ASSESSING PERFORMANCE OF MULTITEMPORAL SAR IMAGE DESPECKLING FILTERS VIA A BENCHMARKING TOOL

Gerardo Di Martino, Alessio Di Simone, Antonio Iodice, Daniele Riccio, Giuseppe Ruello

University of Naples "Federico II", Naples, Italy

IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium | 978-1-7281-6374-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/IGARSS39084.2020.9323137

ABSTRACT

In this work we present a novel benchmarking framework for assessing the quality of despeckled multitemporal SAR images on an objective basis. Taking cues from a recent work providing a useful tool for measuring the performance of single-channel despeckling filters, here we extend that analysis to multitemporal filters and datasets. Due to the lack of reference speckle-free real SAR images, the quantitative performance parameters introduced here are evaluated on simulated time series of canonical scenes obtained through a reliable and well-assessed SAR simulator which accounts for both geometrical and electromagnetic properties of the scattering surface. First we analyze image quality over datasets with homogeneous reflectivity in time. Then, for a more realistic performance assessment in practical situations, ad-hoc quality parameters are introduced to measure the effects of time-varying scene characteristics on the despeckled dataset. The consistency of the proposed framework is tested on state-of-the-art multitemporal despeckling algorithms.

***Index Terms*—** Synthetic aperture radar (SAR), speckle, multitemporal filtering, SAR simulation

1. INTRODUCTION

In the last decade, the radar and remote sensing scientific community has benefited from the development of numerous satellite systems conceived for Earth Observation (EO) and remote sensing applications. This is especially true for synthetic aperture radar (SAR) thanks to the launch of different sensor constellations (COSMO-SkyMed, TerraSAR-X, RADARSAT, Sentinel-1), most of which are still operational. As a result, the end-users can now benefit from a huge amount of data acquired at different resolution (both in space and time), frequency, look angle and polarization that can be fruitfully exploited in advanced image processing techniques, e.g., machine learning. Among others, the Copernicus Sentinel-1 mission represented a breakpoint in multitemporal analysis of SAR imagery thanks to the large catalogue of data accessible at no charge. SAR multitemporal series acquired on the same scene and spanning over years are nowadays available and allowed users to gather more accurate and comprehensive information in both spatial and temporal domains. As a matter of fact, numerous applications using SAR time series have been investigated so far, e.g., information retrieval [1], change detection [2] and classification [3, 4].

As for single-channel images, speckle noise represents the main disturbance in multitemporal SAR data, thus significantly impairing the interpretation by human users and the processing by automatic algorithms. To mitigate the degrading effects of speckle on SAR time series, intense research activity has been carried out in the recent past [5–8].

Despite the key relevance of reliable despeckling algorithms, very little effort has been focused on the development of standard procedures for an objective and comprehensive assessment of the filter performance in a wide variety of contexts. Actually, such a topic has been investigated for single-channel despeckling algorithms in [9], where different quantitative measures were introduced to assess image quality in several respects, such as edge smoothing, speckle reduction, texture and target features preservation.

In this paper, we extend that analysis to multi-band SAR imagery, with the aim of providing a framework for benchmarking competing multitemporal despeckling techniques, by defining a set of canonical scenes and suitable performance measures in analogy with [9].

The remainder of this work is as follows: Section 2 introduces the proposed benchmarking framework and the set of performance parameters to be used in the different canonical scenes. Some experimental results obtained with state-of-the-art multitemporal filters are shown and discussed in Section 3. Finally, main conclusions are briefly sketched in Section 4.

2. PROPOSED FRAMEWORK

A multitemporal image stack is composed by N SAR images (hereafter referred to as "bands") acquired at N different time instants over the same scene. Assuming fully developed speckle, each band z_i of the noisy stack $\mathbf{z}$ can be expressed as

$$z_i(s) = x_i(s)n_i(s) \tag{1}$$

where $x_i(s)$ and $n_i(s)$ stand for the noise-free reflectance map, i.e., the reference image, and the speckle component.

Due to the lack of reference speckle-free real-world SAR images, the benchmarking framework proposed here relies on

the well-assessed SAR raw signal simulator described in [10], that provides realistic SAR images accounting for both geometrical and electromagnetic properties of the illuminated surface - macroscopic (i.e., much larger than the operating wavelength) and microscopic (on the order of the wavelength or below) roughness, medium complex dielectric constant - along with system parameters - frequency, polarization, look angle, spatial resolution.

Here we analyze the filter behavior in presence of both a constant (in time) and time-varying reflectance. Accordingly, we define two different canonical situations, which are described in the following sections.

In addition, we assume here uncorrelated speckle noise along the bands.

2.1. Unperturbed case

In this first test case, we propose a method to evaluate filter performance in presence of a scene with constant reflectance x along the bands. In other words, the single-look image stack consists of independent and identically-distributed 2-dimensional processes. Despite this ideal behavior, such a scenario approximates a scattering surface whose reflectance varies on time scales much larger than the overall acquisition time of the image stack, i.e., the time interval between the first and the last image acquisition times. For instance, this may be the case for bare soils and man-made structures.

When applied on a stationary (in time) image stack, it is reasonable that a proper despeckling procedure fulfills the following conditions:

- For N approaching infinity, the filtered dataset $\widehat{\mathbf{x}}_N$ converges (within a determined convergence mode) to the reference image x (which is the same among all the bands), i.e.:
$$\lim_{N\to\infty} d(\widehat{\mathbf{x}}_N, x) = 0 \tag{2}$$
where $d(\cdot,\cdot)$ denotes the selected convergence metric or distance.
- The filter should converge rapidly in terms of number of bands. Accordingly, for a fixed threshold α_{TH}, we here introduce a *convergence rate* (CR) defined as the lowest number of bands $N_{\alpha_{TH}}$ such that:
$$\frac{d(\widehat{\mathbf{x}}_N, x) - d(\widehat{\mathbf{x}}_{N-1}, x)}{d(\widehat{\mathbf{x}}_{N-1}, x)} \leq \alpha_{TH} \tag{3}$$

Different convergence modes and metrics may be defined also depending on the features of the scene to be analyzed. In this work we adopt the classical mean squared convergence due to its ability to capture the overall despeckling capabilities of the filter regardless of the illuminated scene. Accordingly, the distance d between the filtered image stack and the reference image is evaluated as the overall mean-square error (MSE) over the N bands (MSE_N), i.e., as

$$\text{MSE}_N(\widehat{\mathbf{x}}_N, x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}[(\widehat{x}_i - x)^2] \tag{4}$$

where $\widehat{x}_i$ is the i-th filtered band and $\mathbb{E}[\cdot]$ stands for the expectation operator.

2.2. Perturbed case

This test case is aimed at providing proper performance indicators when the image stack exhibits appreciable variations of the scene reflectivity during the overall observation time,

Due to the large number of parameters influencing the SAR representation of the illuminated scene, it is easily understood that a scene reflectivity might exhibit any arbitrary SAR response. Therefore, a proper modeling and simulation of all possible temporal heterogeneities is meaningless. Taking the cue from the work in [9], here we define a canonical temporal variation of the scene reflectivity which allows us to define proper and objective performance metrics and, on the other hand, has physical motivations.

The perturbed dataset is then created as follows: the first $N-1$ images have been simulated according to the "homogeneous" case described in [9]. Therefore, they represent independent SAR images of a flat surface with homogeneous (in both time and space) electromagnetic and geometrical parameters. Hereafter, they are referred to as *original* bands. The temporal variation of the scene reflectivity is accounted for by introducing a corner reflector in the scattering area. Accordingly, the last image of the stack is simulated as the "corner" case of [9]. This is hereafter referred to as the *perturbed* band. Such a dataset may be representative of bright point targets appearing and disappearing during the observation interval, e.g., cars.

The presence of a bright feature in the perturbed band may affect despeckling capabilities in the original bands as multitemporal filters average along the time axis as well. Therefore, in such a scenario, it is reasonable to require that the multitemporal despeckling filter exhibits a sensitivity to the perturbation as low as possible in the original bands, while it should retain as accurate as possible the reflectivity variation features in the perturbed band.

Accordingly, proper metrics have to be defined for both the perturbed and original bands. The capability of the filter to keep the corner features in the perturbed band is here evaluated with the corner measures C_{NN} and C_{BG} defined according to [9] as:

$$C_{NN} = 10\log_{10}\frac{x_{CF}}{x_{NN}}, \quad C_{BG} = 10\log_{10}\frac{x_{CF}}{x_{BG}} \tag{5}$$

where x_{CF} is the intensity measured in the corner reflector site, x_{NN} is the average intensity computed in the eight-connected nearest neighbors, and x_{BG} is the background average intensity.

In order to assess despeckling capabilities in the original bands, we here propose a *perturbation sensitivity* (PS) measure on the i-th original band, which is defined as:

$$\mathrm{PS}(\widehat{x}_i, x_i) = 10\log_{10}(\mathrm{MSE}(\widehat{x}_i, x_i)/\mathrm{MSE}(\widehat{x}_{i,nopt}, x_i)) \quad (6)$$

where $\widehat{x}_{i,nopt}$ is the i-th filtered band of the corresponding unperturbed dataset. Finally, the PS for the overall despeckled image stack is defined by averaging (6), i.e.,

$$\mathrm{PS}(\widehat{\mathbf{x}}_N, \mathbf{x}) = \frac{1}{N-1}\sum_{i=1}^{N-1}\mathrm{PS}(\widehat{x}_i, x_i) \quad (7)$$

3. EXPERIMENTAL RESULTS

In this Section, we discuss some details related to the practical implementation of the proposed benchmarking framework and test it on two despeckling algorithms, namely the multitemporal SAR-BM3D (MSAR-BM3D) and RABASAR, which were first presented in [7] and [8], respectively. The choice of such filters has been driven by their state-of-the-art despeckling capabilities and their public availability.

For both unperturbed and perturbed cases, the reference image has been obtained according to the procedure described in [9], i.e., through pure multilook over 512 independent single-look images.

Furthermore, the single-look image stack relevant to the unperturbed scenario has been created according to the "squares" case of [9]. Accordingly, the scattering surface is a flat area whose electromagnetic parameters vary over four different squares.

In addition, in all experiments, default parameters as defined in the original works are used for both filters.

In the unperturbed case, the limit in (2) is approximated assuming $N = 64$, whereas $\alpha_{TH} = 0.1$ is used for the evaluation of CR. Furthermore, N is set equal to 8 in the perturbed case. Finally, all the images have been normalized to the mean intensity value of the reference.

Image results relevant to the unperturbed and perturbed scenarios are shown in Fig. 1 and Fig. 2, respectively. In particular, Fig. 1 shows the first band of the unperturbed dataset, whereas Fig. 2 shows the original band 1 (a-d) and the perturbed band 8 (e-h), in order to provide visual information about the capability of the filter to retain the temporal variation (i.e., the corner features) in the perturbed band and its influence on the other bands. Synthetic performance parameters are listed in Table 1 and Table 2.

In the unperturbed case, MSAR-BM3D has better convergence properties as it is evident from Fig. 1. Indeed, its MSE value is significantly lower than that obtained by the RABASAR filter, which also introduces visible artifacts, see Fig. 1(d). However, the latter converges much faster than the former as demonstrated by the CR measure.

Table 1. Measures for the unperturbed case

	MSE_{64}	CR
Reference	0	–
Noisy	0.2510	–
MSAR-BM3D	4.92×10^{-4}	32
RABASAR	16.61×10^{-4}	2

Table 2. Measures for the perturbed case

	C_{NN}	C_{BG}	PS
Reference	7.74	36.54	–
Noisy	7.71	36.68	0
MSAR-BM3D	7.62	37.26	0.55
RABASAR	7.66	36.30	12.06

In presence of a perturbation within the image stack, the RABASAR filter is able to preserve the perturbation features better than MSAR-BM3D as it is evident both visually in Fig. 2(e-h) and quantitatively in Table 2 (see C_{NN} and C_{BG} measures). However, the impact of the perturbation on the unperturbed bands is much stronger when using RABASAR than MSAR-BM3D, see PS measure in Table 2.

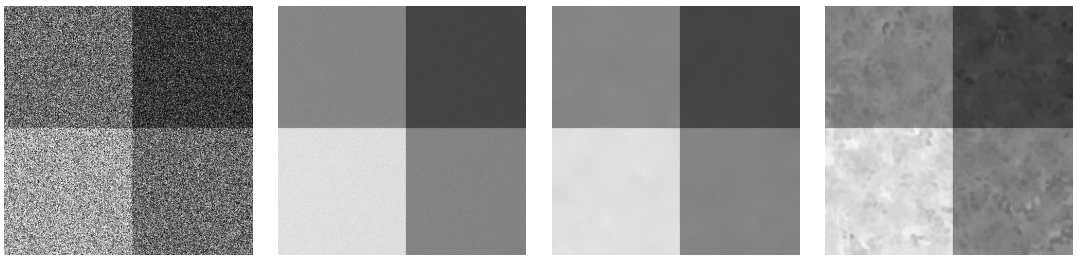
Finally, it is noteworthy that such results have been obtained assuming uncorrelated speckle noise along the bands and that filters performances might be influenced by the temporal correlation of the image stack.

4. CONCLUSIONS

In this work we have presented a new methodology for evaluating the despeckling performance of SAR multitemporal filters. Starting from a previous benchmarking framework for single-channel despeckling algorithms, we have defined two canonical datasets aimed at capturing filters behaviors in presence and absence of temporal changes in the scene reflectivity. Quantitative performance metrics have been defined according to reasonable requirements a multitemporal despeckling algorithm should have in both cases. Finally, the proposed framework has been run over simulated SAR images and applied on two state-of-the-art filters in order to illustrate its capability to catch the main filters properties. Future research lines may concern extensions of the proposed benchmarking framework to other canonical perturbations of the scene reflectivity, as well as the application of a wider set of despeckling techniques and the analysis of time-correlated bands.

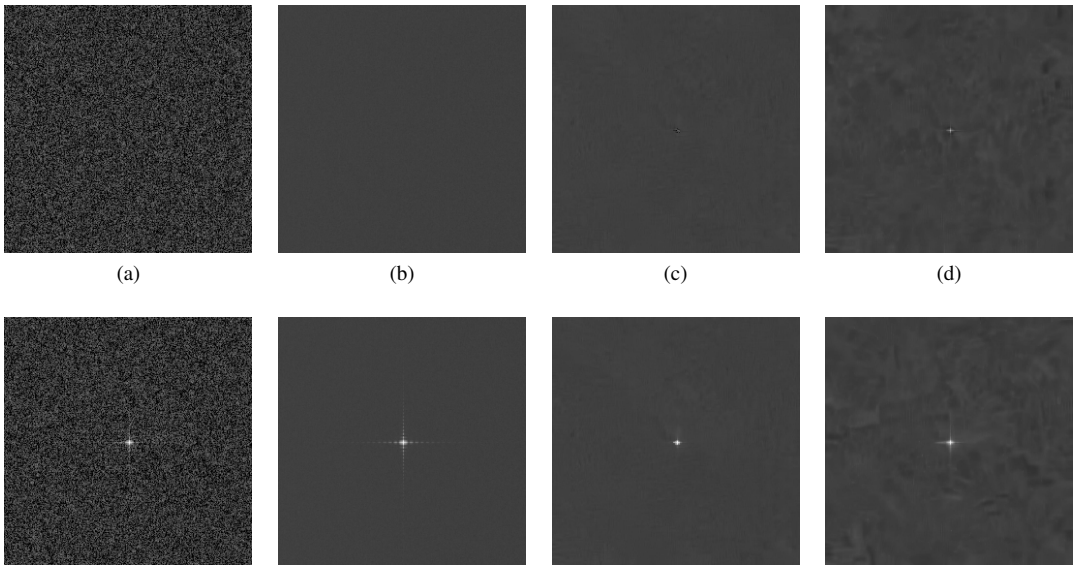
5. REFERENCES

[1] S. Paloscia *et al.*, "The contribution of multitemporal SAR data in assessing hydrological parameters," *IEEE*



(a) (b) (c) (d)

Fig. 1. Unperturbed case (band 1): (a) Single-look SAR image. (b) Reference. (c) MSAR-BM3D. (d) RABASAR.



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 2. Perturbed case, (a-d) original band (band 1), (e-h) perturbed band (band 8): (a,e) Single-look SAR image. (b,f) Reference. (c,g) MSAR-BM3D. (d,h) RABASAR.

Geosci. Remote Sens. Lett., vol. 1, no. 3, pp. 201–205, July 2004.

[2] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2963–2972, 2005.

[3] L. Bruzzone *et al.*, "An advanced system for the automatic classification of multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1321–1334, 2004.

[4] H. Skriver *et al.*, "Crop classification using short-revisit multitemporal SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 2, pp. 423–431, June 2011.

[5] S. Quegan and J. J. Yu, "Filtering of multichannel SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2373–2379, 2001.

[6] X. Su *et al.*, "Two-step multitemporal nonlocal means for synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6181–6196, 2014.

[7] G. Chierchia *et al.*, "Multitemporal SAR image despeckling based on block-matching and collaborative filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5467–5480, 2017.

[8] W. Zhao *et al.*, "Ratio-based multitemporal SAR images denoising: RABASAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3552–3565, 2019.

[9] G. Di Martino *et al.*, "Benchmarking framework for SAR despeckling," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1596–1615, 2014.

[10] G. Franceschetti *et al.*, "SARAS: A synthetic aperture radar (SAR) raw signal simulator," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 1, pp. 110–123, 1992.