Ratio-based multi-temporal SAR images denoising

Weiying Zhao, Loïc Denis, Charles-Alban Deledalle, Henri Maître, Jean-Marie Nicolas, Florence Tupin

Ratio-based multi-temporal SAR images denoising

Weiying Zhao, Loïc Denis, Charles-Alban Deledalle, Henri Maître, Jean-Marie Nicolas, Florence Tupin, *Senior Member, IEEE*

***Abstract*—In this paper, we propose a generic multi-temporal SAR despeckling method to extend any single-image speckle reduction algorithm to multi-temporal stacks. Our method, RAtio-BAsed multi-temporal SAR despeckling (RABASAR), is based on ratios and fully exploits a "super-image" (i.e. temporal mean) in the process. The proposed approach can be divided into three steps: 1) calculation of the "super-image" through temporal averaging; 2) denoising the ratio images formed through dividing the noisy images by the "super-image"; 3) computing denoised images by multiplying the denoised ratio images with the "super-image". Thanks to the spatial stationarity improvement in the ratio images, denoising these ratio images with a speckle-reduction method is more effective than denoising the original multi-temporal stack. The data volume to be processed is also reduced compared to other methods through the use of the "super-image". The comparison with several state-of-the-art reference methods shows numerically (peak signal-noise-ratio, structure similarity index) and visually better results both on simulated and real SAR stacks. The proposed ratio-based denoising framework successfully extends single-image SAR denoising methods in order to exploit the temporal information of a time series.**

***Index Terms*—Multi-temporal SAR series, ratio image, super-image, SAR, speckle reduction**

I. Introduction

SYNTHETIC aperture radar (SAR) imaging is widely used in monitoring land surfaces, disasters or the environment, due to its advantages of all-time acquisition capability, its sensitivity to geometric structures, its penetration characteristics, etc. However, the system-inherent speckle noise visually corrupts the appearance of images and severely diminishes the analysis and interpretation of SAR images. Therefore, a preliminary speckle reduction step is often necessary for the successful exploitation of SAR images.

The main problems of SAR image despeckling are spatial resolution preservation, edge and texture restoration, and point-like targets preservation. Spatial multilooking is a common way of reducing speckle fluctuations in a single SAR image, at the cost of a significant spatial resolution loss [1]. To effectively estimate the noise-free reflectivity as well as preserve the spatial resolution, many single-channel SAR speckle reduction methods have been proposed during the past decades. Detailed introductions of the methods are given by Touzi *et al.* [2], Argenti *et al.* [1] and Deledalle *et al.* [3]. These methods mainly belong to four categories: Bayesian methods in the spatial domain (Lee filter [4], Lee refined filter [5]), Bayesian methods in a transformed domain [6], selection-based filtering (IDAN [7], PPB [8] and NL-SAR [9]) and sparse-based approaches [10]. Most of the single-SAR-image denoising methods are designed through the combination of using different domain information (spatial domain or transformed domain), with different estimation criteria and different reflectivity or speckle probability distribution models [1]. Many state-of-the-art single-image despeckling methods perform a weighted average of some surrounding pixels values to estimate speckle-free values. This can induce a bias (a spatial resolution loss) when participating pixel candidates are not well selected or the corresponding weights are not well assigned. In addition, even using powerful spatial denoising approaches (such as SAR-BM3D [11] or NL-SAR [9]), the smallest and least contrasted structures can be damaged.

Recently, convolutional neural networks have shown a high capability of denoising data affected by additive white Gaussian noise (AWGN) [12]. Application to SAR images has also been proposed, either through homomorphic approach [13] or directly applying Gamma distribution based methods [14]. Unlike traditional SAR image denoising approaches, these methods predict the noise-free value through the estimation of the speckle component. These recent techniques reach comparable despeckling results w.r.t. state-of-the-art approaches, both in terms of signal to noise ratio (simulated data) and spatial feature preservation. However, the training of these networks is time consuming.

With the launch of new generations of satellite radar systems (Cosmo-SkyMed, TerraSAR-X, ALOS-2, Sentinel-1, etc.) [15], more and more SAR images, with shorter revisit time or higher resolution, are now available. With multi-temporal images, both spatial and temporal information can be exploited in the denoising process, which gives the potential of providing better denoising results than using only a single image. This is the path followed by several kinds of multi-temporal denoising methods, proposing temporal weighted average [16], [17], [18], temporal weighted average in transformed domain (M-TSF [19] and MSAR-BM3D [20]), change detection aware multi-temporal average (L$\hat{e}$'s method [21], [22] and 2SPPB [23]) and filtering using three-dimensional adaptive neighborhoods [24], [25]. A drawback of such approaches is the increased computational complexity with longer time series.

Multi-temporal denoising methods take advantage of the increasing availability of SAR image time-series to solve the spatial denoising problems, for the benefit of resolution preservation. Most of multi-temporal denoising methods use the whole time series in order to denoise a given image. In this paper, we take a different approach by forming a summary of the multi-temporal series through a "super-image", and using

W. Zhao, H. Maître, J.-M. Nicolas, F. Tupin are with LTCI, Télécom ParisTech, Université Paris-Saclay, 75013 Paris, France

L. Denis is with Univ Lyon, UJM-Saint-Etienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, SAINT-ETIENNE, France

C.-A. Deledalle is with IMB, CNRS, Univ. Bordeaux, Bordeaux INP, F-33405 Talence, France

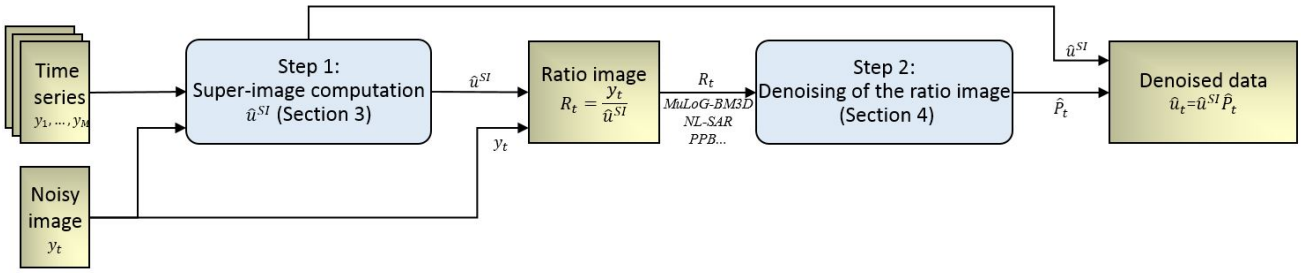


Fig. 1. Multi-temporal SAR image denoising framework. Using y_t for the calculation of super-image $\hat{u}^{SI}$is optional.

this "super-image" to denoise the SAR images at each date.

This ratio-based multi-temporal denoising method fully exploits the significant information of the multi-temporal stack through the "super-image". After forming the ratio image between the noisy image and the "super-image", the proposed method takes advantage of the state-of-the-art single-image speckle reduction methods to denoise the ratio image. We consider multi-temporal images acquired on the same orbit (i.e., either all ascending orbits, or all descending orbits), with similar incidence angles, and which have been accurately registered [26].

The contributions of this paper are the following:

- we provide a generic ratio-based multi-temporal denoising method, called RABASAR, that can embed any state-of-the-art speckle reduction method designed to process a single SAR image;
- we demonstrate the interest of the proposed method on synthetic and real SAR time series.

The remainder of this paper is organized as follows. In section II, we introduce the general framework of the proposed method. Section III presents different ways to compute the "super-image". Then the filtering of the ratio image is described in section IV. Experimental results are presented in section V. Finally, some conclusions and perspectives are drawn in section VI.

II. Principle of the proposed method

Temporal averaging (also called temporal multi-looking) of multi-temporal stacks of images produces images with reduced speckle where spatial resolution is preserved. Temporal averaging provides a so-called "super-image". In this paper, we propose to exploit this super-image to build a ratio-based denoising framework (Fig.1).

The proposed method mainly contains three steps.

1) In the first step, a super-image is calculated using a series of well registered and calibrated SAR images. Averaging temporal intensity samples is the simplest way to obtain an image with speckle reduction. It corresponds to the maximum likelihood estimation of the reflectivity if the variables are i.i.d. (independent and identically distributed, i.e. if temporal fluctuations are independent and only caused by fully-developed speckle). In practice, the influence of the SAR imaging system and the changes of the remote sensing areas (e.g. farmland or building areas) should be taken into account. Even for temporally stable areas, because of possible different temporal correlations, the noise reduction may be variable in space.
   For areas with seasonal changes, a typical evolution of vegetation areas, the changed areas are rather homogeneously distributed and with low fluctuations. On the contrary, for abrupt or impulsive changes [27], as often caused by human activities, the mean value may be seriously influenced by changes. Hence, it can be interesting to average unchanged or similar temporal points through binary weights, as done for instance in [23], [22]. Computation of the super-image, denoted as $\hat{u}^{SI}$ in the sequel, will be presented in Section III.
2) In the second step, using the obtained super-image, we calculate the ratio R_t between the image y_t at time t and the super-image:
$$R_t = \frac{y_t}{\hat{u}^{SI}} \tag{1}$$
   R_t is named the ratio image at time t; it contains the residual speckle noise between the two images, and the radiometric shifts when changes occurs.
   In stable areas, if we have an infinite series of images, the super image will be close to the reflectivity of this area with $\hat{u}^{SI} = u$ [28]. Then, the ratio contains fully developed speckle approximately following a Gamma distribution with unit mean. On the contrary, in case of a change in some area at time t, it is likely that this change will appear in the ratio image. Any speckle filtering technique can be applied to the ratio image. A good speckle reduction method is expected to preserve the possible changes at time t. This point will be discussed in Section IV.
3) In the third step, the filtered image is recovered by multiplying the denoised ratio image with the original super-image. Experimental results on simulated and real SAR data will be presented in section V.

III. Super-image computation

With a multi-temporal stack, the super-image can be calculated through different ways. First, different Hölder means (such as arithmetic or geometric means) could be chosen; then such means may be applied either on the intensity data or on the amplitude data. We may expect from these choices to enhance different pieces of information [29]. In this paper, registered and radiometrically corrected intensity SAR images

are used. We propose to use the arithmetic mean for its good properties [29] with the option of using binary weights to discard the intensity at some dates when a change occurred.

A. Statistics of SAR images

We briefly recall in this section the statistics of fully developed speckle. Under Goodman's hypothesis [30], the fully developed intensity speckle follows a Gamma distribution $\mathcal{G}[u, L]$ depending on the number of looks L and the mean reflectivity u of the scene:

$$\mathcal{G}[u, L](y) = \frac{L}{u\Gamma(L)}\left(\frac{Ly}{u}\right)^{L-1} e^{-\frac{Ly}{u}} \tag{2}$$

Speckle in coherent processed SAR data acts like a multiplicative noise and the speckle model can be expressed as [4]:

$$y = uv \tag{3}$$

with

$$\mathbb{E}[y] = u \tag{4}$$

and

$$\mathrm{Var}[y] = u^2/L, \tag{5}$$

and v follows a Gamma distribution $\mathcal{G}[1, L]$, $\mathbb{E}[\cdot]$ is the expectation operator, $\mathrm{Var}[\cdot]$ represents the variance operator. With the increase of the number of looks L, the variance u^2/L decreases. A multiplicative signal model can be used to describe both intensity and amplitude data.

B. Arithmetic mean

Given a time series of M intensity values $[y_1(s), y_2(s), y_3(s), \cdots, y_M(s)]$ indexed by time t, the arithmetic mean is calculated at location s by:

$$\hat{u}^{AM}(s) = \frac{1}{M}\sum_{t=1}^{M} y_t(s) \qquad 1 \leq t \leq M \tag{6}$$

Theoretically, with *i)* no change in the time series $[u_1(s) = u_2(s) = \cdots = u_t(s) = \cdots = u_M(s)]$ and *ii)* M large enough, averaging the temporal intensity data is a simple yet effective approach to reduce the speckle [31]. The arithmetic mean $\hat{u}^{AM}(s)$ is equal to the maximum likelihood estimation (no matter how large M is) of $u(s)$, and the multi-look image $\hat{u}^{AM}$ follows a Gamma distribution $\mathcal{G}[u, LM]$. In practice, the resulting equivalent number of looks (ENL) may be less than the theoretical value ($L \times M$ with i.i.d. variables) if there are temporal correlations in the time series, especially in case of images in interferometric configuration.

When there are changes in the time series samples $\{y_t(s)\}_{t=1}^{M}$, the arithmetic mean $\hat{u}^{AM}(s)$ has no physical meaning and therefore does not correspond to the scene reflectivity (since it varies in time). We can account for this mismatch between the temporal average and the underlying reflectivity through a correcting factor or by averaging only the unchanged temporal samples.

C. Binary weighted arithmetic mean

Instead of computing blindly the super-image on the time series and then relying on the ratio image to recover the reflectivity of each specific date, another option is to compute a dedicated super-image. In this case, only samples with similar and stable reflectivities are averaged. To detect these samples, a patch-based similarity estimation $w_{t,t'}(s)$ based on the generalized likelihood ratio (GLR) test may be used.

It uses pixel-wise comparisons on a patch taken at date t and date t' with the following formula [8], [23]:

$$w_{t,t'}(s) = \sum_k \left(\log\left(\sqrt{\frac{y_t(s+k)}{y_{t'}(s+k)}} + \sqrt{\frac{y_{t'}(s+k)}{y_t(s+k)}} \right) - \log 2 \right) \tag{7}$$

where $y_t(s+k)$ is the value in the noisy patch at date t. The sum is taken over all pixel shifts k such that pixels with index $s+k$ are located inside a patch centered on s (small square window). Based on the analysis in [8], a size of 7×7 is chosen for the patch. Then, a binary weight $\varphi[w_{t,t'}(s)]$, expressing whether there are temporal changes or not, is computed as:

$$\varphi[w_{t,t'}(s)] = \begin{cases} 1, & \text{if } w_{t,t'}(s) < \sigma \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

with σ as a threshold. σ is computed with pure speckle data as $\sigma = \mathrm{quantile}(w_{t,t'}(s), \alpha)$, with a value α taken as $\alpha = 0.92$ as proposed in [8].

Then, the binary weighted arithmetic mean (denoted by BWAM in the sequel) is calculated through:

$$u_t^{BWAM}(s) = \frac{1}{\sum_{t'=1}^{M}\varphi[w_{t,t'}(s)]}\sum_{t'=1}^{M}\varphi[w_{t,t'}(s)]y_{t'}(s) \tag{9}$$

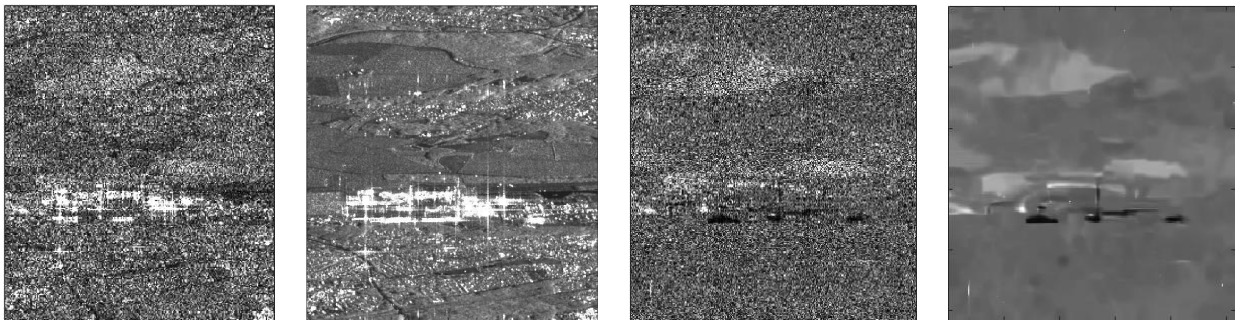
The computation of this BWAM image should be done for each date and each pixel in the temporal stack through a patch-based comparison of the whole stack. This u_t^{BWAM} image is more faithful to the reflectivity of the image at time t at the cost of an increased computational time and a reduced number of looks in u_t^{BWAM}. The interest of using such a temporal mean will be evaluated in Section V.

D. Denoising the super-image

When the number of temporal candidates used for the computation of the super-image is not large enough, the spatial information can be taken into account to improve the estimation of the speckle-free signal. The super-image may have a spatially varying ENL because of unstable areas, because of varying coherence, and because of the use of the binary weights. Let L denote the local equivalent number of looks. To robustly estimate L^{AM} in the arithmetic mean image calculated with media or high resolution temporal SAR images, the log-cumulant method is used [32].

Empirical expressions for the first and second order log-cumulant estimators are (for N samples):

$$\hat{k}_1 = \frac{1}{N}\sum_{i=1}^{N}[\log(\sqrt{\hat{u}_i})] \tag{10}$$



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

Fig. 2. (a) Sentinel-1 noisy image, (b) Arithmetic mean image, (c) Ratio image, (d) MuLoG-BM3D denoised result on ratio image. Appearing (dark areas in (d)) and disappearing buildings (clear areas) are located in the middle of the image.

$$\hat{k}_2 = \frac{1}{N}\sum_{i=1}^{N}[\log(\sqrt{\hat{u}_i}) - \hat{k}_1]^2 \tag{11}$$

Then, we can obtain an estimation of the ENL using the following relationship (theoretical expression):

$$\hat{k}_2 = \frac{1}{4}\psi(1, \hat{L}^{AM}) \tag{12}$$

where $\psi(1, \hat{L}^{AM})$ is the first-order Polygamma function [32]. Note that the traditional estimation method (by means of the ratio $\mathbb{E}(y_i)^2/\mathrm{Var}(y_i)$), the moment estimation method or the maximum likelihood (ML) estimation method could also be used for the ENL estimation [33].

This ENL is locally estimated using a sliding window, and it is an important parameter in statistical modeling of multilook SAR images. A global ENL for the super-image may be deduced from the local ENL as its maximum value. In practice, this global value may be useful to speed up a denoising process applied on the super-image. In this paper, the arithmetic mean image is filtered using MuLoG-BM3D [34]. Any other spatial adaptive denoising method could be used.

Therefore, 4 super-images can be computed: arithmetic mean image (AM, according to Eq.(6)), binary weighted arithmetic mean image (BWAM, according to Eq.(9)) and their denoised versions denoted as DAM and DBWAM. The interest of this step will be evaluated in Section V.

IV. Ratio image denoising

The ratio image (Eq.(1)) should contain both the variations between the super-image $\hat{u}^{SI}$ and u_t and the speckle of the noisy image y_t. Nonetheless, the spatial homogeneity is largely improved in the ratio image compared to the original noisy image.

At this point, any denoising method can be used to denoise the ratio image R_t. In this paper, MuLoG [34] has been chosen. Unlike homomorphic approaches, and inspired by [35], MuLoG relies on the Fisher-Tippett distribution to approximate the reflectivities of the log-transformed data. To lighten the notation, R is the ratio image of a date t, $r = \log R$ and P is the true ratio with $\rho = \log P$.

Its estimation is based on a MAP optimization approach as:

$$\rho \in \underset{\rho\in\mathbb{R}^n}{\arg\min}(-\log p_r(r|\rho) + \mathcal{R}(\rho)) \tag{13}$$

with

$$-\log p_r(r|\rho) = L\sum_{h=1}^{N}\rho_h + e^{r_h-\rho_h} + Cst \tag{14}$$

and $\mathcal{R}(\rho) = -\log p_\rho(\rho)$ is a prior term enforcing some regularity on the solution, N is the number of pixels and Cst is the constant term. Then, the denoised data is reconstructed by $\hat{P} = \exp(\rho)$. In this paper, we consider MuLoG+BM3D that uses BM3D [36] as an implicit regularization term $\mathcal{R}(\rho)$.

An example is shown in Fig.2. A typical filtering result on a ratio image is presented on Fig.2(d) along with the original image, super-image and the ratio image.

After obtaining the estimated noise free value $\hat{P}_t$ of the ratio, we obtain the denoised image $\hat{P}_t$ through:

$$\hat{u}_t = \hat{u}^{SI}\hat{P}_t \tag{15}$$

Algorithm Multi-temporal speckle reduction (RABASAR)

Input: M co-registered SAR images $\{y_1, y_2, \cdots, y_M\}$
Output: M images with reduced speckle $\{\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_M\}$

Step 1: computation of the super-image
for each pixel s **do**
  $\hat{u}(s) \leftarrow \frac{1}{M}\sum_{t=1}^{M} y_t(s)$
end for
(optional: denoise the multi-looked image)
if denoise super-image **then**
  $\hat{L} \leftarrow$ `estimate_looks`$(\hat{u})$
  $\hat{u} \leftarrow$ `MuLoG_BM3D`$(\hat{u}, L = \hat{L})$
end if
for t from 1 to M **do**
  Step 2: denoising of the ratio image
  $\hat{P}_t \leftarrow$ `MuLoG_BM3D`$(y_t/\hat{u}, L = 1)$
  Step 3: computation of the restored image
  $\hat{u}_t \leftarrow \hat{u} \cdot \hat{P}_t$
end for
return $\{\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_M\}$

V. Experimental results

To evaluate the performances of the proposed method, different experiments have been conducted on simulated and real SAR images. The influence of different super-images on the RABASAR denoising results are commented (section

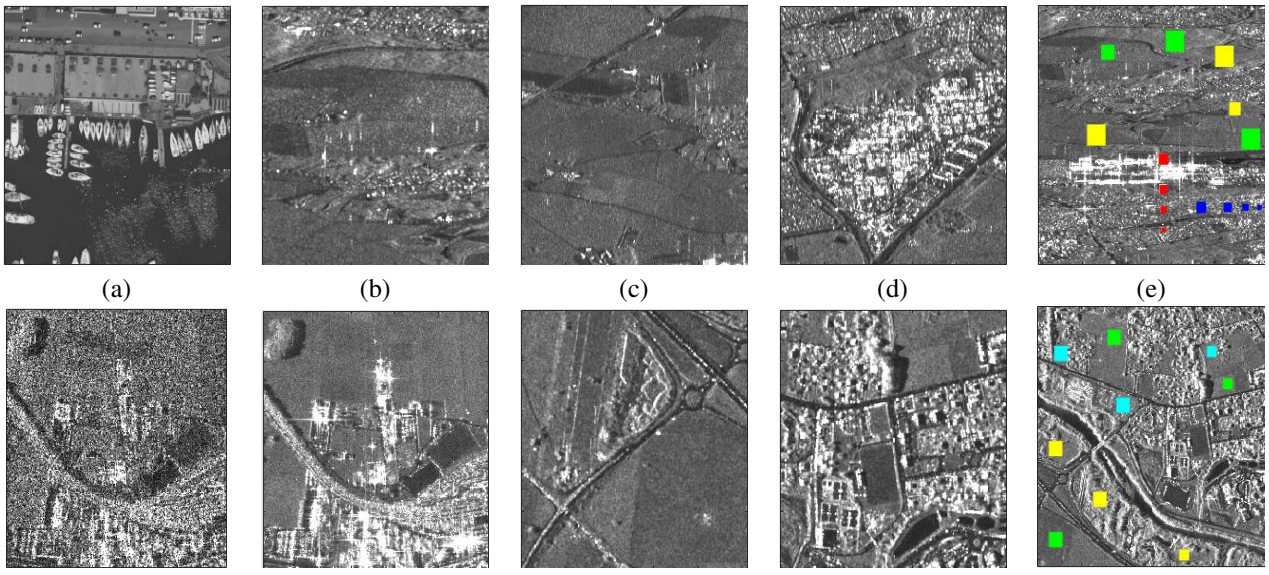


Fig. 3. Different types of images: (a) optical image, (b-e) arithmetic mean of Sentinel-1 images over different land covers, (f-g) noisy and arithmetic mean of TerraSAR-X images, (h-j) arithmetic mean of TerraSAR-X images over different areas. Different colors represent different changes: Green=farmland, yellow=forest, red=disappearing then appearing, blue=appearing then disappearing, cyan=cycle changes.

V-C). Then, the ratio-based multi-temporal denoising methods are compared with some methods selected from the literature: UTA [16], [17], NLTF[20], 2SPPB [23] and MSAR-BM3D [20] (section V-D).

A. Data presentation

1) Real data: The proposed methods are tested on a time series of 69 descending Sentinel-1 IW SLC Level-1 images acquired from December 24, 2014 to May 6, 2017 with VV polarization over Saclay area, South of Paris[1]. Saclay plateau is mostly an agricultural area with pieces of forests and dispersed academic buildings. In the last five years, many new buildings and infrastructures have been under development.

In addition, 26 single-look TerraSAR-X images (13 images are sensed in 2009 and the other 13 images in 2011) acquired over Saint-Gervais-les-Bains, South-East of Geneva, (project DLR-MTH0232) are used. These images are taken over a highly mountainous countryside, with a narrow inhabited valley concentrating many human artifacts (roads, bridges, dams).

2) Simulated data: Simulated SAR images are usually obtained through equation (3), by multiplying a reflectivity map with a random Gamma distributed noise.

Many simulations are based on reflectivity maps obtained from optical images. However, real SAR images exhibit strong and persistent scatterers, especially in urban areas which can hardly be simulated using optical images. Therefore, we propose to use the arithmetic mean image of long time-series of SAR images, considered as a noise free image (a reflectivity map u) to create realistic simulations of SAR images. This map u is multiplied by a gamma distributed noise v_t providing an image $y_t = uv_t$ of the series. For specific applications, dedicated sequences will be created over various areas: forests, farmlands, building areas etc. as shown in Fig.3. Different temporal changes may also be simulated as shown in (Fig.3 (e) and (j)). Changed values are extracted from the corresponding real SAR time series.

It often happens that SAR pixels are not spatially independent because of a slight over-sampling creating a spatial correlation. This spatial correlation should be taken into account during the parameter estimation. However, most of the denoising methods are based on the hypothesis that the speckle component is i.i.d.. When blindly applied to correlated data, reduced performances may be expected from these methods. Therefore, we recommend to perform a spatial decorrelation before despeckling. In this paper, the noisy TerraSAR-X images are decorrelated using the method proposed in [37] and the Sentinel-1 images are decorrelated by resampling because of its special acquisition model (the beam both steering in range direction and steering from backward to forward in azimuth direction). All the real SAR images are well registered using subpixel image registration applied on the single look complex data [26].

B. Evaluation methods

Measurement of speckle reduction performances is a challenging task, especially when noise-free data are unavailable. Visually checking the despeckling results is an immediate and important way for quality evaluation, but it lacks objectivity. To overcome this limitation, peak signal-noise-ratio (PSNR) and structure similarity (SSIM) indexes may be used, but their limits have been abundantly discussed.

1) PSNR: Peak signal-noise-ratio: PSNR is a commonly used approach to evaluate the quality of noise-free reflectivity restoration results. The denoising results can be quantified by the PSNR with:

[1]All the Sentinel-1 images can be downloaded from Copernicus Open Access Hub (https://sentinels.copernicus.eu/web/sentinel/ sentinel-data-access/accessto-sentinel-data).

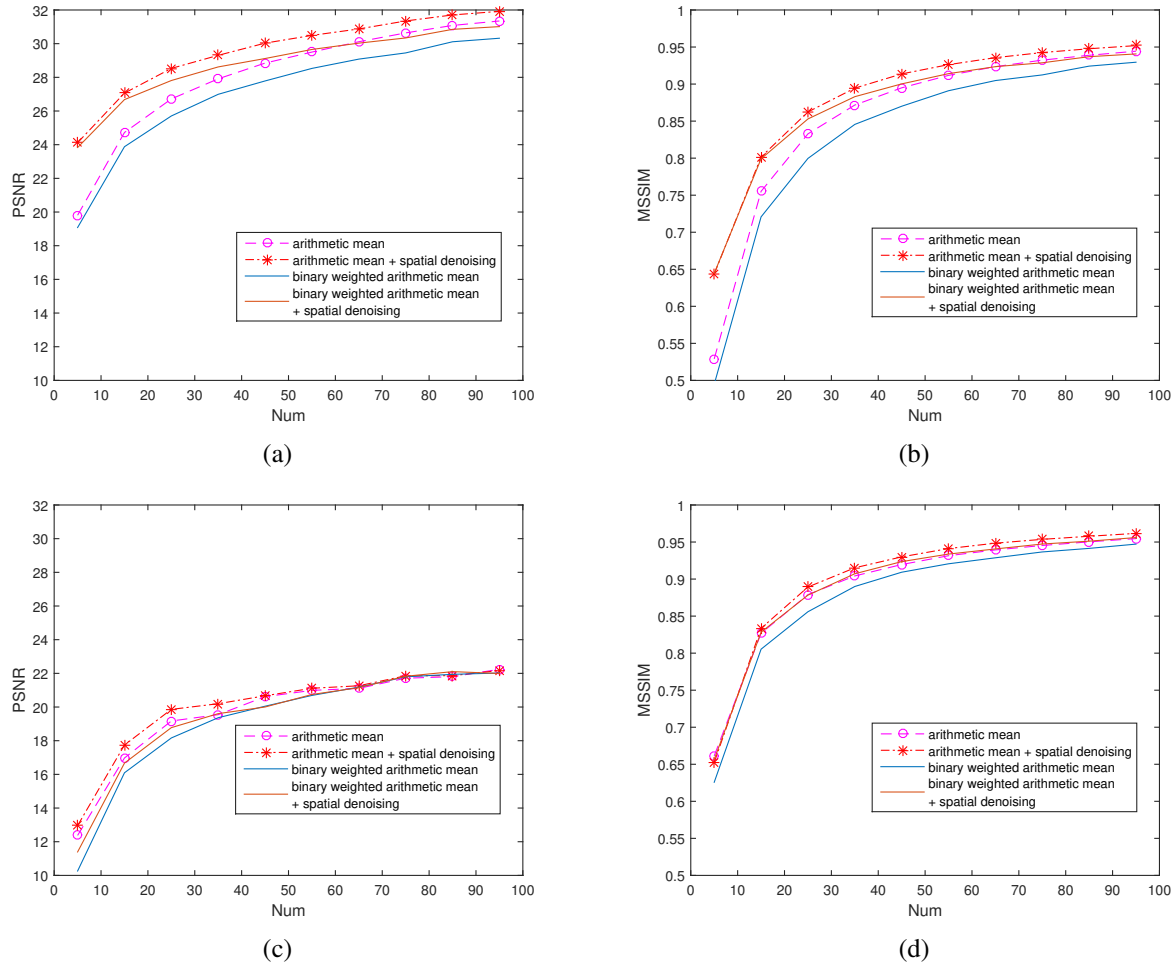


Fig. 4. RABASAR performances with different super-images. (a) PSNR in case of images without change (Fig.3(b)), (b) MSSIM under the same conditions. (c) PSNR as a function of the number of images in the case of existing changes in the stack (Fig.3(e)), (d) MSSIM in the same conditions,

$$\mathrm{PSNR} = 10 \cdot \log_{10} \frac{|u_A|^2_{max}}{\mathbb{E}[(u_A(s) - \hat{u}_A(s))^2]} \tag{16}$$

where $|u_A|_{max}$ is the maximum amplitude value in the noise free data, $\mathbb{E}[\cdot]$ represents the spatial average and $\hat{u}_A$ is the denoised amplitude value.

2) MSSIM: Mean structure similarity index: To evaluate the preservation of image features, SSIM index [38] (structural similarity index measurement) is often preferred to PSNR. From the SSIM, we derive the mean structural similarity index measurement (MSSIM) values which provide a comprehensive measure over the whole image:

$$\mathrm{MSSIM} = \frac{1}{N} \sum_{i=1}^{N} \left[\frac{2 \cdot \widehat{\mathbb{E}}[u_A] \cdot \widehat{\mathbb{E}}[\hat{u}_A] + \alpha_1}{\widehat{\mathbb{E}}[u_A^2] + \widehat{\mathbb{E}}[\hat{u}_A^2] + \alpha_1} \frac{2 \cdot \widehat{\mathrm{Cov}}[u_A, \hat{u}_A] + \alpha_2}{\widehat{\mathrm{Var}}[u_A] + \widehat{\mathrm{Var}}[\hat{u}_A] + \alpha_2} \right]$$

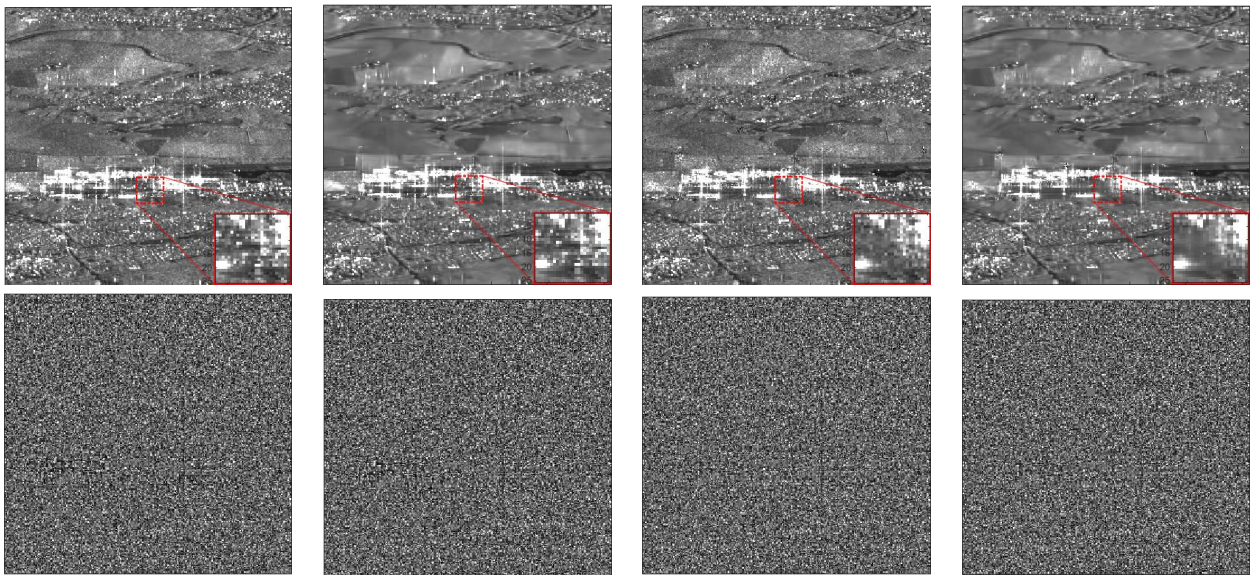
where u_A and $\hat{u}_A$ are noise free and denoised patches, $\mathrm{Cov}[\cdot, \cdot]$ is the measure of covariance, α_1 and α_2 are suitable constants, N is the number of local windows in the image.

C. Which super-image gives the best denoising?

As proposed earlier (see Section III-D), different super-images (AM, BWAM and their spatially denoised versions) may be computed with the same multi-temporal stack. This section presents quantitative (Fig.4) and qualitative (Fig.5) results of the denoising stage, obtained by using different super-images with different stack lengths. The parameters used with each super-image are shown in Table I. The ENL of the super-image is estimated using the log-cumulant method with a window size 30×30, and the maximum estimated ENL is used. The ENL of the ratio image is supposed to be the same as the one of the noisy image (L).

TABLE I
PARAMETER SETTING WITH DIFFERENT SUPER-IMAGES (L IS THE ENL OF THE NOISY IMAGE)

Super-image	Denoising step (MuLog-BM3D)	ENL estimation	Similarity window [39]
AM	Step 2	L	
DAM	Step 1-2	log-cumulant	
BWAM	Step 2	L	7×7
DBWAM	Step 1-2	log-cumulant	7×7



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

Fig. 5. (a-d) RABASAR denoising results on real Sentinel-1 images over the Saclay area (higher row) and corresponding ratio-images (lower row) based on the use of: (a) AM, (b) DAM, (c) BWAM, (d) DBWAM. 69 Sentinel-1 images are used.

1) Using synthetic Sentinel-1 images: Synthetic images allow to evaluate numerically the performances obtained with four different super-images by measuring PSNR and MSSIM values. The number of images in the time series varies from 5 to 95.

For temporal images without changes (Fig.4(a-b)), using denoised super-images provides better PSNR and MSSIM values when using a small number of images, but this benefit disappears with long stacks of images (more than 60). When using denoised super-images, the obtained PSNR/MSSIM values are seriously influenced by the used ENL. Smaller ENL estimation window size is recommended when fewer temporal images are available. Because of the i.i.d. time series variables, we set the ENL equals to the number of used temporal synthetic images. We recommend using a spatially adaptive ENL for the denoising of BWAM, because different number of temporal samples may be used for its computation.

When changes occur (Fig.4(c-d)), RABASAR can provide better PSNR with AM and DAM for short stacks (under 30 images). All the strategies are almost similar when more images are used. With temporally changed images, RABASAR PSNR values are much lower than that using temporally stable images. Using BWAM image provides slightly worse MSSIM values while using DAM provides better MSSIM.

2) Using Sentinel-1 images: The temporal series of 69 Sentinel-1 images on the Saclay area is used to test the method on real images. Figure 5 can be used to visually assess the noise reduction efficiency of using different "super-images". RABASAR provides good denoising results with the four different super-images. The use of an additional spatial filtering step to form the super-image seems beneficial in terms of restoration quality: the obtained images are smoother.

For results which used AM and DAM, small areas with low values were sometimes smoothed leading to the apparition of new points in the denoised results (Fig.5(a-b) red rectangular). This phenomenon is obvious for impulsive and abrupt changes in building areas. Using BWAM and DBWAM reduces this problem (Fig.5(c-d)). In some changing parts of the image, using BWAM, however, leads to poor filtering results because of using only few similar points to compute the super-image.

3) Computation time: The computation time of the algorithm depends on the adopted strategy, i.e. the type of super-image and the choice of the spatial denoising method. Using denoised super-images is far more time-consuming than using unprocessed super-images. Additionally, because of the dedicated super-image computed for each noisy data, using BWAM image takes more time than using AM.

D. Denoising performances of RABASAR compared to existing methods

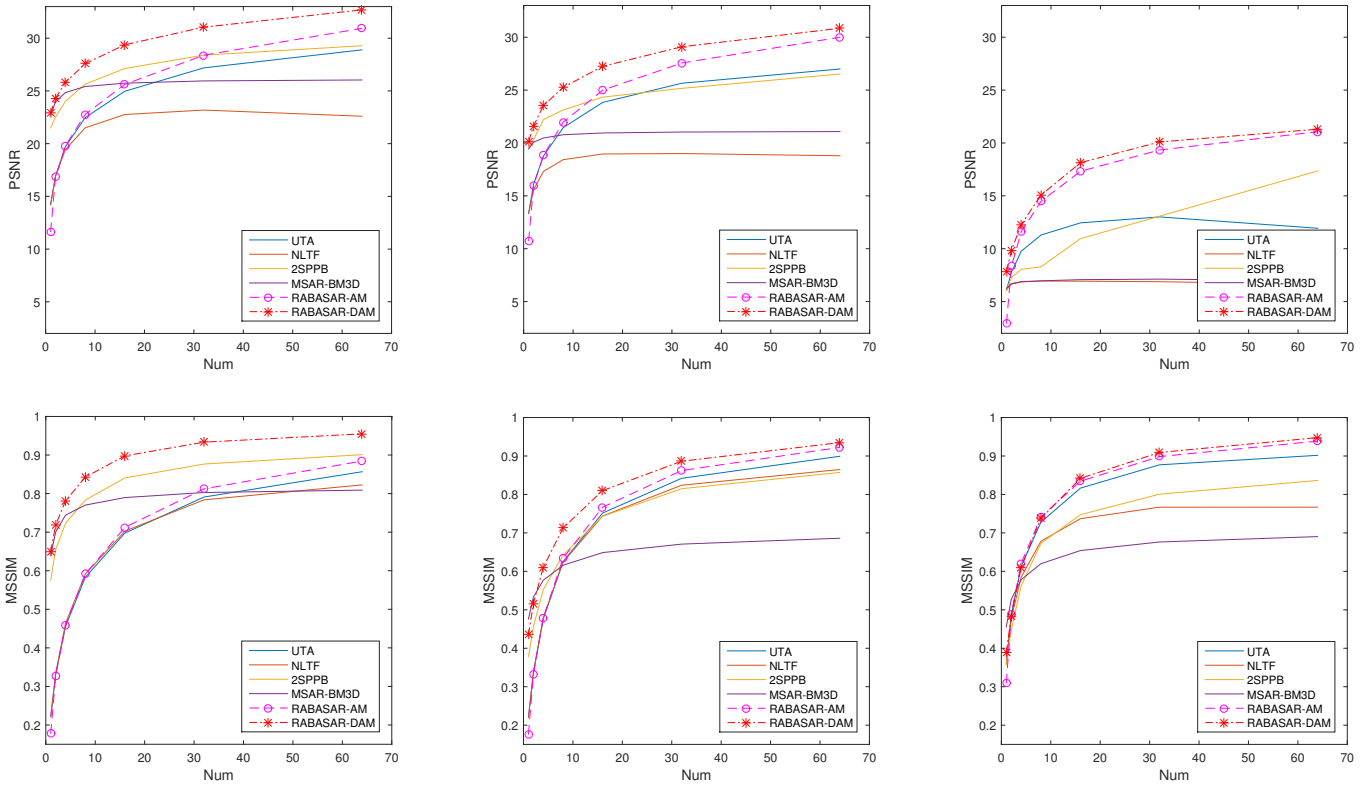
The proposed method is compared with state-of-the-art multi-temporal denoising methods, both with synthetic SAR and with real SAR images. Numerical and visual results are provided when comparing RABASAR with the chosen methods: UTA [17], NLTF [20], 2SPPB [23] and MSAR-BM3D [20].

1) Quantitative comparison: As in the previous sections, to quantitatively compare the filtering performance of these methods, PSNR and MSSIM are computed. ROC curves are also presented, but with only RABASAR-AM and RABASAR-DAM to make the demonstration clearer. Since MSAR-BM3D released code requires a number of time series equal to a power of 2, we only tested such cases, to keep the original performances of this method.

To make a fair comparison with other methods, three kinds of temporal data are tested (Fig.6). In the case where no change or bright echo exist in the simulated image, we see, from figure 6, that RABASAR generally provides better PSNR than other filtering methods. This may be explained because, under these hypotheses, the ratio image is very homogeneous.

TABLE II
NUMERICAL RESULTS PROVIDED BY DIFFERENT METHODS WITH 32 IMAGES IN THE SERIES. EVALUATION IS MADE BY PSNR AND MSSIM. FOR PSNR AND MSSIM, LARGER VALUES EXPRESS BETTER DENOISING RESULTS.

Sentinel-1	Evaluation methods	UTA	NLTF	2SPPB	MSAR -BM3D	RABASAR -AM	RABASAR -DAM	RABASAR -BWAM	RABASAR -DBWAM
without changes	PSNR	25.65	19.00	25.17	21.03	27.56	**29.10**	26.64	28.40
	MSSIM	0.84	0.82	0.81	0.67	0.86	**0.89**	0.84	0.88
with changes	PSNR	13.01	6.88	13.11	7.12	19.34	**20.10**	19.00	19.50
	MSSIM	0.88	0.77	0.80	0.68	0.90	**0.91**	0.88	0.90



(a) (b) (c)

Fig. 6. Comparison of different temporal denoising methods: UTA, NLTF, 2SPPB, MSAR-BM3D, RABASAR-AM and RABASAR-DAM. PSNR (upper row) and MSSIM (lower row) as a function of the number of images: (a) using temporal simulated data without changes (Fig.3(a)), (b) with unchanged realistic synthetic SAR images, (c) with changed realistic synthetic SAR images.

With fewer images in the stack, MSAR-BM3D and 2SPPB also provide competitive PSNR and MSSIM. However, with the increase of the number of images, MSAR-BM3D curve does not rise as fast as 2SPPB. When using less than 4 synthetic images, MSAR-BM3D provides the best MSSIM values.

With an increasing number of images without change, UTA PSNR and MSSIM values keep increasing. Due to the use of a sliding window to estimate the reflectivity, UTA does not give good results when few images are used. Since there are large building areas in the time series with changes, it suffers from blurred boundaries. Table II presents some results about PSNR and MSSIM results.

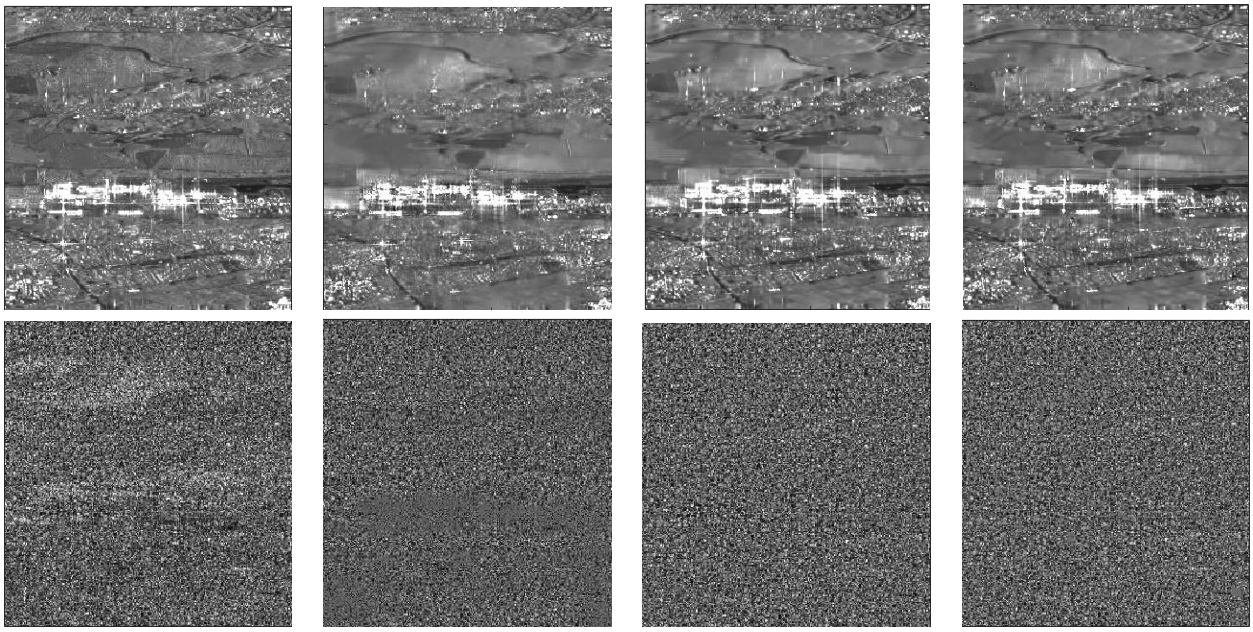
2) Synthetic SAR data with different ENL: To evaluate the effects of the ENL on the denoising performances, different denoising methods are applied on the data selected when creating the simulated samples over various landscapes (see Section V-A2). Results are presented in Table III. In the case of long time series (here 64 temporal images are used), the various RABASAR methods always provide better results than other methods. When there are changes in the time series, using BWAM may be better than using AM.

3) Comparison for real SAR image denoising: This section presents results obtained when denoising real Sentinel-1 and TerraSAR-X images. Since we do not have noise free images, the denoised results are compared using their ratio with noisy image.

Compared to 2SPPB and MSAR-BM3D methods, RABASAR-DAM and RABASAR-DBWAM provide better denoising results by taking both the ratio results and the texture characteristics into account. 2SPPB method does not give good results in homogeneous areas, which leads to an obvious bias in the ratio image. For example, there are two bright areas in the ratio of 2SPPB denoised Sentinel-1 image

TABLE III
PSNR-MSSIM VALUES PROVIDED BY DIFFERENT DENOISING METHODS USING DIFFERENT DATA ISSUED FROM THE SIMULATED SAR IMAGES (SEE FIGURE 3). THE NUMBER OF IMAGES IN THE SERIES IS 64. SENTINEL-1 IMAGES ARE EXPRESSED BY S-, TERRASAR-X IMAGES ARE DENOTED WITH T-.

Data	ENL	UTA	2SPPB	NLTF	MSAR-BM3D	RABASAR-AM	RABASAR-BWAM
S-farmland	1	27.98-0.88	27.52-0.83	20.69-0.86	23.19-0.62	**30.50-0.90**	29.74-0.88
S-farmland	4	33.62-0.97	30.13-0.88	29.52-0.97	26.90-0.78	**36.55-0.97**	**36.55-0.97**
S-farmland	8	37.13-0.98	31.80-0.89	35.48-0.99	29.36-0.85	39.50-0.98	**39.50-0.99**
S-building	1	20.47-0.94	18.68-0.89	10.57-0.76	13.36-0.77	**26.17-0.96**	25.02-0.95
S-building	4	25.84-0.98	26.94-0.94	19.73-0.98	19.26-0.89	**31.97-0.99**	**31.97-0.99**
S-building	8	28.86-0.99	31.84-0.96	23.69-0.99	21.66-0.92	**35.72-0.99**	35.67-0.99
S-changes	1	10.03-0.90	17.67-0.84	06.18-0.77	05.68-0.69	19.59-0.94	**21.14-0.93**
S-changes	4	14.17-0.96	23.88-0.89	11.50-0.96	12.19-0.85	24.97-0.98	**26.05-0.98**
S-changes	8	16.48-0.97	27.17-0.93	14.93-0.98	16.34-0.90	27.87-0.99	**28.86-0.99**
T-farmland	1	27.79-0.89	27.34-0.85	21.26-0.87	24.22-0.69	**30.57-0.92**	29.29-0.89
T-farmland	4	33.49-0.97	29.34-0.89	29.62-0.98	26.95-0.82	**36.24-0.98**	36.24-0.98
T-farmland	8	36.47-0.98	31.00-0.91	35.94-0.99	29.32-0.87	**39.54-0.99**	39.54-0.99
T-building	1	21.32-0.94	19.43-0.90	11.36-0.80	14.15-0.81	**26.04-0.96**	24.83-0.95
T-building	4	26.09-0.98	25.73-0.94	21.60-0.99	19.16-0.90	**31.83-0.99**	**31.83-0.99**
T-building	8	28.34-0.99	30.13-0.96	26.71-0.99	21.50-0.94	**34.96-0.99**	**34.96-0.99**
T-changes	1	21.16-0.94	19.07-0.85	12.60-0.74	15.45-0.76	**27.00-0.96**	26.20-0.95
T-changes	4	25.39-0.98	26.42-0.93	20.26-0.97	19.46-0.88	**32.92-0.98**	32.86-0.98
T-changes	8	29.34-0.99	31.43-0.96	24.44-0.98	22.27-0.92	35.99-0.99	**36.00-0.99**



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

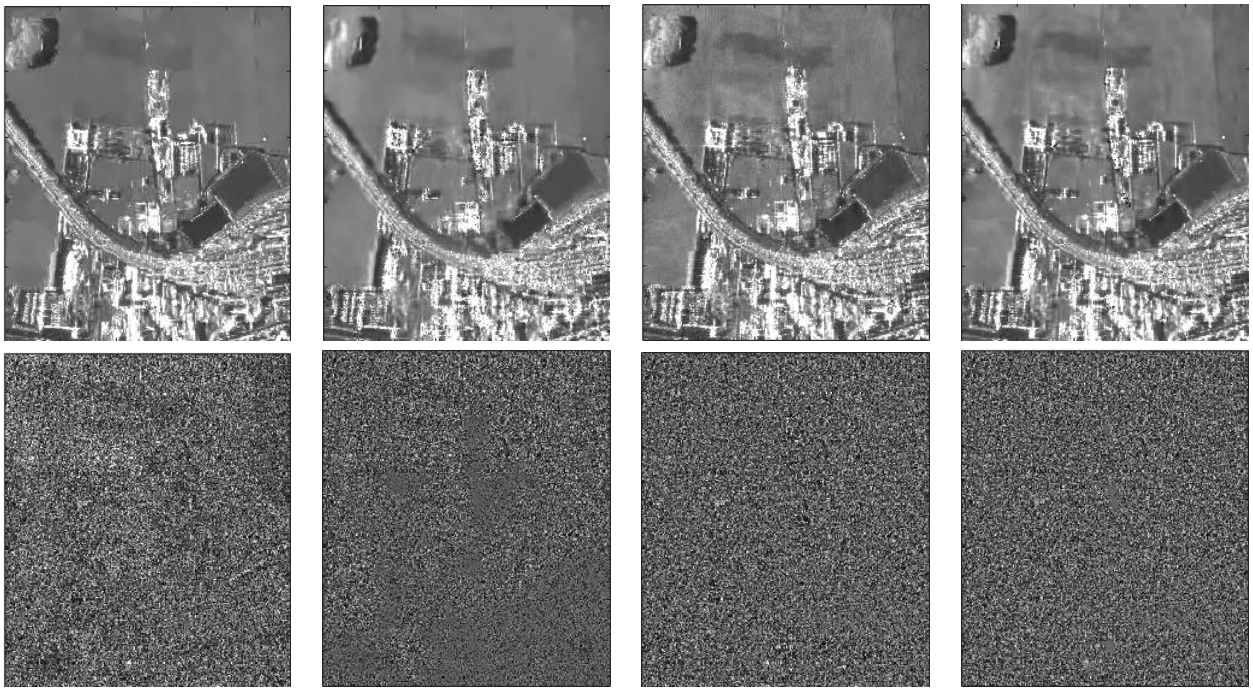
Fig. 7. Denoising real Sentinel-1 images over the region of Saclay (the original noisy image is available in figure 2(a)) : (a) filtered with 2SPPB, (b) with MSAR-BM3D, (c) with RABASAR-DAM, (d) with RABASAR-DBWAM. Top row: denoised results; Lower row: residual ratio images.

(Fig.7(a)). Since MSAR-BM3D method detects the bright points in advance and prohibits any denoising for these points [20], building areas in the ratio image are very smooth and their denoising results are still influenced by noise (Fig.7(b) and Fig.8(b)).

VI. CONCLUSION AND FUTURE WORK

This paper has proposed a ratio-based multi-temporal denoising framework. During the restoration of each SAR image, it exploits the temporal information through a super-image. The use of few kinds of super-images has been analyzed. RABASAR can provide better PSNR and MSSIM values when using a spatially denoised super-image. With the increase of the number of images in the time series, the differences of using different super-images decrease. When there are changes in the time series, using a binary weighted arithmetic mean can also provide good results. Based on the processing of simulated time-series, actual Sentinel-1 stacks and TerraSAR-X stacks, the qualitative and quantitative comparison with UTA, NLTF, MSAR-BM3D and 2SPPB methods showed the potential of RABASAR to better preserve structures in multi-temporal SAR images while efficiently removing speckle. Besides, the super-image can be easily updated when a new data becomes available so as to process new images on-line.

The proposed method can use other super-images such as geometric mean, quadratic mean, etc., without modifying



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

Fig. 8. Different denoising results with TerraSAR-X images over Sain-Gervais. The original noisy image is available in figure 3(f). First row: the denoised images; Second row: residual ratio images. (a) with 2SPPB, (b) with MSAR-BM3D, (c) with RABASAR-DAM, (d) with RABASAR-DBWAM.

the calculation procedure. The ratio image denoising method can be replaced by any other single-image speckle reduction method. Future work will be devoted to the updating framework, specially for the "re-computation" of the super-image and to the application of denoised results.

ACKNOWLEDGMENT

The authors would like to thank China Scholarship Council and the CNES (French Space Agency, project DAJ/AR/IB-2016-10117102) for funding .

REFERENCES

[1] F. Argenti, A. Lapini, T. Bianchi, and L. Alparone, "A tutorial on speckle reduction in Synthetic Aperture Radar images," *IEEE Geoscience and remote sensing magazine*, vol. 1, no. 3, pp. 6–35, 2013.

[2] R. Touzi, "A review of speckle filtering in the context of estimation theory," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 11, pp. 2392–2404, 2002.

[3] C. Deledalle, L. Denis, G. Poggi, F. Tupin, and L. Verdoliva, "Exploiting patch similarity for SAR image processing: the nonlocal paradigm," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 69–78, 2014.

[4] J. Lee, "Speckle analysis and smoothing of Synthetic Aperture Radar images," *Computer graphics and image processing*, vol. 17, no. 1, pp. 24–32, 1981.

[5] ——, "Refined filtering of image noise using local statistics," *Computer graphics and image processing*, vol. 15, no. 4, pp. 380–389, 1981.

[6] P. Meer, Park, R.H., and K. Cho, "Multiresolution adaptive image smoothing," *CVGIP: Graphical Models and Image Processing*, vol. 56, no. 2, pp. 140–148, 1994.

[7] G. Vasile, E. Trouvé, J.-S. Lee, and V. Buzuloiu, "Intensity-driven adaptive-neighborhood technique for polarimetric and interferometric SAR parameters estimation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1609–1621, 2006.

[8] C. Deledalle, L. Denis, and F. Tupin, "Iterative weighted maximum likelihood denoising with probabilistic patch-based weights," *IEEE Transactions on Image Processing*, vol. 18, no. 12, pp. 2661–2672, 2009.

[9] C. Deledalle, L. Denis, F. Tupin, Reigber, and M. A. and Jäger, "NL-SAR: A unified nonlocal framework for resolution-preserving (Pol)(In) SAR denoising," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2021–2038, 2015.

[10] S. Foucher, "SAR image filtering via learned dictionaries and sparse representations," in *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, vol. 1. IEEE, 2008, pp. I–229.

[11] S. Parrilli, M. Poderico, C. Angelino, and L. Verdoliva, "A nonlocal SAR image denoising algorithm based on LLMMSE wavelet shrinkage," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 606–616, 2012.

[12] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.

[13] G. Chierchia, D. Cozzolino, G. Poggi, and L. Verdoliva, "SAR image despeckling through convolutional neural networks," *arXiv preprint arXiv:1704.00275*, 2017.

[14] P. Wang, H. Zhang, and V. Patel, "SAR image despeckling using a convolutional neural network," *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1763–1767, 2017.

[15] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 1, pp. 6–43, 2013.

[16] J. Lee, M. Grunes, and S. Mango, "Speckle reduction in multipolarization, multifrequency SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 29, no. 4, pp. 535–544, 1991.

[17] S. Quegan, T. Le Toan, J. Yu, F. Ribbes, and N. Floury, "Multitemporal ERS SAR analysis applied to forest mapping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2, pp. 741–753, 2000.

[18] S. Quegan and J. Yu, "Filtering of multichannel SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 11, pp. 2373–2379, 2001.

[19] D. Coltuc, E. Trouvé, F. Bujor, N. Classeau, and J. Rudant, "Time-space filtering of multitemporal SAR images," *In Geoscience and Remote Sensing Symposium, Proceedings. IGARSS 2000.*, vol. 7, pp. 2909–2911, 2000.

[20] G. Chierchia, M. El Gheche, G. Scarpa, and L. Verdoliva, "Multitemporal SAR image despeckling based on block-matching and collaborative filtering," *IEEE Transactions on Geoscience and Remote Sensing*, 2017.

[21] T. Lê, A. Atto, E. Trouvé, and J. Nicolas, "Adaptive multitemporal SAR image filtering based on the change detection matrix," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 10, pp. 1826–1830, 2014.

[22] T. Lê, A. 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 Journal of Photogrammetry and Remote Sensing*, vol. 107, pp. 64–76, 2015.

[23] X. Su, C. Deledalle, F. Tupin, and H. Sun, "Two-step multitemporal nonlocal means for Synthetic Aperture Radar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6181–6196, 2014.

[24] M. Ciuc, P. Bolon, E. Trouvé, V. Buzuloiu, and J. Rudant, "Adaptive-neighborhood speckle removal in multitemporal Synthetic Aperture Radar images," *Applied Optics*, vol. 40, no. 32, pp. 5954–5966, 2001.

[25] S. Lobry, L. Denis, and F. Tupin, "Multitemporal SAR image decomposition into strong scatterers, background, and speckle," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 8, pp. 3419–3429, 2016.

[26] J. Nicolas, E. Trouve, R. Fallourd, F. Vernier, F. Tupin, O. Harant, M. Gay, and L. Moreau, "A first comparison of Cosmo-Skymed and TerraSAR-X data over Chamonix Mont-Blanc test-site," *In Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 5586–5589, Jul. 2012.

[27] X. Su, C. Deledalle, F. Tupin, and H. Sun, "NORCAMA: Change analysis in SAR time series by likelihood ratio change matrix clustering," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 101, pp. 247–261, 2015.

[28] L. Gomez, R. Ospina, and A. Frery, "Unassisted quantitative evaluation of despeckling filters," *Remote Sensing*, vol. 9, no. 4, p. 389, 2017.

[29] G. Quin, B. Pinel-Puyssegur, J. Nicolas, and P. Loreaux, "MIMOSA: An automatic change detection method for SAR time series," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5349–5363, 2014.

[30] J. Goodman, *Speckle phenomena in optics: theory and applications*. Roberts and Company Publishers, 2007.

[31] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar images*. SciTech Publishing, 2004.

[32] C. Tison, J. Nicolas, F. Tupin, and H. Maître, "A new statistical model for markovian classification of urban areas in high-resolution SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 10, pp. 2046–2057, 2004.

[33] J. Nicolas and S. Anfinsen, "Introduction to second kind statistics: Application of log-moments and log-cumulants to the analysis of radar image distributions," *Trait. Signal*, vol. 19, no. 3, pp. 139–167, 2002.

[34] C. Deledalle, L. Denis, S. Tabti, and F. Tupin, "MuLoG, or how to apply Gaussian denoisers to multi-channel SAR speckle reduction?" *IEEE Transactions on Image Processing*, 2017.

[35] J. M. Bioucas-Dias and M. A. Figueiredo, "Multiplicative noise removal using variable splitting and constrained optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 7, pp. 1720–1730, 2010.

[36] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[37] R. Abergel, L. Denis, S. Ladjal, and F. Tupin, "Sidelobes suppression and strong targets extraction in single look complex SAR images with several applications," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018.

[38] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[39] A. Buades, B. Coll, and J. Morel, "A non-local algorithm for image denoising," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2. IEEE, 2005, pp. 60–65.