\mathcal{N}_s = \{s_1, \dots, s_{N_2}\}$  the positions of the most similar patches (which include  $s$  itself),<sup>1</sup> and by  $\mathcal{W}_{N_1} = \{p_1, \dots, p_{N_1^2}\}$  the spatial coordinates within an  $N_1 \times N_1$  window, the means and variances are estimated as

$$\mu_i(s) = \frac{1}{N_2 N_1^2} \sum_{n=1}^{N_2} \sum_{j=1}^{N_1^2} z_i(s_n + p_j) \quad (7)$$

$$\sigma_i^2(s) = \frac{1}{N_2 N_1^2} \sum_{n=1}^{N_2} \sum_{j=1}^{N_1^2} [z_i(s_n + p_j) - \mu_i(s)]^2. \quad (8)$$

It is worth noting that the 3-D patch does not need to be homogeneous. The correlation coefficients are estimated as

$$\rho_{i,k}(s) = \frac{\sum_{n=1}^{N_2} \sum_{j=1}^{N_1^2} [z_i(s_n + p_j) - \mu_i(s)][z_k(s_n + p_j) - \mu_k(s)]}{\sigma_i(s)\sigma_k(s)N_2 N_1^2}. \quad (9)$$

Once the statistics are estimated, the weight vector  $\alpha(s)$  is computed as in (4), and then the pixels in the selected blocks are filtered as in (2), yielding for every  $s_n \in \mathcal{N}_s$  and  $p_j \in \mathcal{W}_{N_1}$

$$\hat{x}_i(s_n + p_j) = \mu_i(s) \sum_{k=1}^M \frac{\alpha_k(s)}{\mu_k(s)} z_k(s_n + p_j). \quad (10)$$

<sup>1</sup>Note that the set  $\mathcal{N}_s$  is independent of the index  $i$ . Hence, the chosen positions are the same in all the temporal components.

This procedure is repeated for each position  $s$  in the image lattice (or a subset of), obtaining multiple filtered versions of the same reflectance  $x_i(s)$ , which are averaged together. Note that our approach is different from the 3-D-oriented versions of nonlocal means [32], [33], which average the pixels along the spatial and temporal directions (within a 3-D neighborhood), whereas we only perform the average along the temporal axis.

The block-matching procedure does a great job in detecting repetitive structures such as roads, rivers, and man-made targets, improving the statistic estimation in correspondence of such regions. However, a more difficult situation occurs for isolated bright targets, like corner reflectors, due to the unavailability of similar patches in their neighborhood. This problem, known as *rare patch effect* [34], [35], causes an impairment of the despeckled signal and gives rise to a noisy area around the point-like structure (noise halo). Different adaptive solutions can be found in the literature to face this issue [28], [34]. In this paper, we detect these bright targets in advance and inhibit any filtering around them. This procedure allows us to preserve important image features, even though it does not avoid the noise halo problem. Target detection is performed based on the ratio between the local variance and the square mean (estimated through a small sliding window)

$$r(s) = \max_{1 \leq i \leq M} \frac{\tilde{\sigma}_i^2(s)}{\tilde{\mu}_i^2(s)}. \quad (11)$$

When the above ratio is larger than a predefined threshold  $\lambda_c$ , the pixels  $z_1(s), \dots, z_M(s)$  are left unaltered.

### III. MULTITEMPORAL SAR-BM3D

The NLTF proposed in Section II can be regarded as an attempt to reduce the speckle without affecting the spatial or temporal resolution. To clarify this concept, Fig. 1 shows a synthetic multitemporal SAR image, along with its despeckled versions obtained using NLTF, SAR-BM3D [1], and a combination of the two.

One can see in Fig. 1(b) that the proposed NLTF effectively reduces the speckle, but does not completely remove it, as no spatial smoothing is performed. Hence, a natural idea consists of integrating NLTF into a despeckling algorithm that actually performs spatial smoothing, such as SAR-BM3D. This intuition is confirmed in Fig. 1(c) and (d), where SAR-BM3D provides a better result when applied to the image prefiltered with NLTF, thanks to the reduced speckle in the latter.

However, a drawback of processing a multitemporal image with SAR-BM3D stems from the fact that each component is despeckled independently. This obviously neglects the temporal correlation of regions that are stationary in time, such as areas with a significant orography (e.g., mountains, craters), natural formations (e.g., rivers, lakes), or man-made structures (e.g., roads, buildings). Of course, the time-stationary targets vary according to the thematic application.

The temporal correlation is the key for improving the despeckling of multitemporal SAR images. To this end, we propose to integrate NLTF with SAR-BM3D, leading to a new algorithm that we called MSAR-BM3D. The processing flow of MSAR-BM3D comprises two passes. In each

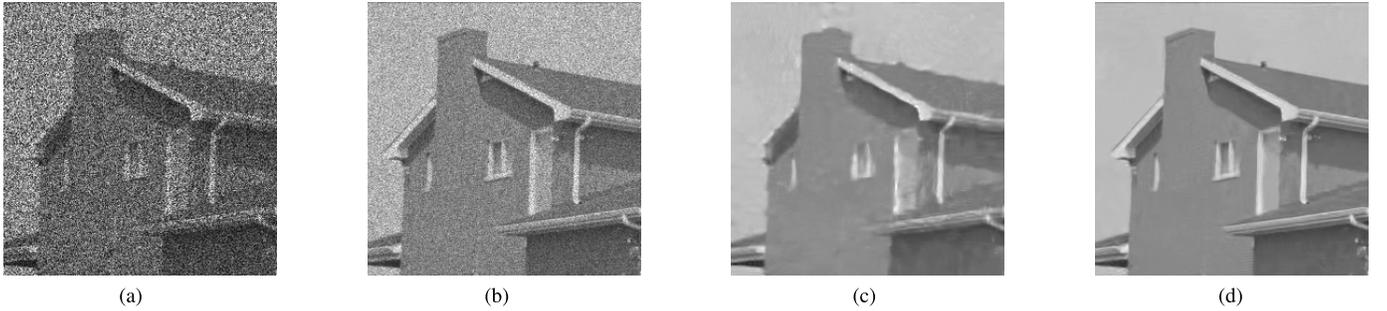


Fig. 1. Despeckling of a synthetic single-look SAR image with eight temporal components (only the first component is shown). (a) Noisy. (b) NLTF. (c) SAR-BM3D. (d) NLTF + SAR-BM3D.

one, the image lattice is scanned block-wise with partial overlapping, and for every visited location, three steps are performed: grouping, collaborative filtering, and aggregation. The first pass, operating on the noisy image, works as follows.

- 1) *Grouping*: For every visited position, the most similar blocks in each temporal component are stacked together.
- 2) *Collaborative Filtering*: Each 4-D group is temporally filtered, transformed along the four axes, hard thresholded, and inverse transformed.
- 3) *Aggregation*: All filtered blocks are returned to their original locations and contribute with suitable weights to the basic estimate of the image.

The second-pass works similarly, with the following differences.

- 1) *Grouping*: The search for similar blocks also integrates the basic estimate provided by the first pass.
- 2) *Collaborative Filtering*: Each 4-D group (of noisy blocks) undergoes a 3-D or 4-D transform without the temporal pre-filtering, an empirical Wiener filtering (driven by the basic estimate), and an inverse transform.
- 3) *Aggregation*: Like the first pass.

The above phases are sketched in Fig. 2. Note that the temporal filtering comes into play exclusively in the first pass, after the grouping step, where the grouped blocks are used to compute the statistics, and then filtered in the temporal direction. Moreover, the stacks undergo collaborative filtering and aggregation in both passes, similar to SAR-BM3D [1].

#### A. Grouping

The grouping can be regarded as an attempt (limited by complexity and data scarcity) to collect blocks with homogeneous data and differing only in the noise realization. To do so, one actually computes the distances between the target block and all the blocks in a surrounding search window, in order to select the most similar ones. In the presence of speckle, the “optimal” distance for comparing noisy patches stems from the generalized likelihood ratio test [30].

For the despeckling of multitemporal images, one could collect similar data by matching 3-D patches within a 3-D search window, as proposed in BM4D [36]. In this paper, however, we adopt a different strategy: the block-matching is carried over the 2-D image obtained as the arithmetic mean of the temporal components. This yields several advantages. First, the number of looks is uniformly increased, and thus the

attenuated speckle ensures that the distances are less noisy, leading to a better localization of similar blocks. Second, the same positions are used to select the blocks in each component, leading to the grouping of 3-D patches with a full temporal extent. Third, the reduced size of patches speeds up the calculations.

More specifically, in the first pass of MSAR-BM3D, we carry out the block matching over the image obtained by averaging the noisy temporal components. Since the temporally averaged pixels have the same number of looks, we stick to the classical formula for comparing noisy patches

$$D_{1st}(a, b) = (2ML - 1) \sum_{j=1}^{N_1^2} \log \left( \frac{a(j)}{b(j)} + \frac{b(j)}{a(j)} \right). \quad (12)$$

In the second pass, we take into account the additional information provided by the first pass using the formula

$$D_{2nd}(a, b, c, d) = \sum_{j=1}^{N_1^2} \left[ (2ML - 1) \log \left( \frac{a(j)}{b(j)} + \frac{b(j)}{a(j)} \right) + \gamma ML \frac{|c(j) - d(j)|^2}{c(j)d(j)} \right] \quad (13)$$

where  $a$  and  $b$  denote two patches of the (temporally averaged) noisy image and  $c$  and  $d$  denote two patches of the (temporally averaged) image estimated in the first pass. The parameter  $\gamma$  weighs the relative importance of noisy data and prefiltered data. Preliminary experiments show  $\gamma = 1$  to be a reasonable choice, as was already observed in [1], so we adhere to this choice. It is worth mentioning that the above “averaged images” only serve for grouping similar patches, and they play no role in the next steps of the algorithm.

Moreover, to further reduce the computational burden, we compute the pixel-wise distances by resorting to a lookup table [25]. This requires a quantization of the noisy image by an  $N_Q$ -level quantizer  $\mathcal{Q}$ . If  $q_n = \mathcal{Q}(x)$  and  $q_m = \mathcal{Q}(y)$ , then

$$\log \left( \frac{x}{y} + \frac{y}{x} \right) \simeq \log \left( \frac{q_n}{q_m} + \frac{q_m}{q_n} \right) = \text{LUT}(n, m). \quad (14)$$

The  $N_Q \times N_Q$  lookup table is computed in advance.

In order to keep a good accuracy, the quantization must be reasonably dense, especially for small values of the input, since the distance diverges when one of them approaches zero.

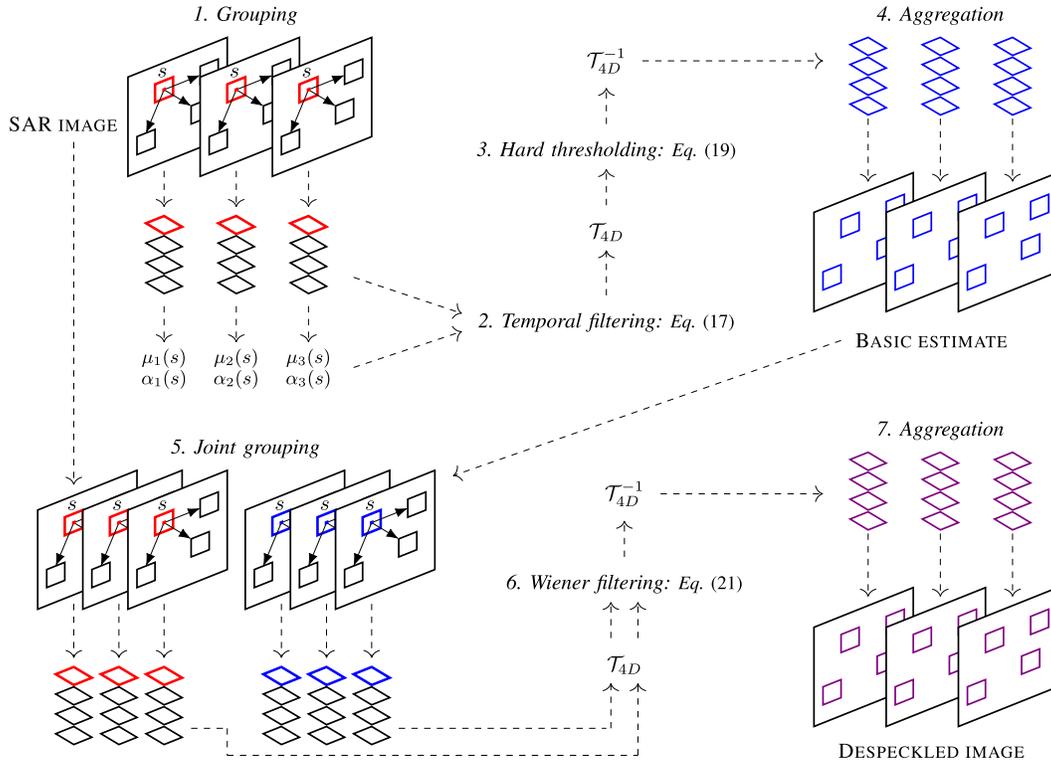


Fig. 2. Flowchart of the principal steps performed by MSAR-BM3D.

Therefore, we carry out a uniform quantization in the log domain, within the range dictated by the extreme values of the input image, that are  $x_{\min} > 0$  and  $x_{\max}$ . Once computed the quantization step as  $\Delta = \log_2(x_{\max}/x_{\min})/(N_Q - 1)$ , we set

$$Q(x) = x_{\min} 2^{n\Delta}, \quad n = \left\lfloor \frac{\log_2(x/x_{\min})}{\Delta} \right\rfloor. \quad (15)$$

With such a quantizer, however, it is not really necessary to compute and store a matrix of distances, since

$$\text{LUT}(n, m) = \log_2[2^{(n-m)\Delta} + 2^{(m-n)\Delta}] \quad (16)$$

depends uniquely on the difference of the indexes. Hence, only a vector of size  $2N_Q - 1$  is needed, thus reducing both memory occupation and access time.

### B. Collaborative Filtering

The collaborative filtering arises from the interplay among the nonlocal paradigm [37], the wavelet shrinkage [38], and the empirical Wiener filter [39]. The principle is that the transformation of grouped blocks yields sparser coefficients than the blocks transformed alone. This allows for better separation of the noise from the clean image, especially if the coefficients are denoised by a two-pass Wiener filter.

For the despeckling of multitemporal images, the first pass involves the prefiltering of grouped blocks, in order to attenuate the speckle through a temporal averaging. Denoting by  $\{s_n\}_{1 \leq n \leq N_2}$  the positions retrieved by the block-matching,

the filtered group can be expressed as

$$\widehat{X}(j, n, i) = \begin{cases} z_i(s_n + p_j) & \text{if } r(s_n + p_j) > \lambda_c \\ \mu_i(s) \sum_{k=1}^M \frac{\alpha_k(s)}{\mu_k(s)} z_k(s_n + p_j) & \text{otherwise} \end{cases} \quad (17)$$

where  $\alpha_k(s)$  and  $\mu_k(s)$ , being calculated as in Section II-B, depend on the position  $s$  of the reference block used in the block-matching phase. The above group is then transformed along the four axes, yielding

$$Y = \mathcal{T}_{4D}(\widehat{X}) \quad (18)$$

and the resulting coefficients are hard-thresholded as follows:

$$\widehat{Y}_{1st}(j, n, i) = \begin{cases} 0, & \text{if } Y(j, n, i) \leq \lambda_{4D} \\ Y(j, n, i), & \text{otherwise.} \end{cases} \quad (19)$$

In the second pass, the temporal prefiltering is omitted, so as to preserve all the information carried by the noisy pixels. The group of noisy blocks is thus transformed right away

$$Z = \mathcal{T}_{4D}(\{\widehat{Y}_{1st}(j, n, i)\}_{j,n,i}) \quad (20)$$

and denoised through a Wiener filtering

$$\widehat{Y}_{2nd}(j, n, i) = \frac{\sigma_{1st}^2(j, n, i)}{\sigma_{1st}^2(j, n, i) + \sigma_{\text{noise}}^2(j, n, i)} Z(j, n, i) \quad (21)$$

where the above variances are inferred from the image obtained at the first pass. Note that the transformation along the temporal axis is optional in the second step, and it can be omitted if one wants to preserve even the smallest temporal changes (at the cost of a slightly inferior despeckling).

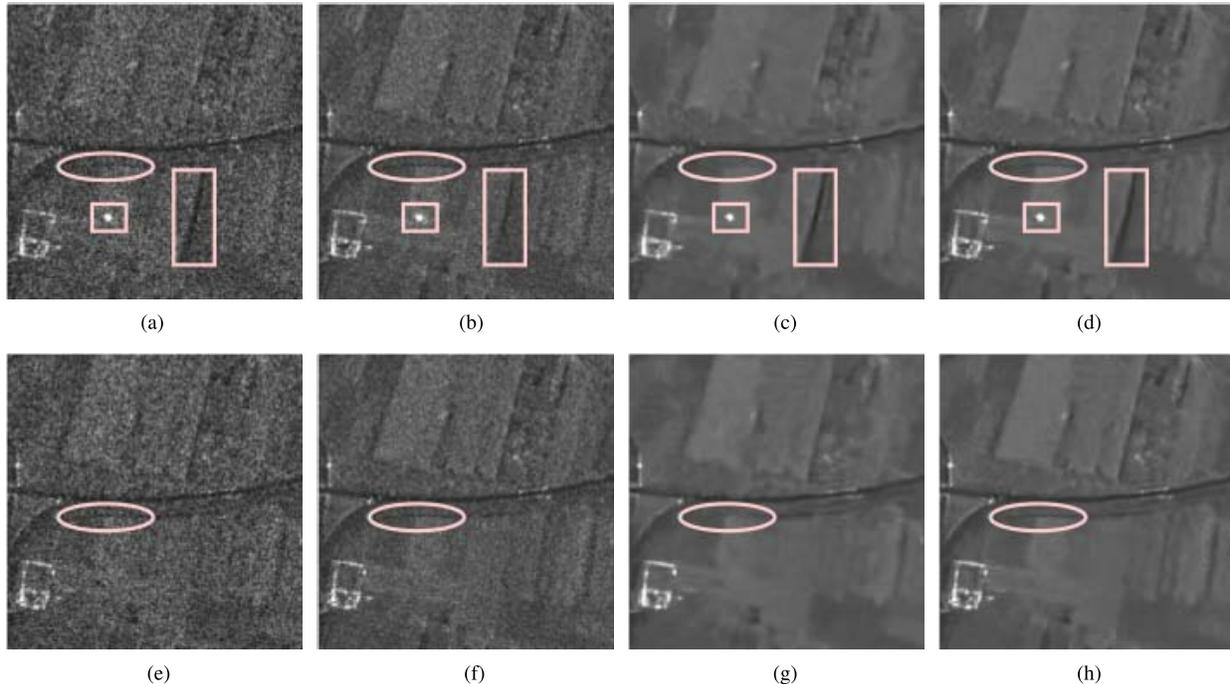


Fig. 3. Despeckling of a realistic SAR image with four components degraded by a one-look simulated speckle (only a zoom of two components is shown). (a) COSMO-SkyMed—component 1. (b) NLTF. (c) SAR-BM3D. (d) MSAR-BM3D. (e) COSMO-SkyMed—component 2. (f) NLTF. (g) SAR-BM3D. (h) MSAR-BM3D.

### C. Aggregation

The aggregation consists of averaging the pixels that, being included in more than one group, are estimated several times

$$\hat{x}_i(s) = \frac{1}{T_s} \sum_{g \in \mathcal{G}(s)} w_g \hat{x}_i(g) \quad (22)$$

where  $\mathcal{G}(s)$  is the set of groups comprising  $s$ ,  $\hat{x}_i(g)$  is the estimate provided by the group  $g$  after an inverse transform,  $w_g$  is the corresponding weight, and  $T_s$  is a normalizing factor. Like in SAR-BM3D, the weights depend on the presumed reliability of the associated group estimate, related in turn to the average noise power of the group after shrinkage.

## IV. EXPERIMENTAL RESULTS

In SAR image denoising, the performance evaluation of different approaches is quite a challenging task, because of the lack of original noiseless signals. Hence, we split the numerical validation in three parts. First, we compare the performance of multilooking, sole temporal or spatial filtering, and spatio-temporal despeckling, using real and simulated SAR images. Second, we present experiments carried out on synthetic SAR images corrupted by simulated speckle, obtaining numerical results and figures that allow a sound comparison among different state-of-the-art algorithms. Third, we finalize the experiments with real SAR image denoising using equivalent number of looks (ENL),  $\alpha\beta$  index [40], and visual inspection. Note that, in order to guarantee reproducibility, the executable code of the proposed technique is available online ([www.grip.unina.it](http://www.grip.unina.it)).

### A. Reference Techniques and Parameter Setting

We compare the proposed techniques with the four state-of-the-art despeckling algorithms listed in the following:

- 1) the unbiased temporal average (UTA) [19];
- 2) the two-step PPB nonlocal filter (2S-PPB) [31];
- 3) the SAR-oriented version of BM3D (SAR-BM3D) [1], adapted to the multitemporal case through a simple component-by-component approach<sup>2</sup>;
- 4) the extension of BM3D to volumetric data (BM4D) [36], adapted to the speckle noise through the homomorphic approach (with mean-bias correction).<sup>3</sup>

Such techniques have been chosen because of their competitive performance and (not least) for the availability of software code to run the experiments.

For all these algorithms, if not stated otherwise, the free parameters are set as suggested in the reference papers. As for the proposed MSAR-BM3D algorithm, in the first pass, we use groups of size  $8 \times 8 \times 16 \times M$  (remember that  $M$  is the number of components), which are transformed with a biorthogonal 2-D discrete wavelet transform (DWT) along the pixels of a block, a Haar 1-D DWT along the blocks of a group, and a Haar 1-D DWT along the components, all with the maximum-level decomposition. Just like in SAR-BM3D, the computational burden is reduced using a relatively small search area of size  $39 \times 39$ , and by selecting reference blocks only on every fourth row and column. Similar choices apply to the second pass, except for the fact that the group dimensions grow to  $8 \times 8 \times 32 \times M$ , and the biorthogonal 2-D DWT is replaced by a spatial discrete cosine transform. Finally, the weights  $w_g$  in (22) are the same as SAR-BM3D.

<sup>2</sup><http://www.grip.unina.it/research/80-sar-despeckling/80-sar-bm3d.html>

<sup>3</sup><http://www.cs.tut.fi/~foi/GCF-BM3D/index.html>

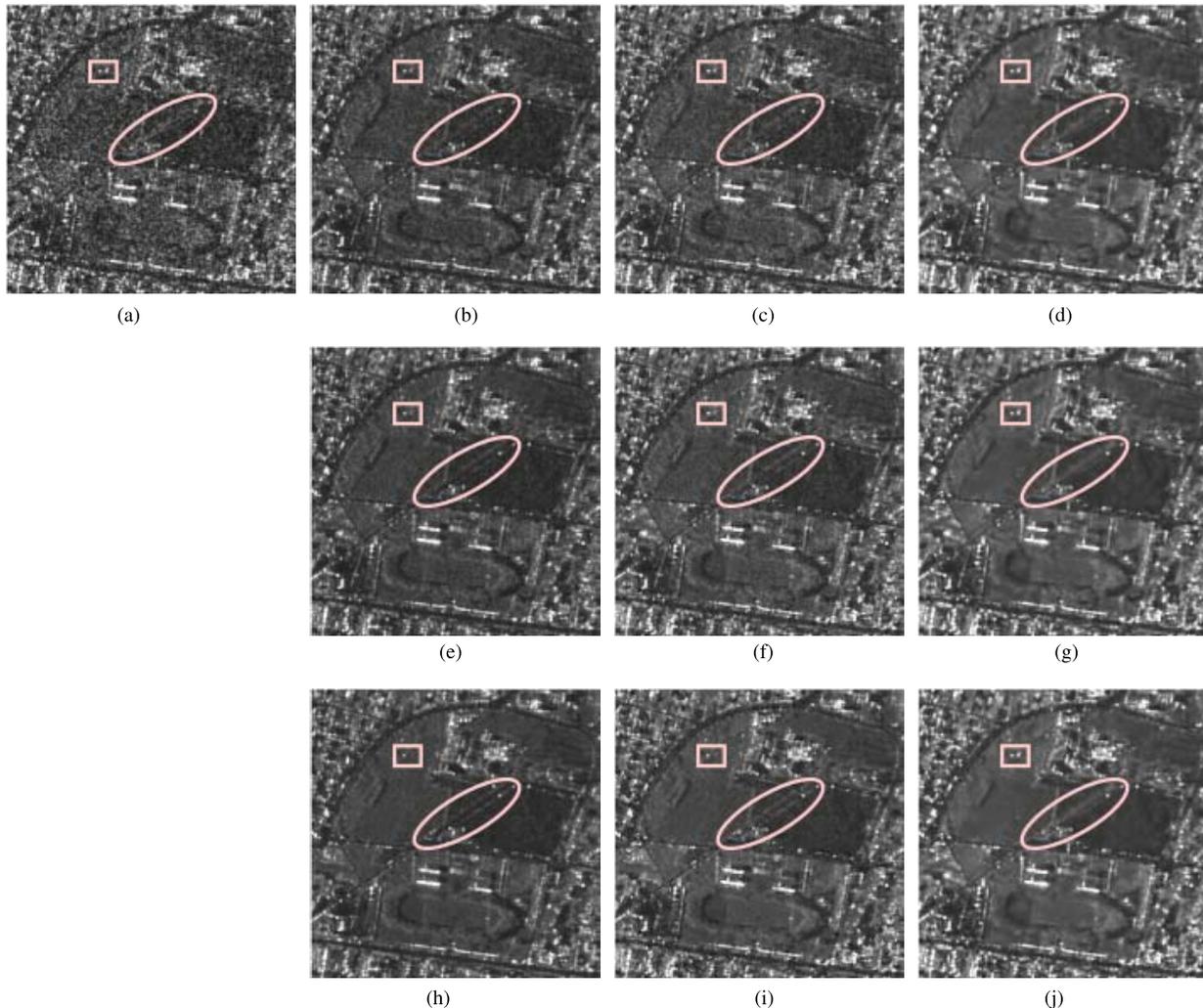


Fig. 4. Despeckling of a single-look COSMO-SkyMed image with an increasing number of components (only a zoom of the first component is shown). (a) COSMO-SkyMed. (b) Multilooking ( $M = 4$ ). (c) NLTF ( $M = 4$ ). (d) MSAR-BM3D ( $M = 4$ ). (e) Multilooking ( $M = 8$ ). (f) NLTF ( $M = 8$ ). (g) MSAR-BM3D ( $M = 8$ ). (h) Multilooking ( $M = 32$ ). (i) NLTF ( $M = 32$ ). (j) MSAR-BM3D ( $M = 32$ ).

### B. Comparison With Temporal or Spatial Filtering

1) *Realistic SAR Images*: In our first set of experiments, we consider a 32-look COSMO-SkyMed image and build a multitemporal image with four identical components,<sup>4</sup> and multiply each of them by a one-look speckle in amplitude format (square-root intensity model). In order to simulate some temporal changes, we add a dark line and a bright target to the first component, which are indicated in Fig. 3(a) by two rectangles, whereas the ellipse highlights a stable object. Fig. 3 presents the results obtained with a temporal filtering (NLTF), a spatial filtering (SAR-BM3D), and a spatio-temporal filtering (MSAR-BM3D). One can observe that the proposed MSAR-BM3D preserves edges better than SAR-BM3D (see the white ellipses), while temporal changes are better restored than with NLTF (see the white rectangles).

2) *Real Multitemporal SAR Images*: In our second set of experiments, we consider a one-look COSMO-SkyMed image with an increasing number of components  $M \in \{4, 8, 32\}$ .

<sup>4</sup>Images taken over the region north of Naples, Italy; pixel size:  $2 \times 2$  meters.

Fig. 4 compares the images obtained with multilooking, NLTF, and MSAR-BM3D. One can see that multilooking cannot preserve temporal changes (see the white ellipses), while the sole temporal filtering does not always guarantee the target preservation (see the white rectangles). Instead, MSAR-BM3D is able to preserve both the temporal and spatial resolutions, while reducing speckle. Also note that, for all the considered techniques, the despeckling ability improves as the number of components  $M$  grows larger.

### C. State-of-the-Art Comparison on Synthetic Images

In our third set of experiments, we generate a number of SAR-like images using optical images multiplied with a simulated speckle in amplitude format (square-root intensity model) [41] with pdfs corresponding to the cases of  $L = 1, 4, 8$  looks. The performance is quantified by the signal-to-noise ratio (SNR) and the structural similarity (SSIM).

In Table I, we present the results obtained on three  $512 \times 512$  images. The best SNR – SSIM for each case is put

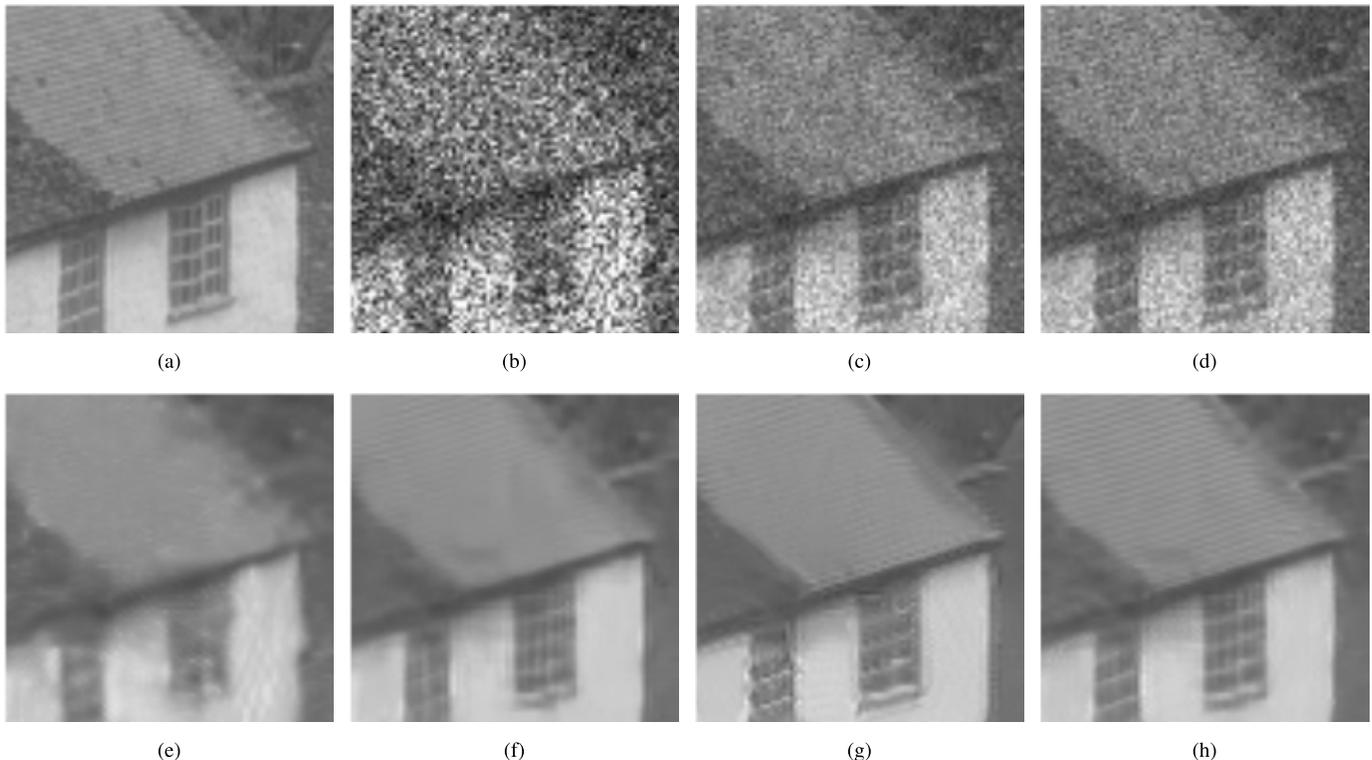


Fig. 5. Despeckling of the image “Hill” (size  $512 \times 512 \times 8$ ) degraded with a single-look simulated speckle. (a) Hill. (b) Noisy with  $L = 1$ . (c) UTA (SNR: 5.83 dB). (d) NLTF (SNR: 6.08 dB). (e) SAR-BM3D (SNR: 12.20 dB). (f) Log-BM4D (SNR: 14.67 dB). (g) 2S-PPB (SNR: 14.90 dB). (h) MSAR-BM3D (SNR: 15.81 dB).

TABLE I

SNR–SSIM INDEXES ARISING FROM THE DESPECKLING OF SYNTHETIC SAR IMAGES WITH EIGHT TEMPORAL COMPONENTS

Looks ( $L$ )	UTA [19]	NLTF	SAR-BM3D [1] ( $\lambda_{3D} = 2.7$ )	log-BM4D [36] ( $\lambda_{4D} = 3.0$ )	2S-PPB [31]	MSAR-BM3D ( $\lambda_{4D} = 3.5$ )
HOUSE						
1	03.50 – 0.314	03.72 – 0.321	12.07 – 0.753	15.11 – 0.837	15.13 – 0.816	<b>16.67 – 0.852</b>
4	10.19 – 0.539	10.46 – 0.549	16.07 – 0.832	19.78 – 0.889	18.40 – 0.872	<b>20.21 – 0.902</b>
8	13.35 – 0.657	13.64 – 0.665	17.93 – 0.861	21.79 – 0.924	19.64 – 0.902	<b>21.95 – 0.928</b>
NAPOLI						
1	05.30 – 0.582	05.48 – 0.589	7.19 – 0.662	10.32 – 0.790	10.11 – 0.804	<b>11.28 – 0.830</b>
4	11.93 – 0.808	12.20 – 0.814	10.25 – 0.800	15.42 – 0.920	12.38 – 0.876	<b>15.58 – 0.922</b>
8	15.09 – 0.885	15.34 – 0.888	11.94 – 0.858	17.88 – 0.952	13.28 – 0.904	<b>17.94 – 0.952</b>
HILL						
1	05.83 – 0.469	06.08 – 0.480	12.20 – 0.650	14.67 – 0.764	14.90 – 0.765	<b>15.81 – 0.797</b>
4	12.42 – 0.711	12.71 – 0.718	15.15 – 0.773	18.97 – 0.884	17.25 – 0.848	<b>19.11 – 0.882</b>
8	15.63 – 0.807	15.89 – 0.812	16.60 – 0.825	20.87 – 0.920	18.53 – 0.884	<b>20.88 – 0.920</b>

in boldface for the sake of clarity. On the left side of the table, one can see that the proposed NLTF performs slightly better than UTA, thanks to the statistics estimated using the nonlocal procedure described in Section II-B. On the right side of the table, we can see that MSAR-BM3D provides consistently the best performance, gaining up to 1.5 dB with respect to log-BM4D, which looks as the second best. The general trend of the SNR is quite similar with respect to the number  $L$  of looks, except that log-BM4D tends to close the gap with MSAR-BM3D as  $L$  grows.

The behavior of the SNR is reflected in the zoom of the denoised images illustrated in Fig. 5 for  $L = 1$ . It is clear that strong noise reduction comes at the price of some loss of details. MSAR-BM3D and 2S-PPB seem to offer the best compromise between these contrasting needs, but the latter

also introduces some artifacts that degrade the image quality.

Beyond this experiment, we also used the benchmark framework proposed in [42], which is based on a physical-level SAR image simulation. We considered five canonical scenes (Homogeneous, Digital Elevation Model (DEM), Squares, Corner, and Building) with related measures to assess the capability of the different algorithms to smooth homogeneous areas, preserve textures, edges, and isolated targets. For a detailed description of the different measures refer to [42]. From Table II, it appears evident that 2S-PPB and MSAR-BM3D give the best performance, which are significantly superior than those of other algorithms. For example, the two bias indicators for the Homogeneous scene, namely, the mean value of the filtered image (MoI) and the mean value of the ratio image (MoR), show that

TABLE II  
PERFORMANCE ON THE BENCHMARK FRAMEWORK PROPOSED IN [42]

	SAR	clean	UTA	NLTF	SAR-BM3D	log-BM4D	2S-PPB	MSAR-BM3D
HOMOGENEOUS								
MoI	0.998	1.000	0.815	0.812	0.978	1.036	0.997	<b>1.000</b>
MoR	-	0.998	1.238	1.239	0.979	0.937	0.985	<b>0.990</b>
ENL	1.00	436.97	6.509	7.390	102.44	174.83	<b>621.81</b>	428.07
DG	0	-	8.655	9.050	19.40	20.45	<b>25.09</b>	24.27
DEM								
MoI	1.003	1.000	0.817	0.882	0.968	0.968	<b>1.003</b>	0.997
MoR	-	1.001	1.236	1.162	0.833	0.908	<b>0.953</b>	0.950
$C_x$	3.54	2.40	2.63	3.07	2.43	2.51	<b>2.40</b>	2.45
DG	0	-	8.27	3.60	5.32	7.95	8.47	<b>10.25</b>
SQUARES								
ES (up)	0.010	-	0.104	0.106	0.036	0.006	0.003	<b>0.001</b>
ES (down)	0.029	-	0.300	0.308	0.113	0.024	0.011	<b>0.005</b>
FOM	0.792	0.993	0.821	0.937	0.847	0.904	<b>0.959</b>	0.953
CORNER								
$C_{NN}$	7.77	7.75	<b>7.77</b>	<b>7.77</b>	7.39	7.68	7.46	7.24
$C_{BG}$	36.50	36.56	37.39	37.41	35.45	38.21	33.61	<b>36.19</b>
BUILDING								
$C_{DR}$	65.90	65.90	66.79	66.81	<b>65.91</b>	68.24	64.93	66.17
BS	0.09	-	<b>0.06</b>	0.07	1.46	8.11	3.15	3.26

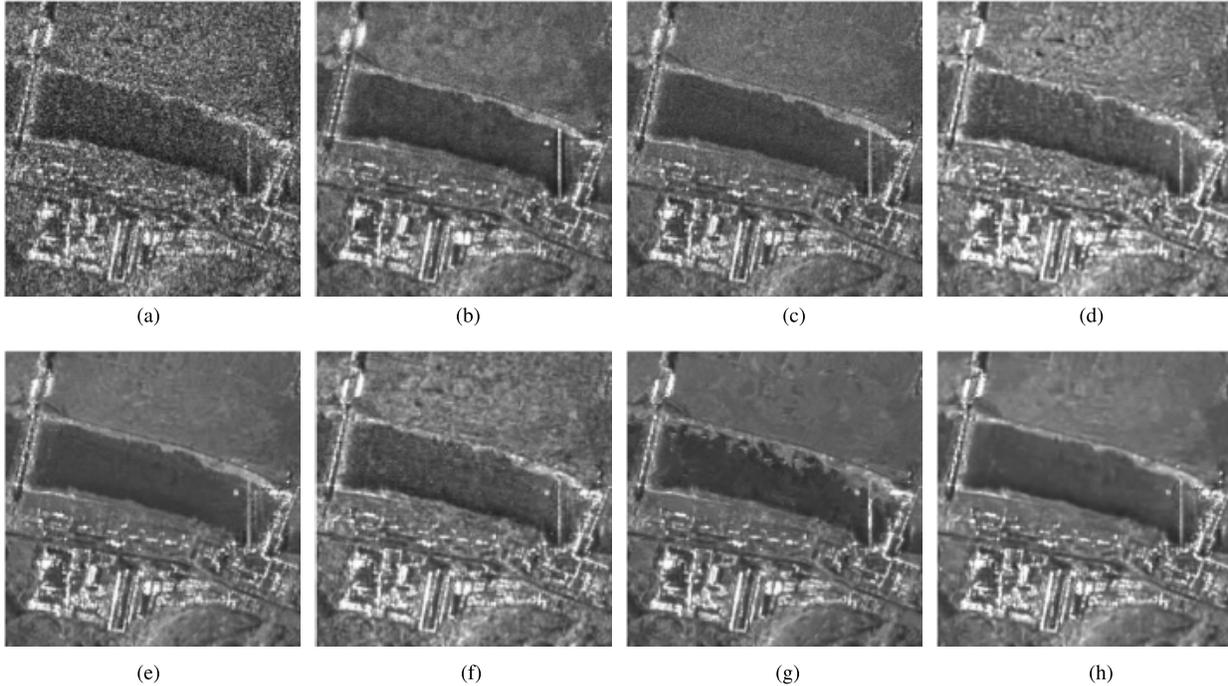


Fig. 6. Despeckling of a single-look COSMO-SkyMed image with eight temporal components (only a zoom of the first component is shown). (a) COSMO-SkyMed. (b) UTA. (c) NLTF. (d) SAR-BM3D. (e) NLTF + SAR-BM3D. (f) Log-BM4D. (g) 2S-PPB. (h) MSAR-BM3D.

MSAR-BM3D is able to preserve radiometric precision better than other methods. In addition, direct measures of performance, like the despeckling gain (DG) and ENL, are also very high, comparable with 2S-PPB that achieves a very high smoothing capability.

For what concerns the texture preservation of DEM, again both 2S-PPB and MSAR-BM3D are by far the best approaches. Furthermore, MSAR-BM3D shows a significant improvement in terms of DG with respect to other methods. Edge preservation measures are also in favor of 2S-PPB and MSAR-BM3D. Specifically, edge smearing (ES) that accounts for edge profile degradation is very low if compared with

UTA and NLTF. Also, the figure of merit (FOM) achieves very high values for the proposed method. Finally, for what concerns isolated target and building preservation, performance measured by means of contrast indicators is good. This confirms that the proposed method is able to guarantee a good despeckling capability without compromising the behavior on corner reflectors and isolated buildings.

#### D. State-of-the-Art Comparison on Multitemporal SAR Images

In our last set of experiments, we consider a single-look COSMO-SkyMed image with eight temporal compo-

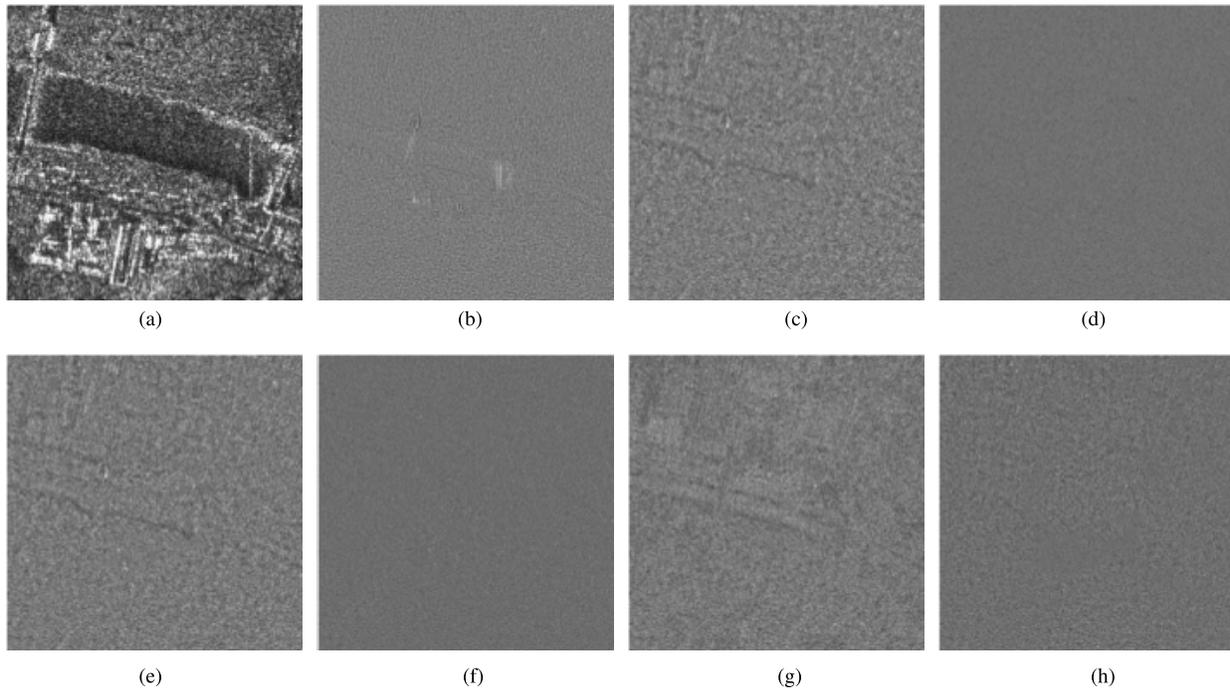


Fig. 7. Ratio images for the results in Fig. 6. (a) COSMO-SkyMed. (b) UTA. (c) NLTF. (d) SAR-BM3D. (e) NLTF + SAR-BM3D. (f) Log-BM4D. (g) 2S-PPB. (h) MSAR-BM3D.

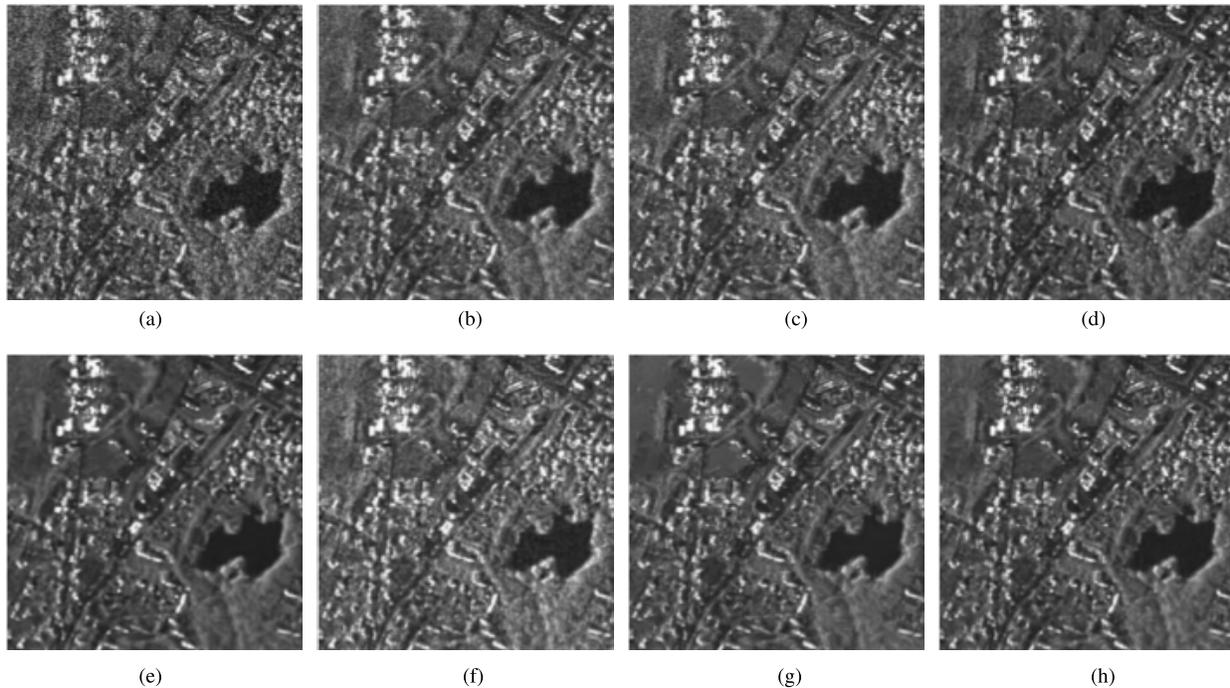


Fig. 8. Despeckling of a single-look TerraSAR-X image with four temporal components (only a zoom of the first component is shown). (a) TerraSAR-X. (b) UTA. (c) NLTF. (d) SAR-BM3D. (e) NLTF + SAR-BM3D. (f) Log-BM4D. (g) 2S-PPB. (h) MSAR-BM3D.

ments, a single-look TerraSAR-X image with four temporal components,<sup>5</sup> and a single-look Sentinel image with six temporal components.<sup>6</sup>

Fig. 6 shows a section drawn from the COSMO-SkyMed image and covering heterogeneous sceneries: urban areas, fields, woods, and water. Despeckling results are quite

consistent, indicating MSAR-BM3D by far as the technique with the strongest speckle rejection ability, followed by 2S-PPB, log-BM4D, and the other temporal filters. This is immediately obvious by visual inspecting the results. Fig. 6 also reports the result obtained using NLTF as a preprocessing step of the standard SAR-BM3D. Not surprisingly, the image despeckled with NLTF + SAR-BM3D presents a similar visual quality as MSAR-BM3D, since they involve the same operations, such as temporal filtering, block-matching, and

<sup>5</sup>Information available at [https://saredu.dlr.de/data/tsx\\_ruhr](https://saredu.dlr.de/data/tsx_ruhr)

<sup>6</sup>Information available at <https://scihub.copernicus.eu/>

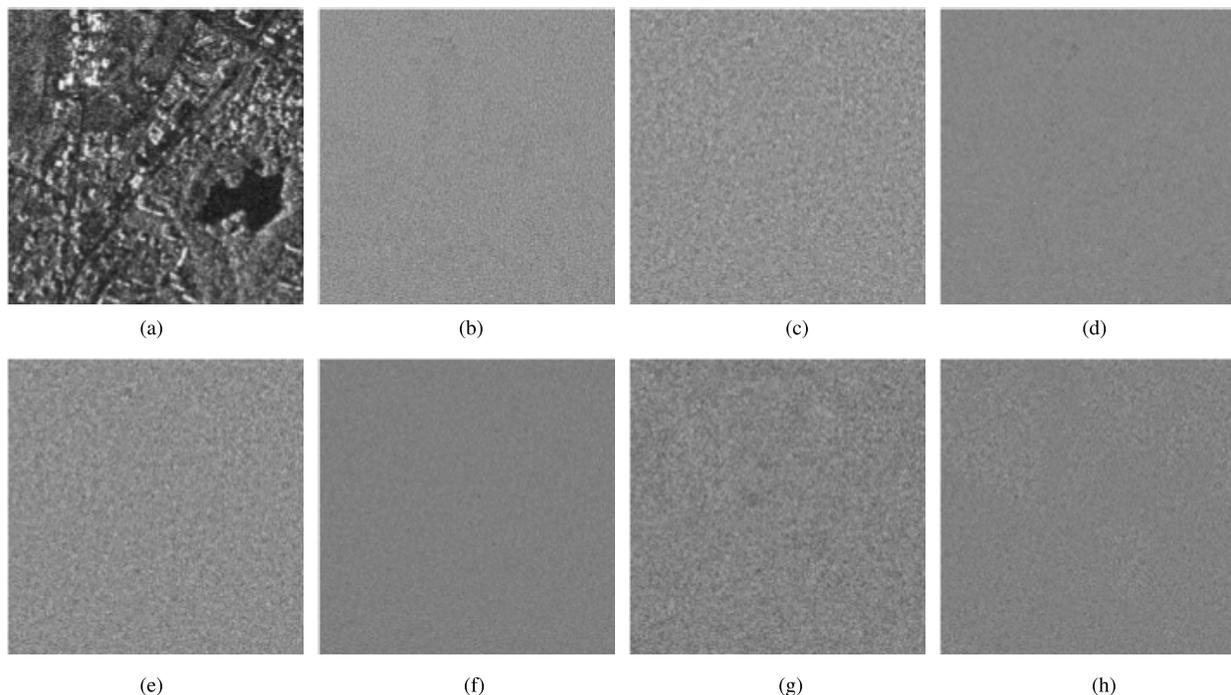


Fig. 9. Ratio images for the results in Fig. 8. (a) TerraSAR-X. (b) UTA. (c) NLTF. (d) SAR-BM3D. (e) NLTF + SAR-BM3D. (f) Log-BM4D. (g) 2S-PPB. (h) MSAR-BM3D.

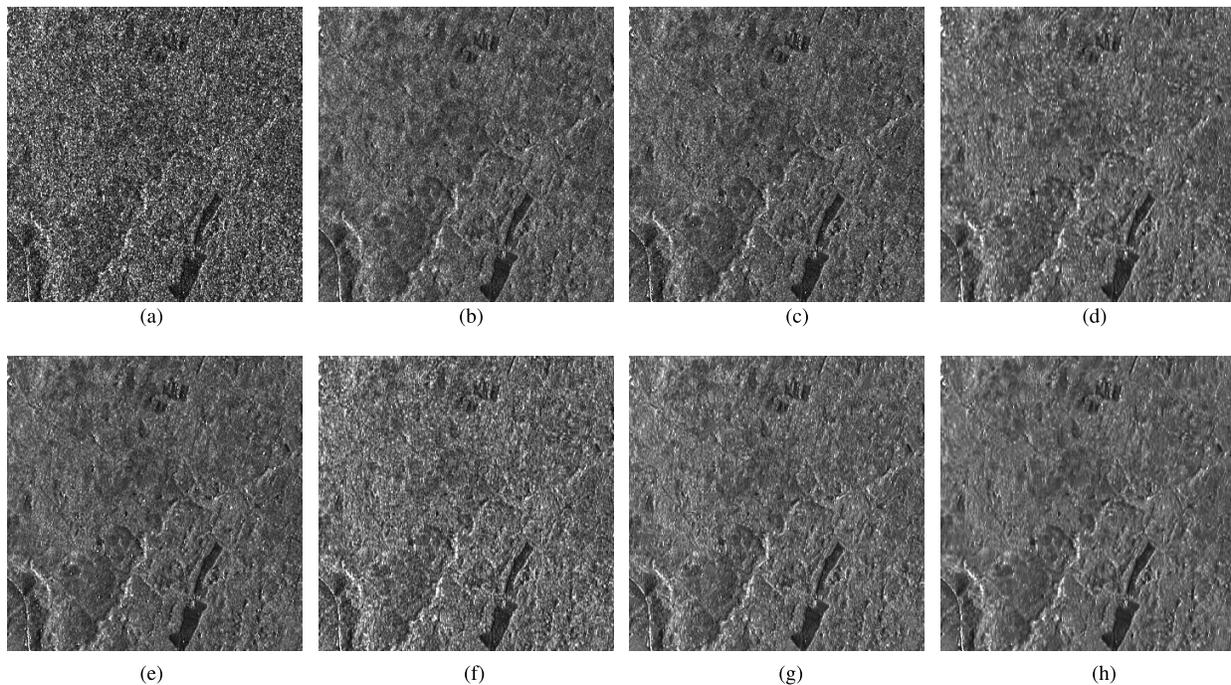


Fig. 10. Despeckling of a single-look Sentinel image with eight temporal components (only a zoom of the first component is shown). (a) Sentinel. (b) UTA. (c) NLTF. (d) SAR-BM3D. (e) NLTF + SAR-BM3D. (f) Log-BM4D. (g) 2S-PPB. (h) MSAR-BM3D.

collaborative filtering. Although the 2S-PPB image may look more pleasant than the other ones, it presents widespread artifacts resembling watercolor strokes. In this regard, MSAR-BM3D seems to be the only technique that guarantees a significant noise reduction without introducing some kinds of artifacts. However, without a noiseless reference, one cannot decide whether such artifacts imply a loss of details.

Fig. 8 gives a similar example for a section drawn from the TerraSAR-X image. The restoration of geometric struc-

tures and details proves the effectiveness of MSAR-BM3D, even though the visual quality of the image despeckled with the latter is similar to the one produced by 2S-PPB. It is worth noting that log-BM4D performs sensibly worse than MSAR-BM3D, NLTF+SAR-BM3D, and 2S-PPB on real SAR images. This is not surprising since the homomorphic approach is well known to be less effective on single-look SAR images, as the speckle is not fully developed for high-resolution images. Remarkably, this behavior was well predicted by the

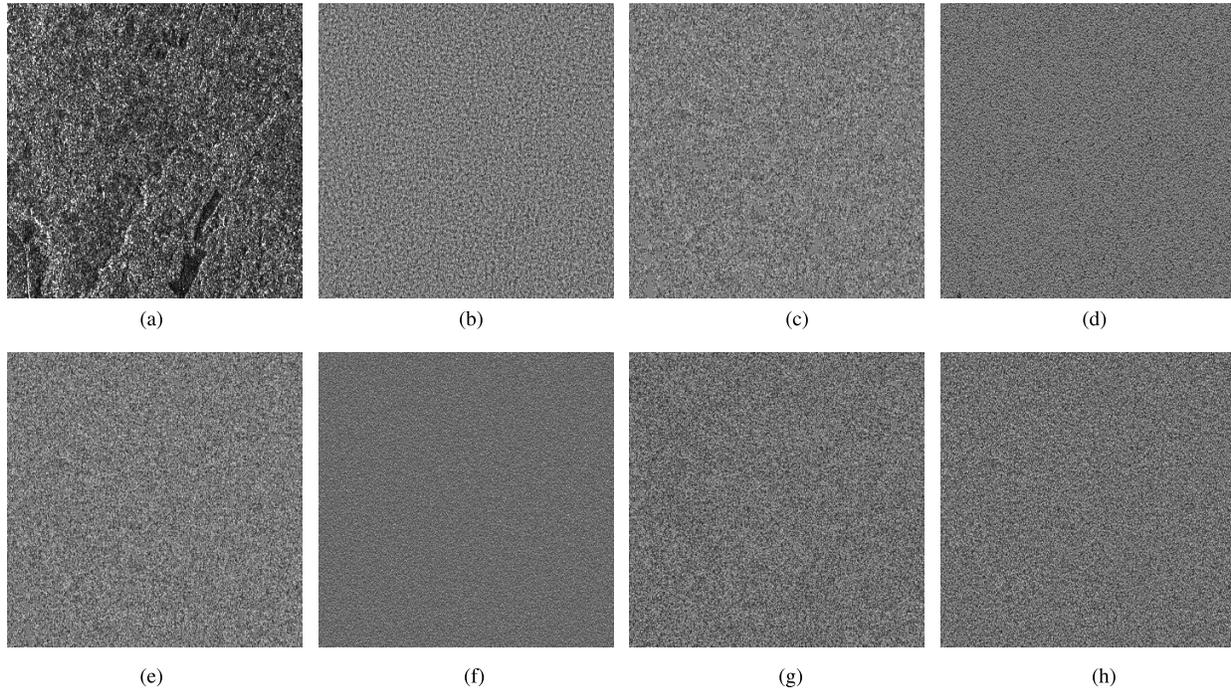


Fig. 11. Ratio images for the results in Fig. 10. (a) Sentinel. (b) UTA. (c) NLTF. (d) SAR-BM3D. (e) NLTF + SAR-BM3D. (f) Log-BM4D. (g) 2S-PPB. (h) MSAR-BM3D.

TABLE III  
PERFORMANCE (ENL AND  $\alpha\beta$  INDEXES) OF THE DESPECKLED SAR IMAGES IN FIGS. 6, 8, AND 10

	SAR	UTA	NLTF	SAR-BM3D	NLTF+SAR-BM3D	log-BM4D	2S-PPB	MSAR-BM3D
COSMO-SKYMED								
ENL (area 1)	1.00	4.64	5.45	6.38	16.71	4.94	<b>45.27</b>	36.12
ENL (area 2)	0.99	4.13	4.68	5.16	12.67	1.93	<b>22.07</b>	18.70
$\alpha\beta$	-	0.44	0.52	0.68	0.64	0.45	0.62	<b>0.40</b>
TERRASAR-X								
ENL (area 1)	1.02	2.57	2.85	3.78	7.73	3.29	<b>23.69</b>	14.05
ENL (area 2)	0.99	2.73	2.81	4.22	7.68	4.31	<b>15.17</b>	15.02
$\alpha\beta$	-	0.21	0.17	0.32	0.16	0.23	0.17	<b>0.12</b>
SENTINEL								
ENL (area 1)	1.02	2.97	2.88	7.87	4.85	3.37	4.62	<b>12.93</b>
ENL (area 2)	1.01	3.13	3.00	6.11	4.72	3.23	5.46	<b>11.69</b>
$\alpha\beta$	-	0.67	0.88	0.71	0.64	1.17	0.47	<b>0.33</b>

TABLE IV  
EXECUTION TIME (IN SECONDS) FOR DESPECKLING THE  $1000 \times 1000 \times 8$  SECTION OF THE COSMO-SKYMED IMAGE SHOWN IN FIG. 6

UTA	NLTF	SAR-BM3D	NLTF+SAR-BM3D	log-BM4D	2S-PPB	MSAR-BM3D
0.5 s	7.5 s	1408 s	1415 s	450 s	717 s	112 s

experiments on simulated data presented in Table I. Finally, Fig. 10 shows a section drawn from the Sentinel image, representing the forest near Dresden. On this distributed scene, it is possible to observe that the proposed method is also able to preserve texture details.

As objective no-reference metrics, we consider the well-known ENL, and the recently proposed  $\alpha\beta$  index [40]

$$\alpha\beta = \{\alpha + |\delta_{\text{ENL}}| + (1 - \alpha)|\delta_{\mu}\} + \beta_{\text{ratio}} \quad (23)$$

where  $\alpha \in [0, 1]$ ,  $\delta_{\text{ENL}} = \text{ENL}_{\text{noisy}} - \text{ENL}_{\text{ratio}}$ , and  $\delta_{\mu}$  is the residue of the mean value of the speckle  $1 - \mu_{\text{ratio}}$ . All these measures need to be estimated in homogeneous areas in the ratio and noisy image. While ENL accounts only for the speckle suppression capability, the  $\alpha\beta$  index can

also measure the amount of detail preservation. In fact, its definition is based on the ratio edge detector, and it is able to measure the remaining geometric content within the ratio image (ideally this index should be equal to zero). Numerical results are reported in Table III both for COSMO-SkyMed and TerraSAR-X images on two selected areas of about  $50 \times 50$  pixels. It can be seen that the higher ENL is achieved by 2S-PPB and MSAR-BM3D, but the latter presents lower  $\alpha\beta$  indexes, suggesting a better detail preservation capability. The simpler UTA and NLTF approaches also show low values for these indexes, but they are also characterized by a very low speckle suppression ability. MSAR-BM3D seems to be the best compromise between speckle suppression and detail preservation. As a further evidence

supporting this claim, Figs. 7, 9, and 11 show the ratio images, which look speckle-like only for MSAR-BM3D and log-BM4D. Some residual structures are instead visible in all the other methods, especially for the COSMO-SkyMed image.

Finally, Table IV reports the computational time of each technique measured on a Intel i5 CPU at 3.20 GHz and 12 GB of RAM (note that 2S-PPB was only made available in MATLAB). UTA and NLTF are obviously the fastest techniques, because of their simpler conception. Among the others, MSAR-BM3D exhibits the lowest complexity, which is four times faster than log-BM4D, and about ten times faster than SAR-BM3D itself, thanks to the lookup table distance computation. This time is probably already acceptable for filtering multitemporal SAR images, but can be much reduced with dedicated hardware.

## V. CONCLUSION

In this paper, we proposed a despeckling algorithm for multitemporal SAR images. Thanks to the nonlocal estimation of the statistics, it can easily handle changes occurring in the temporal direction and, at the same time, it is able to obtain a good despeckling capability in the spatial dimension. This can be achieved by means of a two-pass procedure that comprises grouping, collaborative filtering, and aggregation. Furthermore, the use of lookup tables for the block-matching helps to limit the increase in computational time due to the nonlocal approach. The experiments carried out both on simulated and on real SAR images show that the algorithm is able to obtain good performance with respect to state-of-the-art approaches. In particular, it is able to suppress speckle and to limit visible artifacts, while preserving fine structures and region boundaries.

## ACKNOWLEDGMENT

The authors would like to thank X. Su, C.-A. Deledalle, F. Tupin, and H. Sun for providing the 2S-PPB code. They would also like to thank G. Poggi for his helpful comments.

## REFERENCES

- [1] 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.
- [2] L. Bruzzone, M. Marconcini, U. Wegmuller, and A. Wiesmann, "An advanced system for the automatic classification of multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1321–1334, Jun. 2004.
- [3] Y. Maghsoudi, M. J. Collins, and D. Leckie, "On the use of feature selection for classifying multitemporal Radarsat-1 images for forest mapping," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 904–908, Sep. 2011.
- [4] R. Gaetano *et al.*, "Exploration of multitemporal COSMO-SkyMed data via interactive tree-structured MRF segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 7, pp. 2763–2775, Jul. 2014.
- [5] G. Moser and S. B. Serpico, "Unsupervised change detection from multichannel SAR data by Markovian data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2114–2128, Jul. 2009.
- [6] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and C. Zoppetti, "Nonparametric change detection in multitemporal SAR images based on mean-shift clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2022–2031, Apr. 2013.
- [7] X. Su, C.-A. Deledalle, F. Tupin, and H. Sun, "NORCAMA: Change analysis in SAR time series by likelihood ratio change matrix clustering," *ISPRS J. Photogram. Remote Sens.*, vol. 101, pp. 247–261, Sep. 2015.
- [8] F. Argenti, A. Lapini, L. Alparone, and T. Bianchi, "A tutorial on speckle reduction in synthetic aperture radar images," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 3, pp. 6–35, Apr. 2013.
- [9] 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, Feb. 1980.
- [10] 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.
- [11] 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, Feb. 1985.
- [12] 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, 1993.
- [13] J. Nicolas, F. Tupin, and H. Maitre, "Smoothing speckled SAR images by using maximum homogeneous region filters: An improved approach," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Sep. 2001, pp. 1503–1505.
- [14] G. Vasile, E. Trouvé, J.-S. Lee, and V. Buzuloiu, "Intensity-driven adaptive-neighborhood technique for polarimetric and interferometric SAR parameters estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1609–1621, Jun. 2006.
- [15] M. Ciuc, P. Bolon, E. Trouvé, V. Buzuloiu, and J. P. Rudant, "Adaptive-neighborhood speckle removal in multitemporal synthetic aperture radar images," *Appl. Opt.*, vol. 40, no. 32, pp. 5954–5966, 2001.
- [16] G. F. De Grandi, M. Leysen, J. S. Lee, and D. Schuler, "Radar reflectivity estimation using multiple SAR scenes of the same target: Technique and applications," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Aug. 1997, pp. 1047–1050.
- [17] J.-S. Lee, "Refined filtering of image noise using local statistics," *Comput. Graph. Image Process.*, vol. 15, no. 4, pp. 380–389, 1981.
- [18] J.-J. Lee, M. R. Grunes, and S. A. Mango, "Speckle reduction in multipolarization, multifrequency SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, no. 4, pp. 535–544, Jul. 1991.
- [19] S. Quegan, T. L. Toan, J. J. Yu, F. Ribbes, and N. Floury, "Multitemporal ERS SAR analysis applied to forest mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 741–753, Mar. 2000.
- [20] S. Quegan and J. J. Yu, "Filtering of multichannel SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2373–2379, Nov. 2001.
- [21] T. T. Le, A. M. Atto, E. Trouvé, and J. Nicolas, "Adaptive multitemporal SAR image filtering based on the change detection matrix," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 10, pp. 1826–1830, Oct. 2014.
- [22] T. T. Le, A. M. Atto, E. Trouvé, A. Solikhin, and V. Pinel, "Change detection matrix for multitemporal filtering and change analysis of SAR and PolSAR image time series," *ISPRS J. Photogram. Remote Sens.*, vol. 107, pp. 64–76, Sep. 2015.
- [23] C.-A. Deledalle, L. Denis, G. Poggi, F. Tupin, and L. Verdoliva, "Exploiting patch similarity for SAR image processing: The nonlocal paradigm," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 69–78, Apr. 2014.
- [24] 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.
- [25] 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.
- [26] J. Chen, Y. Chen, W. An, Y. Cui, and J. Yang, "Nonlocal filtering for polarimetric SAR data: A pretest approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1744–1754, May 2011.
- [27] C. A. Deledalle, L. Denis, and F. Tupin, "NL-InSAR: Nonlocal interferogram estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1441–1452, Apr. 2011.
- [28] C.-A. Deledalle, L. Denis, F. Tupin, A. Reigber, and M. Jäger, "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.
- [29] P. Sica, D. Reale, L. Verdoliva, G. Poggi, and G. Fornaro, "Nonlocal adaptive multilooking in SAR multipass differential interferometric processing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 4, pp. 1727–1742, Apr. 2015.

- [30] C. A. Deledalle, L. Denis, and F. Tupin, "How to compare noisy patches? Patch similarity beyond Gaussian noise," *Int. J. Comput. Vis.*, vol. 99, no. 1, pp. 86–102, Jan. 2012.
- [31] X. Su, C. Deledalle, F. Tupin, and H. Sun, "Two-step multitemporal nonlocal means for synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6181–6196, Nov. 2014.
- [32] M. Mahmoudi and G. Sapiro, "Fast image and video denoising via nonlocal means of similar neighborhoods," *IEEE Signal Process. Lett.*, vol. 12, no. 12, pp. 839–842, Dec. 2005.
- [33] P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, "An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 27, no. 4, pp. 425–441, Apr. 2008.
- [34] C. A. Deledalle, V. Duval, and J. Salmon, "Non-local methods with shape-adaptive patches (NLM-SAP)," *J. Math. Imag. Vis.*, vol. 43, no. 2, pp. 103–120, Jun. 2012.
- [35] C. Sutour, C.-A. Deledalle, and J.-F. Aujol, "Adaptive regularization of the NL-means: Application to image and video denoising," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3506–3521, Aug. 2014.
- [36] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 119–133, Apr. 2013.
- [37] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2. San Diego, CA, USA, Jun. 2005, pp. 60–65.
- [38] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [39] S. Ghael, A. Sayeed, and R. Baraniuk, "Improved wavelet denoising via empirical Wiener filtering," *Proc. SPIE*, vol. 3169, pp. 389–399, Oct. 1997.
- [40] L. Gomez, M. E. Buemi, J. C. Jacobo-Berlles, and M. E. Mejail, "A new image quality index for objectively evaluating despeckling filtering in SAR images," *J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 3, pp. 1291–1307, Apr. 2016.
- [41] 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.
- [42] 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.



**Giovanni Chierchia** (M'14) received the Engineering Degree in computer science from the University of Naples Federico II, Naples, Italy, in 2010, and the Ph.D. degree from Telecom ParisTech, Paris, France, in 2015.

From 2010 to 2011, he was a Research Engineer with Telecom ParisTech. In 2015, he was a Post-Doctoral Researcher with the University of Naples Federico II. Since 2015, he has been an Assistant Professor with ESIEE Paris, Université Paris Est, Noisy-le-Grand, France. His research interests include convex optimization, image restoration, and supervised learning.



**Mireille El Gheche** received the Engineering Degree in telecommunications from Lebanese University, Tripoli, Lebanon, the master's degree in radio-communication from SupElec, Paris, France, in 2010, and the Ph.D. degree from Université Paris Est, Marne-la-Vallée, France, in 2014.

Since 2015, she has been a Post-Doctoral Researcher with IMS and IMB Laboratories, Université de Bordeaux, Talence, France. Her research interests include convex and nonconvex optimization applied to inverse problems, such as image restoration, stereo matching, super-resolution, and horizon volume reconstruction from seismic data.



**Giuseppe Scarpa** (M'15) received the Laurea (M.S.) degree in telecommunication engineering and the Ph.D. degree in electronic and telecommunication engineering from the University of Naples Federico II, Naples, Italy, in 2001 and 2005, respectively.

He was a Visiting Student at the INRIA Institute, Sophie Antipolis, France. He was a Research Fellow with the UTIA Institute, Czech Academy of Sciences, Prague, in 2005, and with the INRIA Institute, in 2006. Since 2006, he has been an Assistant Professor with the Department of Electrical Engineering and Information Technology, University of Naples Federico II. His research interests include image analysis, and, in particular, segmentation, texture modeling and classification, object detection, pansharpening, and feature extraction, with applications in both remote sensing and medical domains.

Prof. Scarpa was a recipient of Marie Curie Scholarship in 2003 and a joint ERCIM Post-Doctoral Fellowship in 2004. He is currently a Senior Area Editor of the IEEE SIGNAL PROCESSING LETTERS.



**Luisa Verdoliva** (M'14) is currently an Assistant Professor of telecommunications with the Department of Electrical Engineering and Information Technology, University of Naples Federico II, Naples, Italy. Her research interests include image processing, in particular restoration of remote-sensing images, both optical and SAR, and digital image forensics.

Dr. Verdoliva has been an elected member of the IEEE Information Forensics and Security Technical Committee since 2016.