



HAL
open science

Two-Step Multitemporal Nonlocal Means for Synthetic Aperture Radar Images

Charles-Alban Deledalle, Florence Tupin, Xin Su

► **To cite this version:**

Charles-Alban Deledalle, Florence Tupin, Xin Su. Two-Step Multitemporal Nonlocal Means for Synthetic Aperture Radar Images. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52, pp.6181 - 6196. 10.1109/tgrs.2013.2295431 . hal-03169399

HAL Id: hal-03169399

<https://telecom-paris.hal.science/hal-03169399v1>

Submitted on 15 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Two-Step Multitemporal Nonlocal Means for Synthetic Aperture Radar Images

Xin Su, Charles-Alban Deledalle, Florence Tupin, *Senior Member, IEEE*, and Hong Sun, *Member, IEEE*

Abstract—This paper presents a denoising approach for multitemporal synthetic aperture radar (SAR) images based on the concept of nonlocal means (NLM). It exploits the information redundancy existing in multitemporal images by a two-step strategy. The first step realizes a nonlocal weighted estimation driven by the redundancy in time, whereas the second step makes use of the nonlocal estimation in space. Using patch similarity misregistration estimation, we also adapted this approach to the case of unregistered SAR images. The experiments illustrate the efficiency of the proposed method to denoise multitemporal images while preserving new information.

Index Terms—Multitemporal denoising, nonlocal means (NLM), synthetic aperture radar (SAR) images.

I. INTRODUCTION

WITH increasing number of synthetic aperture radar (SAR) systems, many SAR images of the same region are nowadays available. Those time-series or multitemporal SAR images provide users more comprehensive information, including both spatial and temporal domains. There are different applications using multitemporal series: denoising and estimation [1]–[4] and change detection [5], [6] and classification [7], [8]. In this paper, we will focus on multitemporal denoising with the purpose of exploiting available information existing in the temporal series as much as possible.

Beginning from decades ago, SAR image denoising (or despeckling) has been a well-studied problem in the image processing and remote sensing community. State-of-the-art algorithms such as those by Lee [9], [10], Kuan *et al.* [11], [12], and Frost *et al.* [13] consider SAR despeckling based on estimation theory. They develop denoising techniques under minimum mean square error or maximum *a posteriori* estimation theory. Touzi [14] proposed a structural multiresolution framework to handle stationary and nonstationary signals and improve the previous approaches by context adaptation.

Another family of approaches based on Wavelet can be readily extended to SAR despeckling after a logarithm transfor-

Manuscript received January 15, 2013; revised June 30, 2013 and October 9, 2013; accepted November 20, 2013.

X. Su and F. Tupin are with the Institut Mines-Télécom, Télécom ParisTech, Centre National de la Recherche Scientifique—Laboratoire Traitement et Communication de l'Information, 75634 Paris, France (e-mail: xin.su@telecom-paristech.fr; florence.tupin@telecom-paristech.fr).

C.-A. Deledalle is with the Institut de Mathématiques de Bordeaux, Centre National de la Recherche Scientifique—Université Bordeaux I, 33405 Talence, France (e-mail: charles-alban.deledalle@math.u-bordeaux1.fr).

H. Sun is with the School of Electronic Information, Wuhan University, Wuhan 430070, China (e-mail: hongsun@whu.edu.cn).

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.2013.2295431

mation. Typical successful examples are [15] and [16]. Considering the heavy-tailed feature of speckle distribution, a heavy-tailed model under Bayesian wavelet shrinkage is proposed in [17]. Recently, based on spatially adaptive Wavelet analysis, Ranjani and Thiruvengadam [18] and Gleich *et al.* [19] used spatial information to better preserve image details, such as edges and textures.

In 2005, Buades *et al.* [20] presented a nonlocal means (NLM) approach, which uses image redundancy to estimate each pixel. The idea is to search similar pixels and average them with different weights expressing the similarity and obtained by patch correlation. Taking inspiration from NLM, Dabov *et al.* [21] proposed block-matching denoising based on 3-D transform (BM3D). With collaborative filtering as the core idea, it works on nonlocal regions (the group of patches collected from the whole image based on their similarity). Katkovnik *et al.* [22] proposed an extension of BM3D with pointwise-based shape-adaptive approach generalizing the advantages of both the BM3D filter and the Pointwise filter [23].

Deledalle *et al.* [24] extended the NLM to SAR image denoising. From a theoretical point of view, this probabilistic patch-based (PPB) algorithm [24] is a more general form of the original NLM in [20]. By considering the distribution of SAR data with the Goodman model [25], they suggest using a generalized likelihood ratio (GLR) to measure the similarity between SAR noisy patches. Analogously, Parrilli *et al.* [26] extended the BM3D for SAR image denoising and proposed a SAR-oriented version of the BM3D filter. It applies the similarity criterion proposed in [24] to group patches and then uses the local linear minimum-mean-square-error estimation.

Most studies on SAR image denoising are dealing with a single image. However, many temporal denoising approaches have been also explored, the simplest one being multilooking. Trouvé *et al.* [27] compared three filters of a multitemporal SAR image: a texture compensation multichannel filter [1] based on the Kuan filter, a time-space filter [2], and an extension of the maximum homogeneous region approach to multitemporal images [3]. Quegan and Yu [28] have a deep insight of the relationship between the desired equivalent number of looks (ENL) and the number of dates in multitemporal images and proposed a multichannel filter based on the intensity covariance matrix. Those denoising approaches of multitemporal images are local methods. Based on the NLM theory, this paper presents a more effective way to exploit the temporal information for multitemporal image denoising. We propose an adaptation of the iterative weighted PPB denoising algorithm [24]. First, a direct extension of PPB for multitemporal SAR images is analyzed to illustrate the limited efficiency of this

direct extension. Based on this analysis, we propose a two-step denoising framework, a temporal averaging step (the first step), and a spatial denoising step (the second step). This denoising framework is also applied to unregistered temporal SAR images by miss-registration estimation.

This paper is organized as follows: Section II briefly introduces the NLM and its extension for SAR images (PPB). A simple comparative experiment is presented in Section III to analyze the problem of the direct extension of PPB. The proposed method is then detailed in Section IV and extended to unregistered images in Section V. This is followed by evaluation in Section VI and conclusion in Section VII.

II. NLM

This section presents two denoising methods based on the NLM approach. They deal with Gaussian and multiplicative speckle noise, respectively.

We first introduce the notations that will be intensively used in the following.

- Bold \mathbf{y} : a patch with size of $\sqrt{K} \times \sqrt{K}$.
- Nonbold y : an image (y_t denotes the image at time t).
- Nonbold \tilde{y} : the temporal mean image.
- Nonbold $y_t(i)$: the i th pixel in image y_t ($y_t(i)$ is the center of patch $\mathbf{y}_t(i)$).
- For the sake of simplicity, \mathbf{y}_1 and \mathbf{y}_2 (with center pixels y_1 and y_2) denote two arbitrary patches of an image, whereas $\mathbf{y}_t(i)$ (with center pixel $y_t(i)$) denotes the i th patch in image y_t . $y_{1,k}$ is the k th pixel in patch \mathbf{y}_1 .
- \mathbf{u} , u and \mathbf{L} , L are similarly defined as representing, respectively, the true parameters and the number of looks. \hat{u} is the estimation of u .

A. NLM for Gaussian Distributed Data

We denote by y the observed image, by $y(i)$ the noisy intensity value at pixel index i , and by $\hat{u}(i)$ the estimation of the actual pixel value $u(i)$ (the *true* value that we are looking for). The NLM perform a weighted average (or weighted maximum-likelihood estimation) of the form

$$\hat{u}(i) = \frac{1}{Z} \sum_{j \in \Omega_i} \exp(-d_{\text{NLM}}(i, j)) \cdot y(j) \quad (1)$$

where Ω_i denotes the candidate pixels used for the estimation, and $Z = \sum_{j \in \Omega_i} \exp(-d_{\text{NLM}}(i, j))$ is a normalization. The weights $\exp(-d_{\text{NLM}}(i, j))$ are based on the weighted similarity between patches extracted at positions i and j , which are defined as

$$d_{\text{NLM}}(i, j) = \frac{S_{\text{NLM}}(\mathbf{y}(i), \mathbf{y}(j))}{h} \quad (2)$$

where $h > 0$ is a filtering parameter, and $\mathbf{y}(i)$ and $\mathbf{y}(j)$ are the two K -dimensional vectors containing the values of the $\sqrt{|K|} \times \sqrt{|K|}$ square patches centered, respectively, at pixel positions i and j (patches will be denoted by bold letters in the following). The similarity function S_{NLM} is chosen as the

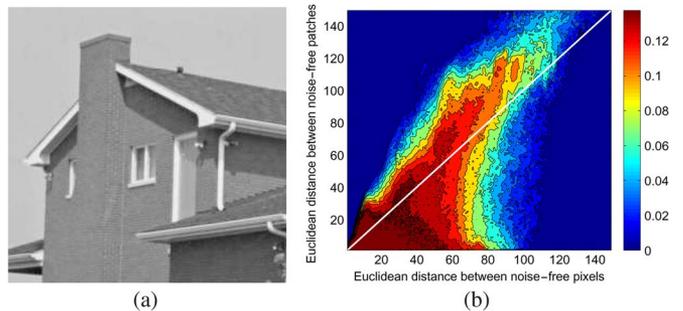


Fig. 1. Two-dimensional histogram of the Euclidean distance between noise-free patches K and noise-free pixels. The ideal case (the assumption that similar patches have similar central pixels [29]) is shown on the white line. (a) Test image *house*. (b) Two-dimensional histogram of the Euclidean distance.

square of the Euclidean distance defined for any pair of patches \mathbf{y}_1 and \mathbf{y}_2 by

$$S_{\text{NLM}}(\mathbf{y}_1, \mathbf{y}_2) = \|\mathbf{y}_1 - \mathbf{y}_2\|^2 = \sum_{k \in K} [y_{1,k} - y_{2,k}]^2. \quad (3)$$

The exponential function transforming similarity into weights is called the kernel.

The pixel similarity $S_{\text{NLM}}(\mathbf{y}(i), \mathbf{y}(j))$ is comparing two small square patches of size K surrounding pixels at positions i and j . $y_{1,k}$ is the value of pixel k in patch \mathbf{y}_1 . This definition is based on the assumption that similar patches have similar central pixels [20], [29]. Fig. 1 shows the 2-D histogram of the Euclidean distance between noise-free pixels $|u(i) - u(j)|$ and the Euclidean distance between noise-free patches $\|\mathbf{u}(i) - \mathbf{u}(j)\|$ measured from the *house* image. From this histogram, an approximate proportional relation between $|u(i) - u(j)|$ and $\|\mathbf{u}(i) - \mathbf{u}(j)\|$ can be found. When Gaussian noise is considered, the similarity between patches is usually defined by the Euclidean distance, as in (3).

B. NLM for Gamma Distributed Data

For the denoising of a single SAR image, Deledalle *et al.* [24] extended the NLM with PPB weights. Since this approach was proposed for the denoising of a single image only, this method will be denoted as 1-PPB. In order to deal with the specific nature of speckle noise as described in [25], they suggest using a similarity criterion based on the GLR for patches in SAR images (more details in [30]), leading to

$$S_{\text{GLR}}(\mathbf{y}_1, \mathbf{y}_2) = \sum_{k \in K} \log \left[\sqrt{\frac{y_{1,k}}{y_{2,k}}} + \sqrt{\frac{y_{2,k}}{y_{1,k}}} \right] \quad (4)$$

where \mathbf{y}_1 and \mathbf{y}_2 are two patches of intensity values. In addition, a refining term is added to the weight $\exp(-d_{1\text{-PPB}}(i, j))$, i.e.,

$$d_{1\text{-PPB}}(i, j) = \frac{S_{\text{GLR}}(\mathbf{y}(i), \mathbf{y}(j))}{h} + \frac{S_{\text{KL}}(\hat{\mathbf{u}}(i), \hat{\mathbf{u}}(j))}{h'} \quad (5)$$

where h and $h' > 0$ are two filtering parameters, and $\hat{\mathbf{u}}(i)$ and $\hat{\mathbf{u}}(j)$ are patches that are iteratively estimated. For the first iteration, $\hat{\mathbf{u}}(i) = 1$, and then, the current estimate is used for $\hat{\mathbf{u}}(i)$. The similarity function S_{KL} used to iteratively refine

the weights is based on the symmetric Kullback–Leibler (KL) divergence defined for any pair of patches \mathbf{u}_1 and \mathbf{u}_2 as

$$S_{\text{KL}}(\mathbf{u}_1, \mathbf{u}_2) = \sum_{k \in K} \frac{[u_{1,k} - u_{2,k}]^2}{u_{1,k}u_{2,k}}. \quad (6)$$

Some papers reference this approach as IT-PPB, but in this paper, 1-PPB stands for this iterative version.

C. Discussion

Compared with the local approaches, NLM has no connection restriction, since those similar pixels may be unconnected to each other. By exploiting patch similarities, the selection of similar pixels is naturally adapted to nonstationarities. For instance, for pixels located on one side of an edge, the weights are all concentrated on the same side of the transition, and the subsequent nonlocal estimation is thus unbiased, while local filters mix samples arising from both sides. A deeper comparison of NLM with local methods, such as intensity-driven adaptive neighborhood, has been done in [31]. In this paper, our contribution mainly focuses on the introduction of temporal information for denoising.

Our aim in this paper is to exploit the efficiency of nonlocal denoising for multitemporal series. We will first consider that images are finely registered and then relax this hypothesis in Section V. Our framework can be applied to interferometric series, but temporal denoising efficiency decreases with the spectrum correlation of the interferometric data.

III. DIRECT EXTENSION AND ANALYSIS

A direct extension of PPB denoising of SAR images is presented in this section. We tested it on a synthetic set of multitemporal SAR images. This comparison experiment will lead us to the approach proposed in the next section.

A. Direct Extension of PPB

The key idea of the PPB denoising in [24] similar to the NLM is to estimate actual pixel intensity with image redundancy. The way to exploit image redundancy is the search of similar pixels in a search window and averaging those similar pixels with different weights. For multitemporal data, a direct extension of PPB can be the definition of a cube search window, by aggregating all the search windows of the different dates. Considering $\{y_{t_1}, y_{t_2}, \dots, y_{t_N}\}$, the stack of multitemporal images, the temporal NLM filter can be defined on the cube $\mathbb{C}_i = \Omega_i \times \{t_1, t_2, \dots, t_N\}$ of all pixel indexes i . Meanwhile, the direct temporal extension of PPB (T-PPB) to estimate the true value $u_t(i)$ is

$$\hat{u}_t(i) = \frac{1}{Z} \sum_{(j,t') \in \mathbb{C}_i} \exp(-d_{\text{T-PPB}}(i_t, j_{t'})) \cdot y_{t'}(j) \quad (7)$$

where

$$d_{\text{T-PPB}}(i_t, j_{t'}) = \frac{S_{\text{GLR}}(\mathbf{y}_t(i), \mathbf{y}_{t'}(j))}{h_T} + \frac{S_{\text{KL}}(\hat{\mathbf{u}}_t(i), \hat{\mathbf{u}}_{t'}(j))}{h'_T}$$

and for a stack of images $\{y_{t_1}, \dots, y_{t_N}\}$, $y_t(i)$ and $\mathbf{y}_t(i)$ are, respectively, the i th pixel value of y_t and its surrounding patch.

B. Comparison

To analyze the interest of this direct extension for the temporal case, we have tested a synthetic set of multitemporal images $\{y_{t_1}, y_{t_2}, y_{t_3}\}$. These images are synthesized supposing there is no change in time (stable case), thus giving three realizations of the same scene. They have been denoised using three methods.

- **Method i (1-PPB)**: Denoise the single image y_{t_1} without using images y_{t_2} and y_{t_3} by the PPB approach.
- **Method ii (T-PPB)**: Denoise y_{t_1} using the set of temporal images $\{y_{t_1}, y_{t_2}, y_{t_3}\}$ by the temporal PPB approach [see (7)].
- **Method iii (M-PPB)**: First, get the mean values \bar{y} of the set of multitemporal images $\{y_{t_1}, y_{t_2}, y_{t_3}\}$ and then denoise this average image \bar{y} by using the PPB approach.

Note that method **i** uses only one image, whereas method **ii** uses the whole image set (three images). Although methods **ii** and **iii** share the same input noisy images, method **iii** first temporally averages the three temporal images. Because the denoising approach of the three methods is PPB and the main difference among them is the input, one can easily predict that method **i** has the poorest results and that methods **ii** and **iii** should have comparable performance.

Fig. 2 shows the comparison results. Indeed, by using three 1-look noisy images, the denoising performance of T-PPB in Fig. 2(d) has improved compared with the result of 1-PPB in Fig. 2(b) (seeing the SNR values). However, it is not sufficient, because M-PPB in Fig. 2(f) has significantly outperformed T-PPB, in spite of the same input noisy images.

C. Analysis

In order to analyze these results, we represent the map of weights $w(i_t, j_{t'})$. Fig. 3 shows the weight maps of several interesting pixels for the three methods. In the noise-free image, we can easily find the similar pixels. However, in method **ii**, more dissimilar pixels have been assigned large weights (bright points in the weight map) than those in method **iii**. More weights are computed, but they are less accurate. Due to the temporal average before denoising, method **iii** reduces the risk in searching similar pixels.

A first solution could be the modification of the kernel that transforms similarity in weights. Indeed, it is shown in [32] that truncated weights for NLM (instead of exponential function) could improve the selection. However, for a 1-look image, the improvement is not sufficient.

Averaging in the temporal domain (multilooking) is considered as the efficient and best unbiased estimator for the hypotheses of independent and identical distribution. Thus, method **iii** naturally performs better than methods **i** and **ii**. The purpose of this experiment is to illustrate the improvement of the introduction of multilooking into PPB. In this paper, we combine the advantages of both approaches (multilooking and PPB) to achieve a multitemporal denoising method.

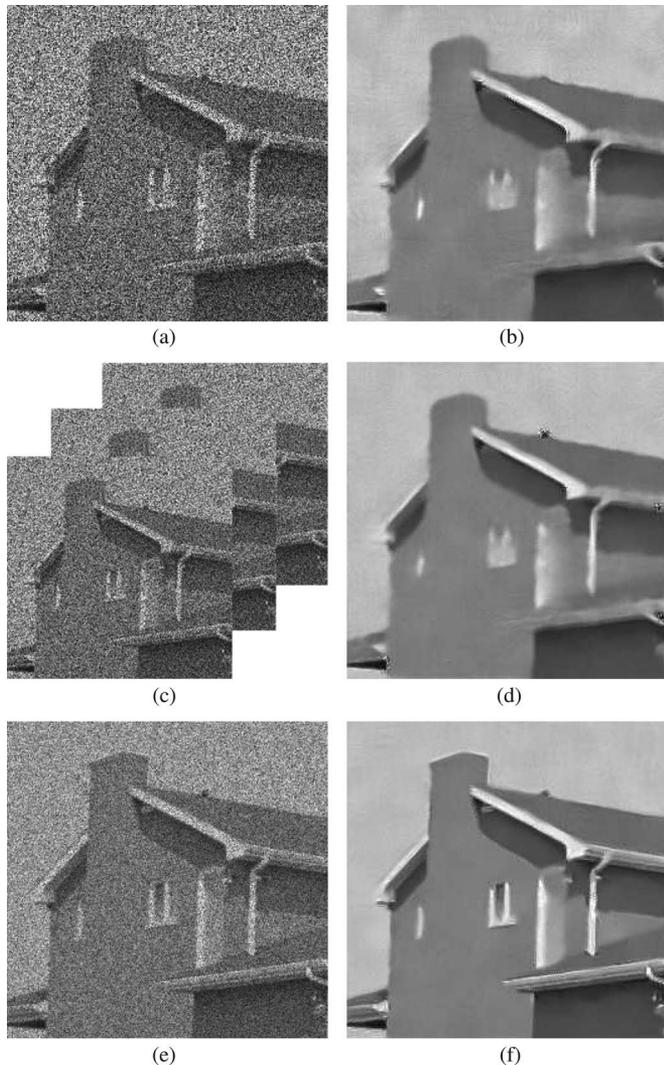


Fig. 2. Denoising results of synthetic images with multiplicative speckle noise (left: input noisy images; right: denoising results). (a) 1-look noisy image. (b) Method **i** 1-PPB result of (a); SNR: 10.79. (c) Three 1-look noisy multitemporal images. (d) Method **ii** T-PPB result of (c); SNR: 11.62. (e) Temporal mean of (c) (a 3-look noisy image). (f) Method **iii** M-PPB result of (e); SNR: 13.81.

IV. TWO-STEP MULTITEMPORAL NLM

Let us go back to the simple comparison experiment in Section III-B. Method **iii** [in Fig. 2(e) and (f)] first temporally averages noisy images, which reduces the noise level and improves the weights, as shown in Fig. 3. Although the real SAR images or video data have problems such as registration (or motion) and temporal changes, there are still lots of cases that can be temporally combined. Taking inspiration from method **iii**, we divide the denoising process into two steps, which deal with temporal and spatial information, respectively. The main purpose of multitemporal denoising is to exploit all available information for temporally stable pixels while keeping the new information as much as possible. Note that in this part, all multiimages are well registered, and only the change detection problem is considered in this section.

In case of stable pixels (no change over time and well registered), method **iii** illustrates the usefulness of temporal

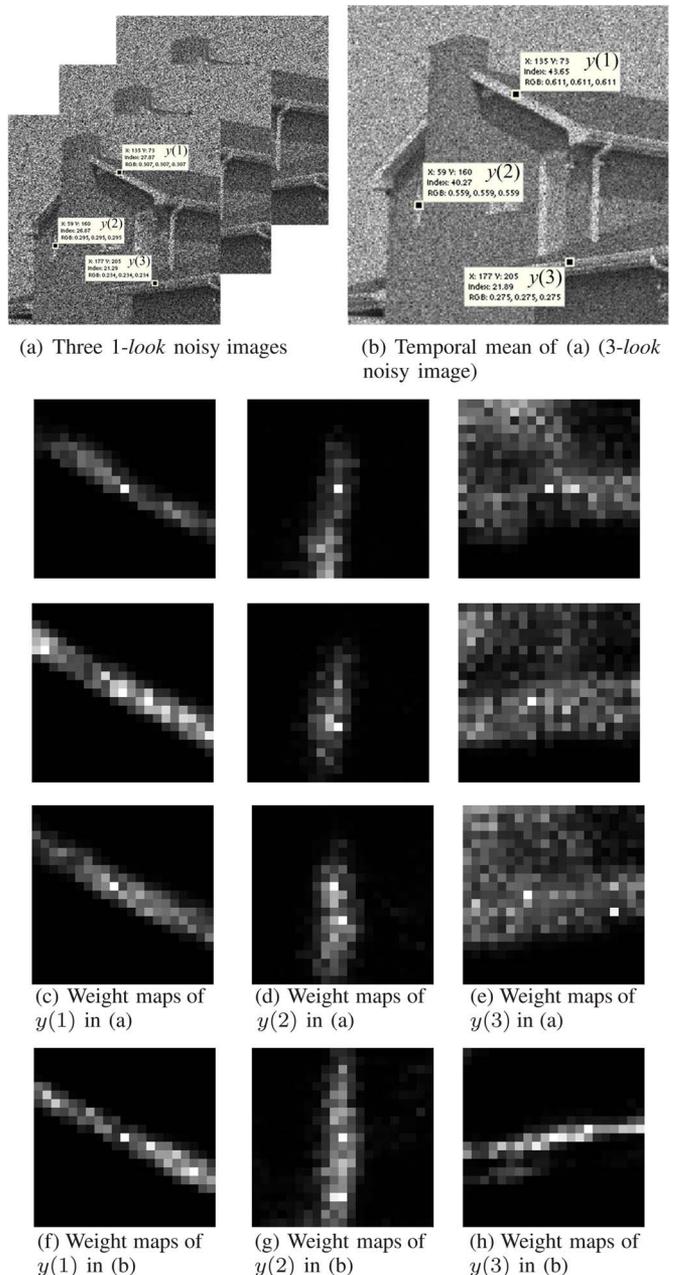


Fig. 3. Weight maps. (a) Three 1-look noisy images. (b) Temporal mean of (a). (c), (d), and (e) Weight maps of $y(1)$, $y(2)$, and $y(3)$ in (a). (f), (g), and (h) Weight maps of $y(1)$, $y(2)$, and $y(3)$ in (b).

averaging. The proposed denoising framework, namely, a two-step (2S-PPB) approach based on NLM and PPB, is to exploit similar pixels in the temporally average image \bar{y} rather than in the stack $\{y_{t_1}, \dots, y_{t_N}\}$ (the premise is that there is no change taking place among multiimages). If pixels located at the same position but at different times (such as $y_t(i)$ and $y_{t'}(i)$) have not changed (in other words, they share a same true value $u_t(i) = u_{t'}(i)$), they can be averaged together to estimate $u_t(i)$. From a probability point of view, this equally weighted average can be considered as an estimation using prior information (relative to the estimate using likelihood information in [24]). This temporal average can be seen as a preliminary multilooking operation. However, the main problem lying in

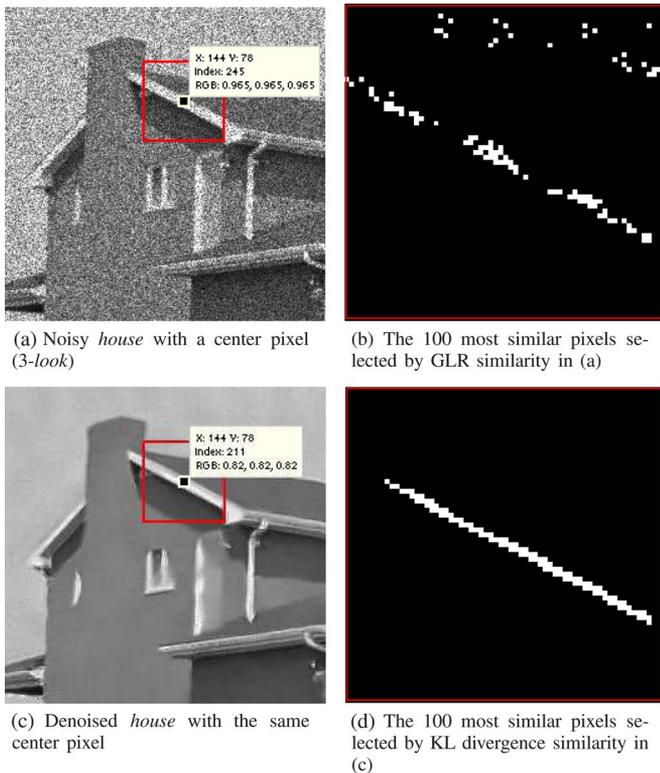


Fig. 4. Selected similar pixels by the GLR criterion and the KL divergence criterion. (a) The 3-look noisy *house* image and the pixel that will be used to illustrate the distributions; the red rectangle is the search window. (b) Selections by the GLR criterion. (c) PPB denoising result of (a). (d) Selections by the KL criterion. The KL criterion in denoised images is more capable of selecting similar pixels than the GLR criterion in noisy images.

this temporal denoising is the use of only stable pixels. Thus, we have to use change detection, which will be detailed in the following.

A. Change Criteria

Most change detection methods are concerned about changes between different terrestrial objects, such as *rivers*, *buildings*, and other artificial objects. However, in image denoising, we theoretically combine the observed pixels from the same true value or reflectivity (coming from the same distribution). Hence, the *changes* in this section are defined as samples coming from different distributions.

Several similarity criteria for noisy patches are compared in [30]. The GLR is proposed for non-Gaussian noise. This criterion can be used to detect changes in times. Another criterion, i.e., the symmetric KL divergence criterion (hereafter referred to as the KL criterion) in denoised images (using 1-PPB on each image), is proposed in [24] to improve the performance of the GLR criterion computed in noisy images. Both criteria, GLR on noisy images and KL divergence on denoised data, have been investigated for similar pixel selection.

Fig. 4 illustrates the performance of the GLR criterion in the noisy image and the KL criterion in the denoised image (using 1-PPB on each image) to find the similar pixels. The results show that the KL criterion in denoised images is more capable of selecting similar pixels than the GLR criterion in

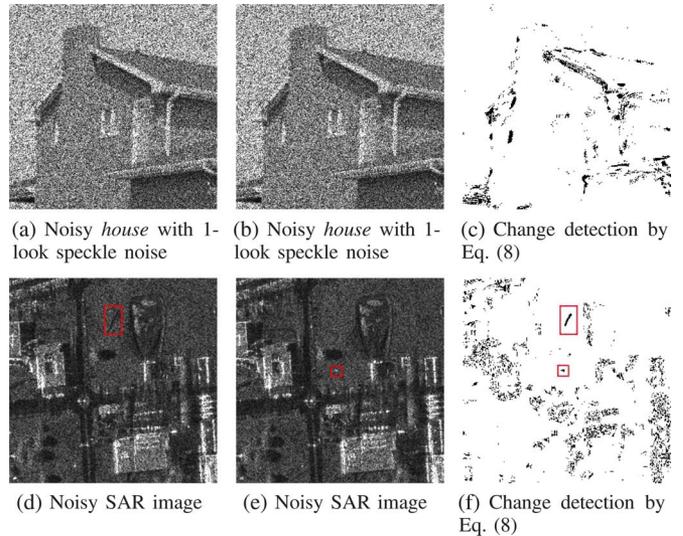


Fig. 5. Change detection by the KL divergence and the GLR criterion. (a) and (b) Noisy *house* images with 1-look speckle noise. (c) Change detection results between (a) and (b). (d) and (e) Noisy SAR images; the red rectangles denote the changes between (d) and (e). (f) Change detection results between (d) and (e). Changes in red rectangles have been found out in detection results.

the noisy image. However, this comparison experiment is not fair, because of the different noise levels in Fig. 4(a) and (c). What is more, the KL criterion in denoised images has an inevitable drawback that the used denoising approaches have great influence on the quality of the selections.

To balance the advantages against the disadvantages of both the GLR in noisy images and the KL criterion in denoised images, we suggest employing both of them to detect the changes between temporal images (see Fig. 5). This suggestion has the similar consideration as in [24], which proposed that the denoising weights are coming from both the noisy images (by the GLR criterion) and the previous denoised images (by KL divergence).

For the sake of simplicity, we use (8) as a binary criterion to define the temporal relation between the pixel values at i_t and the pixel value at $j_{t'}$ as

$$\varphi[d_{1st}(i_t, j_{t'})] = \begin{cases} 1, & \text{if } d_{1st}(i_t, j_{t'}) < 2 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

$$d_{1st}(i_t, j_{t'}) = \frac{S_{GLR}(\mathbf{y}_t(i), \mathbf{y}_{t'}(j))}{h_{1st}} + \frac{S_{KL}(\hat{\mathbf{u}}_t(i), \hat{\mathbf{u}}_{t'}(j))}{h'_{1st}}$$

where $\hat{\mathbf{u}}_t(i)$ and $\hat{\mathbf{u}}_{t'}(j)$ are patches extracted in the denoised result, respectively, in \hat{u}_t^i and $\hat{u}_{t'}^j$ using 1-PPB. In order to ensure that both GLR and KL criteria have the same contribution on change detection, we normalize the GLR and KL terms in (8) by parameters $h_{1st} = \text{quantile}(S_{GLR}, \alpha)$ and $h'_{1st} = \text{quantile}(S_{KL}, \alpha)$, respectively, where for any similarity measure f , $\text{quantile}(f, \alpha)$ denotes the α -quantile of the *pure* distribution of f (i.e., the distribution when patches have the same underlying reflectivity). In practice, we have chosen $\alpha = 0.99$, such that a binary weight $\varphi[d_{1st}(i_t, j_{t'})] = 1$ means that pixel $y_t(i)$ and $y_{t'}(j)$ have a high probability to be realizations coming from the same underlying reflectivity. If α keeps constant, a threshold larger than “2” will add more changed pixels

in the temporal step. On the contrary, a smaller threshold may neglect some similar (unchanged) pixels in the temporal step.

B. Two-Step Denoising

The following formula shows the temporal averaging process with the change criterion (binary weights):

$$\tilde{y}_t(i) = \frac{1}{Z} \sum_{t' \in [t_1, t_N]} \varphi[d_{1st}(i_t, i_{t'})] \cdot y_{t'}(i) \quad (9)$$

where $Z = \sum_{t' \in [t_1, t_N]} \varphi[d_{1st}(i_t, i_{t'})]$.

Denoising on the temporal estimate $\tilde{y}_t(i)$ is comparable to (7). However, different pixels $\tilde{y}_t(i)$ may have different (equivalent) number of *looks* depending on the number of averaged data. The resulting number of *looks* is temporally and spatially varying and is given by

$$\tilde{L}_t(i) = \sum_{t' \in [t_1, t_N]} \varphi[d_{1st}(i_t, i_{t'})] \cdot L_{t'} \quad (10)$$

in which $L_{t'}$ is the original spatially invariant (equivalent) number of *looks* of the image $y_{t'}$. In this case, the similarity between patches $\tilde{\mathbf{y}}_t(i)$ and $\tilde{\mathbf{y}}_t(j)$ has to be modified to take into account the spatially varying number of *looks*. For any patches $\tilde{\mathbf{y}}_1$ and $\tilde{\mathbf{y}}_2$, the GLR gives (see Appendix A)

$$S_{GLR}(\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2) = \sum_{k \in K} \left[(\tilde{L}_{1,k} + \tilde{L}_{2,k}) \log[\tilde{L}_{1,k} \tilde{y}_{1,k} + \tilde{L}_{2,k} \tilde{y}_{2,k}] \right. \\ \left. - (\tilde{L}_{1,k} + \tilde{L}_{2,k}) \log(\tilde{L}_{1,k} + \tilde{L}_{2,k}) \right. \\ \left. - \tilde{L}_{1,k} \log[\tilde{y}_{1,k}] - \tilde{L}_{2,k} \log[\tilde{y}_{2,k}] \right] \quad (11)$$

where \tilde{L}_1 and \tilde{L}_2 are, respectively, the (equivalent) number of *looks* of patches $\tilde{\mathbf{y}}_1$ and $\tilde{\mathbf{y}}_2$. Then, the estimate $\hat{u}_t(i)$ using \tilde{y}_t will be

$$\hat{u}_t(i) = \frac{1}{Z} \sum_{j \in \Omega_i} \exp(-d_{2nd}(i_t, j_t)) \cdot \tilde{y}_t(j) \quad (12)$$

where

$$d_{2nd}(i_t, j_t) = \frac{S_{GLR}(\tilde{\mathbf{y}}_t(i), \tilde{\mathbf{y}}_t(j))}{h_{2nd}} + \frac{S_{KL}(\hat{\mathbf{u}}'_t(i), \hat{\mathbf{u}}'_t(j))}{h'_{2nd}}$$

and S_{KL} plays the same role as in (7) for iteratively refining the weights by using patches $\hat{\mathbf{u}}(i)$ and $\hat{\mathbf{u}}(j)$ extracted from the previous estimate $\hat{\mathbf{u}}'_t$. However, by considering the different number of looks, the similarity function S_{KL} has also to be modified according to the KL divergence, where for any pair of patches \mathbf{u}_1 and \mathbf{u}_2 (see Appendix B for details)

$$S_{KL}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) = \sum_{k \in K} \tilde{L}_{1,k} \frac{\tilde{u}_{2,k}}{\tilde{u}_{1,k}} + \tilde{L}_{2,k} \frac{\tilde{u}_{1,k}}{\tilde{u}_{2,k}} - \tilde{L}_{1,k} - \tilde{L}_{2,k} \\ + (\tilde{L}_{1,k} - \tilde{L}_{2,k}) \left[\psi(\tilde{L}_{1,k}) - \psi(\tilde{L}_{2,k}) + \ln(\tilde{u}_{1,k}) - \ln(\tilde{u}_{2,k}) \right]. \quad (13)$$

Algorithm 1 summarizes the steps to denoise temporal SAR images by the proposed 2S-PPB method.

Algorithm 1 The proposed 2S-PPB algorithm.

Input:

Well-registered temporal SAR images $\{y_{t_1}, y_{t_2}, \dots, y_{t_N}\}$.
A date t_1 of interest.

Output:

\hat{u}_{t_1} : the denoising result of image y_{t_1} .

———— **Step 1 (Temporal step):** ————

- 1: **for** each y_t in $\{y_{t_1}, y_{t_2}, \dots, y_{t_N}\}$ **do**
- 2: denoise y_t using the 1-PPB approach;
 (see Section II-B)
- 3: obtain predenoised results \hat{u}'_t ;
- 4: **for** each pixel index i **do**
- 5: compute change criterion $\varphi[d_{1st}(i_{t_1}, i_t)]$; [see (8)]
- 6: **end for**
- 7: **end for**
- 8: **for** each pixel index i **do**
- 9: initialize $\tilde{y}_{t_1}(i) = 0, \tilde{L}_{t_1}(i) = 0$;
- 10: **for** each y_t in $\{y_{t_1}, y_{t_2}, \dots, y_{t_N}\}$ **do**
- 11: $\tilde{y}_{t_1}(i) = \tilde{y}_{t_1}(i) + \varphi[d_{1st}(i_{t_1}, i_t)] \cdot y_t(i)$; [see (9)]
- 12: $\tilde{L}_{t_1}(i) = \tilde{L}_{t_1}(i) + \varphi[d_{1st}(i_{t_1}, i_t)] \cdot L_t$
- 13: **end for**
- 14: $\tilde{y}_{t_1}(i) = \tilde{y}_{t_1}(i) / \tilde{L}_{t_1}(i)$; [see (9)]
- 15: **end for**

———— **Step 2 (Spatial step):** ————

- 16: denoise $\tilde{y}_{t_1}(i)$ using (11), (12), and (13) in Section IV-B.
This denoising step is similar to the PPB approach.
 - 17: Denoised result \hat{u}_{t_1} ;
-

V. TWO-STEP MULTITEMPORAL NLM FOR UNREGISTERED SAR IMAGES

The two-step multitemporal NLM proposed in the previous section deals with multitemporal SAR images that have been well registered. However, in practice, it is not always easy to get well-registered images because of the lack of accuracy of sensor parameters or of terrain deviation. Here, we propose an adaptation of the two-step multitemporal NLM to deal with unregistered temporal SAR images.

A. Miss-Registration Estimation

From the simple comparison experiment in Section III-B, we have seen that the direct extension of 2-D denoising approaches to well-registered temporal SAR images is not optimal. Hence, unlike in [33], we try to consider the *offset* caused by miss-registration between temporal SAR images using patch similarity. This offset between remote sensing SAR images caused by miss-registration is much simpler than the complex scene changes in video. Moreover, our aim is not accurate registration but only accurate detection of similar pixels to perform temporal averaging.

Let y_t and $y_{t'}$ denote two temporal images without registration, and $y_t(i)$ and $y_{t'}(i)$ are the i th pixel in y_t and $y_{t'}$, respectively. Note that $y_t(i)$ and $y_{t'}(i)$ are both located at i in

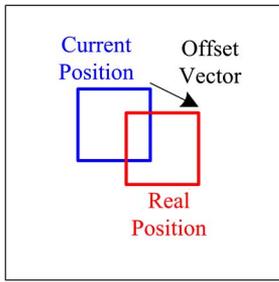


Fig. 6. Assumption of image registration. In the intensity image registration, all pixels in a small patch share the same offset to their real position.

images, but they do not denote the same position in the geographic coordinate system before image registration. Suppose that $y_{t'}(i + \vec{v}_{tt'}(i))$ is the $[i + \vec{v}_{tt'}(i)]$ th pixel in $y_{t'}$ that shares the same geographic location with $y_t(i)$ in y_t , and $\vec{v}_{tt'}(i)$ denotes the offset between $y_t(i)$ and $y_{t'}(i + \vec{v}_{tt'}(i))$. If no change takes place, $y_{t'}(i + \vec{v}_{tt'}(i))$ should be the most similar pixel of $y_t(i)$ in image $y_{t'}$. Consequently, the idea to estimate this offset $\vec{v}_{tt'}(i)$ caused by miss-registration is to use the pixel similarity.

Using the conclusion in Section IV-A, we combine the KL divergence and the GLR criterion as the similarity between patches, illustrated in (8). (Note that this KL divergence criterion is from the 1-PPB denoising results on each image.) For each pixel $y_t(i)$ in y_t , we measure the similarity between $y_t(i)$ and the pixels in region $A_{t'}(i)$ (an image patch with $y_{t'}(i)$ as the center in $y_{t'}$, shown in Fig. 7). The similarity between $y_t(i)$ and $y_{t'}(j)$ is $d_{1st}(i_t, j_{t'})$. Each pair of pixel $y_t(i+n)$ in y_t and region $A_{t'}(i+n)$ in $y_{t'}$ has a similarity map $M_{tt'}(i+n)$ ($0 \leq n \leq N_A - 1$, N_A is the size of region A), as shown in the top-right image in Fig. 7. Thus

$$M_{tt'}(i) = \{d_{1st}(i_t, j_{t'})\}_{j_{t'} \in A_{t'}(i)}. \quad (14)$$

The offset $\vec{v}_{tt'}(i)$ should be the offset between the center of map $M_{tt'}(i)$ and the position of minimum value in $M_{tt'}(i)$. However, it leads to a poor performance because of noise. Based on the assumption in Fig. 6, we suggest that the offset $\vec{v}_{tt'}(i)$ caused by miss-registration can be estimated from an average similarity map $\bar{M}_{tt'}(i)$, which is the mean of similarity maps $M_{tt'}(i+n)$ ($1 \leq n \leq N_A$). The estimation of offset $\vec{v}_{tt'}(i)$ is from the center of map $\bar{M}_{tt'}(i)$ to the position of minimum value in $\bar{M}_{tt'}(i)$, i.e.,

$$\begin{aligned} \bar{M}_{tt'}(i) &= \{\bar{d}_{1st}(i_t, j_{t'})\}_{j_{t'} \in A_{t'}(i)} \\ \bar{d}_{1st}(i_t, j_{t'}) &= \sum_{i_t \in A_t(i)} d_{1st}(i_t, j_{t'}). \end{aligned} \quad (15)$$

We test the miss-registration estimation on real SAR images. Fig. 14(b) is the temporal average without image registration or miss-registration estimation, which is blurred. After miss-registration estimation, the temporal average is illustrated in Fig. 14(c).

B. Two-Step Denoising

The miss-registration estimation entitles the proposed 2S-PPB denoising approach to deal with unregistered multitem-

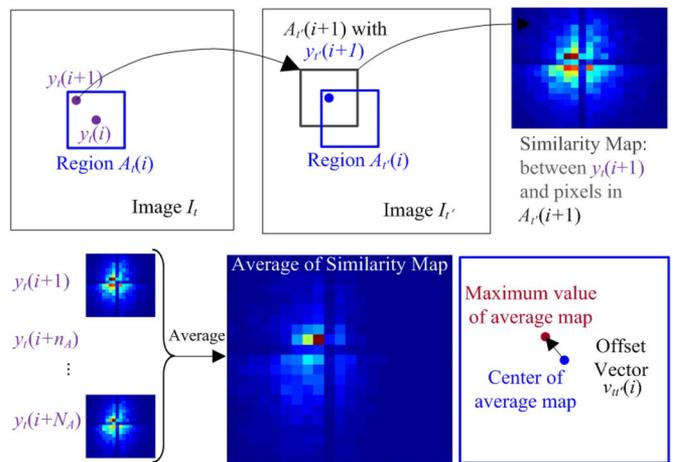


Fig. 7. Sketch map of the miss-registration estimation for unregistered images.

poral images. The denoising process is exactly the same as detailed in Section IV, except that the candidate pixels $y_{t'}(i)$ are replaced by $y_{t'}(i + \vec{v}_{tt'}(i))$ found out by the miss-registration estimation (see Fig. 7). Algorithm 2 summarizes the steps to denoise unregistered temporal SAR images by the proposed 2S-PPB method.

Algorithm 2 The proposed 2S-PPB algorithm (unregistered temporal images).

Input:

Unregistered temporal SAR images $\{y_{t_1}, y_{t_2}, \dots, y_{t_N}\}$.
A date t_1 of interest.

Output:

\hat{u}_{t_1} : the denoising results of image y_{t_1} .

———— **Miss-Registration estimation:** ————

- 1: **for** each y_t in $\{y_{t_1}, y_{t_2}, \dots, y_{t_N}\}$ **do**
- 2: denoise y_t using the 1-PPB approach;
 (see Section II-B)
- 3: obtain predenoised results \hat{u}'_t ;
- 4: **end for**
- 5: **for** each pixel index i **do**
- 6: **for** each image y_t in $\{y_{t_1}, y_{t_2}, \dots, y_{t_N}\}$ **do**
- 7: $\bar{M}_{t_1 t}(i) = 0$;
- 8: **for** each pixel index $i+n$ in region $A_{t_1}(i)$ **do**
- 9: compute similarity map $M_{t_1 t}(i+n)$; [see (15)]
- 10: $\bar{M}_{t_1 t}(i) = \bar{M}_{t_1 t}(i) + M_{t_1 t}(i+n)$;
- 11: **end for**
- 12: find the offset vector $\vec{v}_{t_1 t}(i)$ in similarity map $\bar{M}_{t_1 t}(i)$;
- 13: **end for**
- 14: **end for**

———— **Step 1 (Temporal step):** ————

- 15: compute \tilde{y}_{t_1} using (8) and (9) by taking into account the vector field $\vec{v}_{t_1 t}$. This denoising step is similar to the one in Algorithm 1 where i_t is substituted to $i_t + \vec{v}_{t_1 t}(i)$.
-

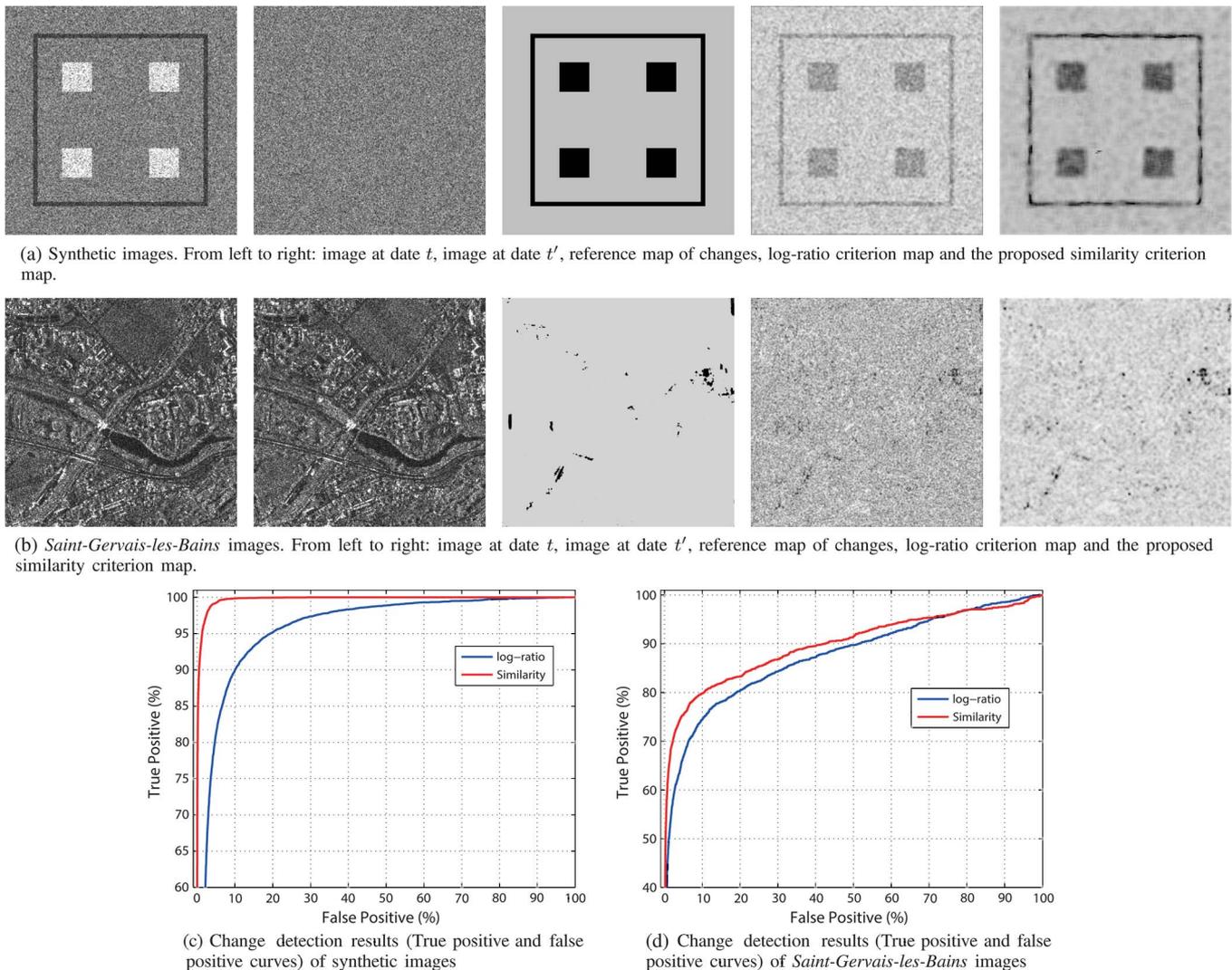


Fig. 8. Change detection criterion comparison. (a) Noisy image at date t . (b) Noisy image at date t' . (c) Reference map of changes. (d) Log-ratio change criterion. (e) Similarity map. (f) Change detection results (true-positive and false-positive curves).

Step 2 (Spatial step):

16: denoise $\tilde{y}_{t_1}(i)$ using (11), (12), and (13) in Section IV-B. This denoising step is similar to the PPB approach.

17: **return** Denoised result \hat{u}_{t_1} ;

VI. EXPERIMENTS

Here, we analyze the performance of the proposed approaches. Before denoising evaluation, the similarity criterion combining GLR and KL criteria used for change detection or offset estimation is studied.

A. Test of Pixel Similarity

The value of d_{1st} given by (8) is a key point of our approach, since it is used both for the change detection step and for the offset estimation step. We compare our change detection criterion with the log-ratio, which is a criterion widely used in SAR images. Synthetic images and real SAR images are used in these comparison experiments: 1) synthetic SAR images

corrupted by a multiplicative 1-look speckle noise in Fig. 8(a); 2) two real SAR images (TerraSAR images in Saint-Gervais-les-Bains, France) sensed in 2009 and 2011, respectively, in Fig. 8(b). Fig. 8(a) and (b) shows noisy images at different times. The reference maps of changes are shown in Fig. 8(c) (for the real SAR images, we manually label the changes taking place between date 1 and date 2). Fig. 8(e) shows our similarity-based change detection criterion, and Fig. 8(d) shows the log-ratio criterion. Our similarity-based change detection criterion has higher receiver operating characteristic curves in the false-positive and true-positive curves in Fig. 8(f).

We also compare our offset estimation with the offset estimation using the intensity tracking method. Intensity tracking with correlation has been widely used in glacier motion estimation tasks on multitemporal SAR images [34], [35]. In this comparison experiment, we simulate a multitemporal image set with an offset $\vec{v}_{tt'} = (3, 3)$ pixels between them, as shown in Fig. 9(a) and (b).

For the offset estimation using intensity tracking, it is self-implemented in our comparison experiments, and it also has a local averaging process similar to that in the proposed method

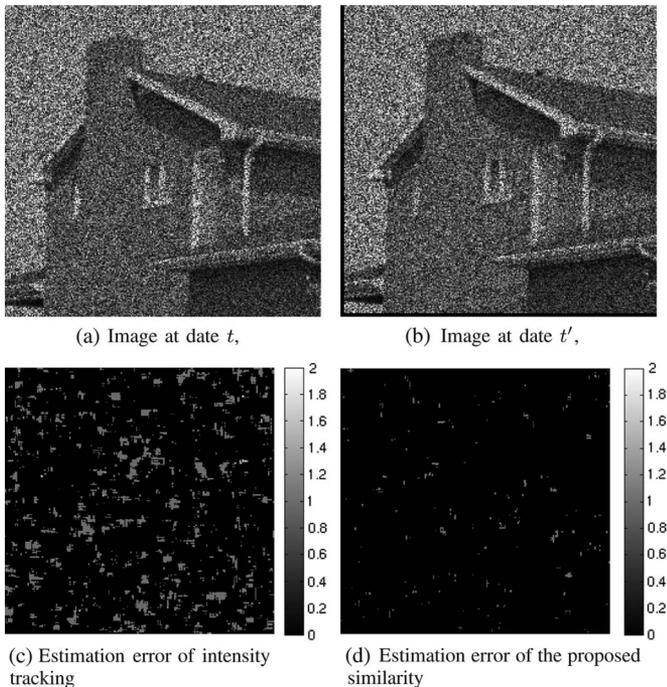


Fig. 9. Offset estimation. (a) Noisy image 1. (b) Noisy image 2 (with an offset $\vec{v}_{tt'} = (3, 3)$ with image 1). (c) Estimation error of intensity tracking [34], [35] (the distance between intensity tracking offset estimation results and the true offset $\vec{v}_{tt'} = (3, 3)$). (d) Estimation error of the proposed similarity (the distance between the proposed offset estimation results and the true offset $\vec{v}_{tt'} = (3, 3)$).

to reduce the effect of speckle noise. The patch size is 7×7 pixels, and the test area is 21×21 pixels. To illustrate the performances, the estimation error is calculated using the Euclidean distance between the true offset $\vec{v}_{tt'} = (3, 3)$ and the estimated offset. Fig. 9(c) and (d) shows that the proposed offset estimation has less estimation error than the log-ratio similarity.

B. Denoising of Well-Registered Images

For all the experiments, we use the parameters as suggested by Deledalle *et al.* in [24]. The search window Ω and patch size K enlarge with the increase in the number of iterations, $\Omega \in \{3 \times 3, 7 \times 7, 11 \times 11, 21 \times 21\}$ and $K \in \{1 \times 1, 3 \times 3, 5 \times 5, 7 \times 7\}$ for all experiments. h_{2nd} , controlling the decay of weights, depends on the distribution of similarity between two noisy patches with the same underlying reflectivity and is defined as the α -quantile of this *pure* distribution ($\alpha = 0.92$). $h'_{2nd} = 0.2|K|$.

The experiments are taken under the MATLAB environment on an Intel Core-2 Quad CPU Q9550 at 2.83-GHz 64-bit computer. The time consumption of the proposed method is about 400 s for a three-date 256×256 temporal image set (PPB needs about 15 s for a 256×256 image).

1) *Synthetic Images*: We present visual and numerical results obtained on synthetic images corrupted by multiplicative Goodman's speckle noise. The classical noise-free images are used: *house*, *lena*, *barbara*, *boat*, and *peppers*. We use the same noise-free image to synthesize a temporal image set, which means that there are no temporal changes. In addition to the proposed 2S-PPB filter, the comparisons that have been tested here are PPB only on single image y_{t_1} (method **i**: 1-PPB), 3-D

adaptive neighborhood filter (3-D-ANF, it is self-implemented) [3] on the multitemporal image set $\{y_{t_1}, \dots, y_{t_N}\}$, and temporal PPB on the multitemporal image set $\{y_{t_1}, \dots, y_{t_N}\}$ (method **ii**: T-PPB). For the 3-D-ANF filter, we use a 3×3 median filter in the first step of 3-D-ANF and 50×50 as the maximum size of adaptive neighborhood in the last step of 3D-ANF.

Fig. 10 only shows the denoising results of the *house* images corrupted by $L = 1$ multiplicative speckle noise. There are three noisy images in the temporal data set. The image obtained by the T-PPB filter is well smoothed compared with the 1-PPB filter. However, the edge and shape preservation has limited improvement. The proposed 2S-PPB filter provides more details of edges as the eaves and windows of *house*, whereas smooth regions are comparable. 3D-ANF has less loss of structural information in stable cases [shown in the ratio map in Fig. 10(b)], whereas it has poor noise reduction in homogeneous regions.

To quantify the denoising qualities, Table I presents numerical results for images corrupted by multiplicative speckle noise with different ENL $L = 1, 3, 5$, and 10 and different number of dates (different number of images in the temporal data set) $N = 1, 2, 3$, and 5. Note that when $N = 1$, there are no T-PPB and 2S-PPB denoising results. The used performance criterion is the signal-to-noise ratio (SNR), i.e.,

$$\text{SNR}(\hat{u}_t, u_t) = 10 \log_{10} \frac{\text{Var}[u_t]}{\text{Mean}[(\hat{u}_t - u_t)^2]}. \quad (16)$$

We observe that the 2S-PPB filter improves on the T-PPB filter for high-noise-level images (i.e., $L = 1$), particularly when N is large. However, for low-noise-level images (i.e., $L > 5$), 2S-PPB has only limited improvement. This is because the similarity of noisy patches in low-noise-level images is efficient enough, and the improvement of patch similarity using temporal average is relatively less important.

2) *Realistic SAR Synthetic Images*: This part presents denoising results of realistic SAR synthetic images. It is a 100-look SAR acquisition identified as *Toulouse* of the CNES in Toulouse (France) sensed by RAMSES and provided by the CNES. We corrupt this 100-look *Toulouse* image by 1-look multiplicative speckle noise to form three temporal images $\{y_{t_1}, y_{t_2}, y_{t_3}\}$. In order to simulate the changes in the multitemporal images, a dark *line* and a bright *target* are added to y_{t_1} , labeled by red rectangles (Region #1 and #2) in Fig. 12(a). y_{t_2} and y_{t_3} are corrupted by different multiplicative speckle noise without the dark *line* and the bright *target* [see Fig. 12(e) and (f)].

Fig. 12 presents the obtained denoising results for the *Toulouse* images. The results of the proposed 2S-PPB filter have better edge and shape preservation with comparable smoothed regions than 1-PPB and T-PPB filters (ENL in Table II). Structural information of new objects (changes) is also better preserved, as shown in Fig. 12(a) (Region #1 and #2). Stable objects, as the dark lines in Region #3 and #4, have been better restored in the results of 2S-PPB than 1-PPB and T-PPB filters.

3) *Real Multitemporal SAR Images*: We also test on real multitemporal SAR data as follows:

- *San Francisco* (well-registered) IGARSS: six single-look TerraSAR images of San Francisco (USA) provided by

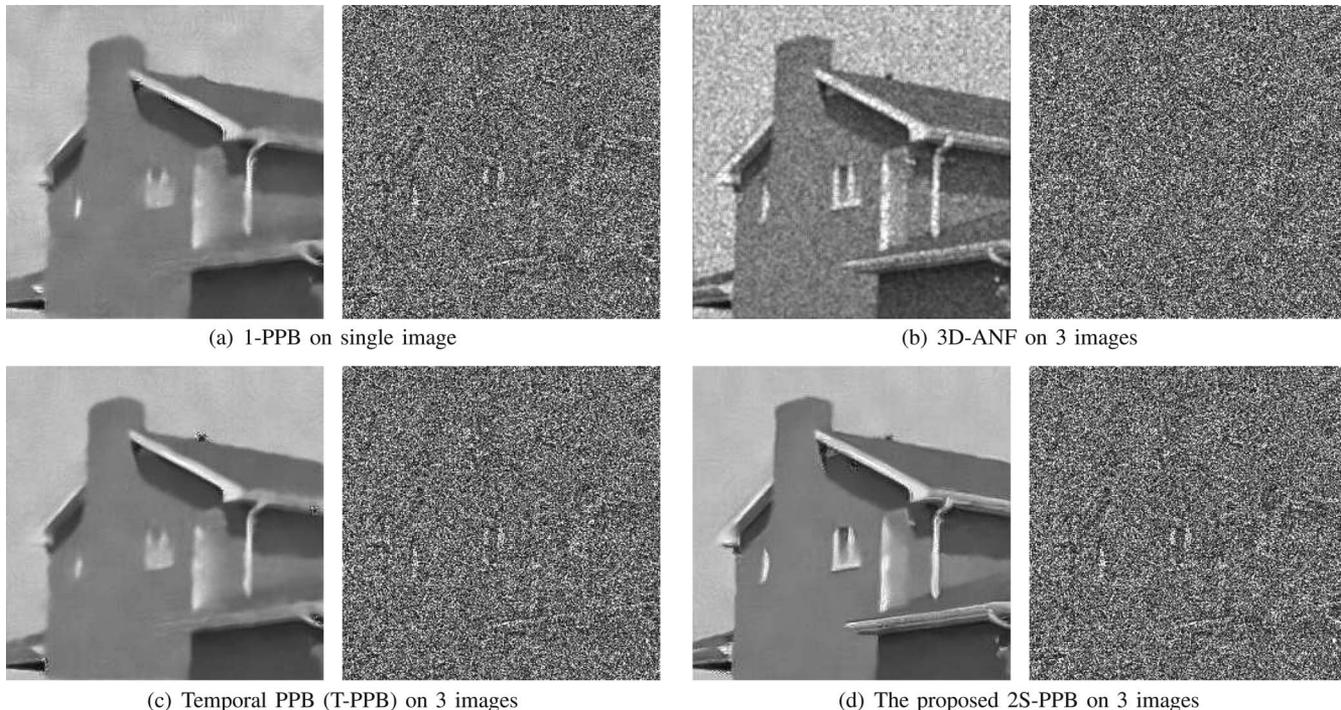


Fig. 10. Denoising results for the *house* images corrupted by multiplicative speckle noise with $L = 1$. From left to right: the denoising results and the ratio of the noisy image to the denoised image. (a) 1-PPB filter. (b) 3D-ANF. (c) T-PPB filter. (d) Proposed 2S-PPB filter.

TABLE I
SNR VALUE OF ESTIMATED IMAGES USING 1-PPB (BROWN), 3D-ANF (BLUE), T-PPB (GREEN), AND 2S-PPB (RED) FILTER FOR IMAGES CORRUPTED BY MULTIPLICATIVE SPECKLE NOISE WITH DIFFERENT ENL $L = 1, 3, 5, 10$ AND DIFFERENT NUMBERS OF IMAGES $N = 1, 2, 3, 5$ IN TEMPORAL DATA SET

		$L = 1$	$L = 3$	$L = 5$	$L = 10$			$L = 1$	$L = 3$	$L = 5$	$L = 10$
<i>peppers</i>	$N = 1$	10.39	13.21	14.29	15.81	$N = 1$	10.71	13.47	14.89	16.69	
		6.58	9.08	10.07	10.97		6.58	8.88	9.68	10.39	
	$N = 2$	8.16	10.35	10.97	11.52	$N = 2$	8.15	9.91	10.38	10.79	
		10.43	13.68	14.99	16.80		11.00	13.91	15.02	16.62	
	$N = 3$	11.51	13.73	14.69	16.20	$N = 3$	12.15	14.53	15.53	16.79	
		9.09	10.86	11.32	11.73		8.91	10.29	10.68	10.95	
	$N = 5$	11.33	13.99	15.37	17.19	$N = 5$	11.77	14.41	15.60	17.29	
		12.15	14.48	15.52	16.84		13.10	15.06	15.99	17.40	
	<i>barbara</i>	$N = 1$	10.03	11.32	11.62	11.90	$N = 1$	9.68	10.65	10.89	11.06
			11.87	14.45	15.74	17.73		12.29	14.77	15.89	18.28
		$N = 2$	12.99	15.20	16.25	17.62	$N = 2$	13.97	15.83	16.78	18.16
			12.37	14.83	16.05	17.60		9.50	11.61	12.62	14.14
$N = 3$		6.89	10.64	12.32	14.18	$N = 3$	6.01	9.15	10.39	11.77	
		9.29	12.81	14.23	15.66		8.01	10.78	11.77	12.66	
$N = 5$		12.90	15.21	16.60	17.81	$N = 5$	9.75	11.87	13.29	15.25	
		13.31	15.63	16.70	18.07		10.75	12.75	13.76	15.22	
<i>lena</i>		$N = 1$	10.62	13.93	15.11	16.32	$N = 1$	9.12	11.59	12.32	13.00
			13.28	15.59	16.77	18.07		9.99	12.62	13.89	15.67
		$N = 2$	14.14	16.13	17.18	18.63	$N = 2$	11.05	13.39	14.34	15.78
			12.26	15.11	16.07	16.86		10.41	12.34	12.86	13.35
	$N = 3$	13.37	15.82	16.80	18.91	$N = 3$	10.97	12.78	14.27	16.41	
		14.80	16.89	17.94	19.18		12.37	14.21	15.24	16.50	
	<i>boat</i>	$N = 1$	12.37	14.83	16.05	17.60	$N = 1$	9.50	11.61	12.62	14.14
			6.89	10.64	12.32	14.18		6.01	9.15	10.39	11.77
		$N = 2$	9.29	12.81	14.23	15.66	$N = 2$	8.01	10.78	11.77	12.66
			12.90	15.21	16.60	17.81		9.75	11.87	13.29	15.25
		$N = 3$	13.31	15.63	16.70	18.07	$N = 3$	10.75	12.75	13.76	15.22
			10.62	13.93	15.11	16.32		9.12	11.59	12.32	13.00
$N = 5$		13.28	15.59	16.77	18.07	$N = 5$	9.99	12.62	13.89	15.67	
		14.14	16.13	17.18	18.63		11.05	13.39	14.34	15.78	
<i>house</i>		$N = 1$	12.26	15.11	16.07	16.86	$N = 1$	10.41	12.34	12.86	13.35
			13.37	15.82	16.80	18.91		10.97	12.78	14.27	16.41
		$N = 2$	14.80	16.89	17.94	19.18	$N = 2$	12.37	14.21	15.24	16.50

IGARSS Fusion Contest 2012 (three images are sensed in 2007 and the other three in 2011);

- *Saint-Gervais-les-Bains* (well-registered): 26 single-look TerraSAR images in Saint-Gervais-les-Bains (France) (13 images are sensed in 2009 and the other 13 images in 2011).

Both of them have been well registered using the sensor parameters.

We assessed the performance of noise reduction in real SAR images by measuring the ENL, i.e.,

$$ENL(\hat{u}) = \frac{[\text{Mean}(\hat{u})]^2}{\text{Var}(\hat{u})} \quad (17)$$

where \hat{u} is the denoised intensity value. The denoising results are shown in Fig. 13. Compared with 1-PPB and T-PPB filters, the 2S-PPB filter reduces the speckle effect comparably in

TABLE II
ENL OF NOISY IMAGES AND ESTIMATED IMAGES USING 1-PPB (BROWN), 3D-ANF (BLUE), T-PPB (GREEN), AND 2S-PPB (RED) FILTERS FOR REAL SAR IMAGES

ENL	<i>Toulouse</i>	<i>San Francisco</i>	<i>S.-G.-Les-B.</i>
	0.83	0.94	0.93
Left green	56.97	41.87	41.48
Region in Fig.11	13.37	10.70	10.79
	89.81	52.31	54.99
	66.78	65.15	497.25
Right green	0.85	0.97	0.92
Region in Fig.11	169.90	16.44	8.79
	21.73	11.78	5.83
	252.33	38.31	13.58
	328.61	69.38	190.08

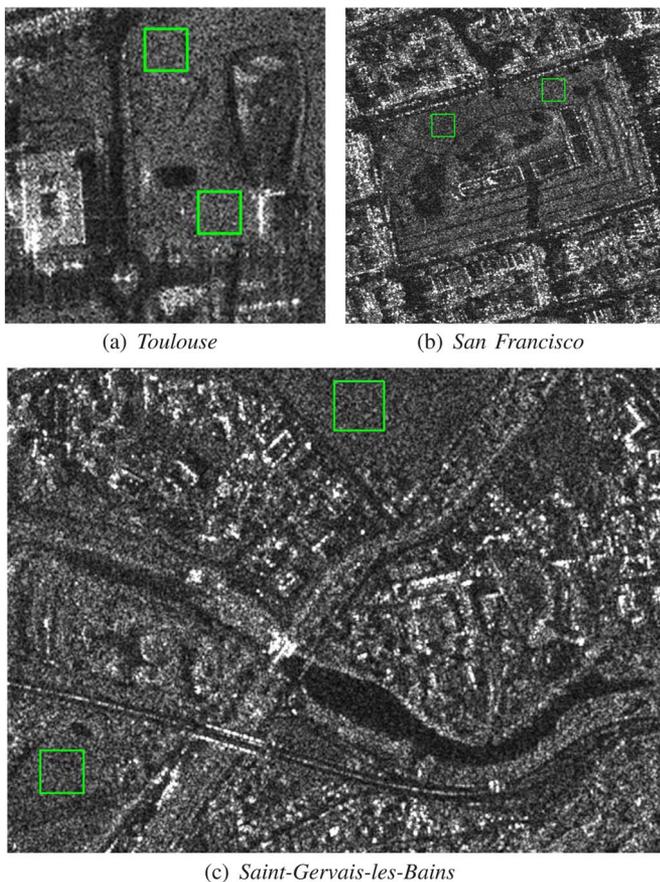


Fig. 11. (a) *Tozulouse*, (b) *San Francisco*, and (c) *Saint-Gervais-les-Bains* noisy images (regions in green rectangles are used to calculated the ENL values).

San Francisco (well-registered). Moreover, more dark and thin streets have been preserved. We also compare the 2S-PPB filter with the classical temporal filtering method, multilooking approach in *Saint-Gervais-les-Bains* (well-registered). Generally, the 2S-PPB filter gets more smoothed results than multilooking, as Region #1 in Fig. 13(c) and (d). More than that, the changes over time can be restored by 2S-PPB, as the changes in Region #2 in Fig. 13(c) and (d). Multilooking tends to ignore the temporal changes (loss of temporal resolution), whereas the 2S-PPB filter can well preserve both the spatial and temporal resolution and reduce the speckle effect. Table II shows the ENL calculated in homogeneous regions (green rectangles in Fig. 11). In general, the proposed filter has higher ENL than other filters.

C. Denoising of Unregistered Images

The real temporal SAR images provided by IGARSS (data information in Section VI-B3) are tested here. The difference is that we manually register those temporal images without using the accurate sensor parameters (see Figs. 12 and 13). Moreover, the miss-registration between two images is about 4–7 pixels (shown in Fig. 14(b), the temporal mean of the temporal images illustrates the unregistration). These unregistered temporal images are identified as *San Francisco* (unregistered). In the miss-registration estimation, the search region $A_t(i)$ is a 21×21 pixel window. Patch size is 7×7 pixels, h and h' have been chosen identical to the filtering step. After miss-registration estimation, the proposed 2S-PPB filter is applied with the same parameters used in Section VI-B.

The temporal mean after miss-registration estimation in Fig. 14(c) shows the performance of the miss-registration estimation. Similarly, for the thin and dark streets in *San Francisco* (unregistered), the proposed 2S-PPB preserves more details than 1-PPB and T-PPB. However, its performance is not as good as the one on well-registered *San Francisco* [in Fig. 14(f)], because of the insufficient estimation of miss-registration. In addition, the miss-registration estimation will fail when the offset between temporal images is too large. Indeed, larger offset estimation needs larger search region $A_t(i)$, but this increases the risk of finding the similar pixels.

VII. CONCLUSION AND DISCUSSION

In this paper, we have presented the limitation of spatio-temporal similarity. Inspired by this phenomenon, we have proposed a two-step denoising framework based on iterative weighted PPB. With the help of miss-registration estimation, we also adapted it to unregistered images. The effective performance in the experiments shows that the proposed method exploits more available information for stable objects while comparably keeping new objects.

Future work will be focused on finding a more efficient change detection method for temporal average. In addition, the adoption of the proposed 2S-PPB for interferometric SAR, polarimetric SAR, and polarimetric interferometric SAR images will be another subject of future work.

APPENDIX A

GLR WITH DIFFERENT NUMBER OF LOOKS

We denote by $p(y, u, L)$ the Gamma distribution, i.e.,

$$p(y, u, L) = \frac{1}{\Gamma(L)} \frac{L}{u} \left(\frac{Ly}{u} \right)^{L-1} \exp \left(-\frac{Ly}{u} \right) \quad (18)$$

where y is the intensity value, u is the noise-free value, and L is the number of looks.

The GLR similarity criterion [30] between y_1 and y_2 is

$$S_{\text{GLR}}(y_1, y_2) = -\log \frac{p(y_1, u_1 = \hat{u}_{12}, L = L_1) p(y_2, u_2 = \hat{u}_{12}, L = L_2)}{p(y_1, u_1 = \hat{u}_1, L = L_1) p(y_2, u_2 = \hat{u}_2, L = L_2)}.$$

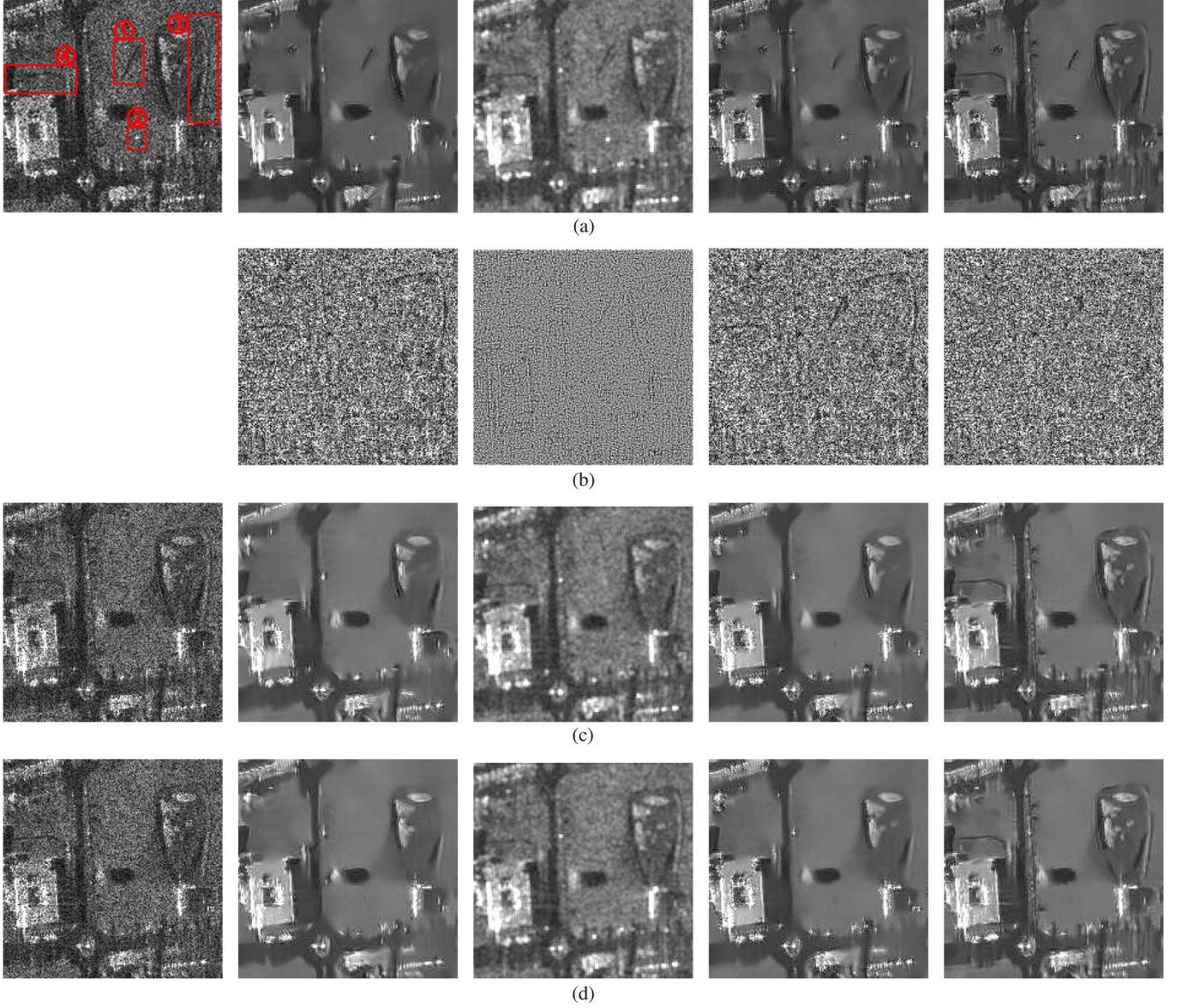


Fig. 12. Denoising results on a zoom of *Toulouse DGA ONERA*. From left to right: noisy image with 1-look multiplicative speckle noise, results by PPB on a single image (1-PPB), results by 3D-ANF on temporal images, results by T-PPB on temporal images, and results by 2S-PPB on temporal images. From top to bottom: image y_{t_1} with new objects (a dark line in Region #1 and a bright target in Region #2), ratio maps of noisy images to denoised images, image y_{t_2} without new objects, and image y_{t_3} without new objects.

Note that in [30], the number of looks L is the same ($L_1 = L_2 = L$). Thus, the parameters of distribution θ only consider the noise-free value u . However, in our case, the number of looks is different, which will have influence on the similarity criterion.

Substitute the Gamma distribution $p(y, u, L)$ in the GLR criterion with $\hat{u}_1 = y_1$, $\hat{u}_2 = y_2$, and $\hat{u}_{12} = (L_1 y_1 + L_2 y_2 / L_1 + L_2)$, i.e.,

$$\begin{aligned}
 S_{\text{GLR}}(y_1, y_2) &= \log \left[\frac{y_1}{\hat{u}_{12}} \left(\frac{\hat{u}_1}{\hat{u}_{12}} \right)^{L_1-1} \exp \left(L_1 - \frac{L_1 y_1}{\hat{u}_{12}} \right) \right]^{-1} \\
 &+ \log \left[\frac{y_2}{\hat{u}_{12}} \left(\frac{\hat{u}_2}{\hat{u}_{12}} \right)^{L_2-1} \exp \left(L_2 - \frac{L_2 y_2}{\hat{u}_{12}} \right) \right]^{-1}
 \end{aligned}$$

$$= L_1 \log \frac{L_1 y_1 + L_2 y_2}{y_1 (L_1 + L_2)} + L_2 \log \frac{L_1 y_1 + L_2 y_2}{y_2 (L_1 + L_2)}.$$

When $L_1 = L_2 = L$, we have

$$S_{\text{GLR}}(y_1, y_2)|_{L_1=L_2=L} = 2L \log \left(\sqrt{\frac{y_1}{y_2}} + \sqrt{\frac{y_2}{y_1}} \right) - 2L \log 2$$

which is the same as the weights formula in [24] (y is the intensity value, and \sqrt{y} is the amplitude value).

APPENDIX B SYMMETRIC KL DIVERGENCE WITH DIFFERENT NUMBER OF LOOKS

For distributions $p(y, u_1, L_1)$ and $p(y, u_2, L_2)$ of a continuous random variable, KL divergence is defined to be the integral

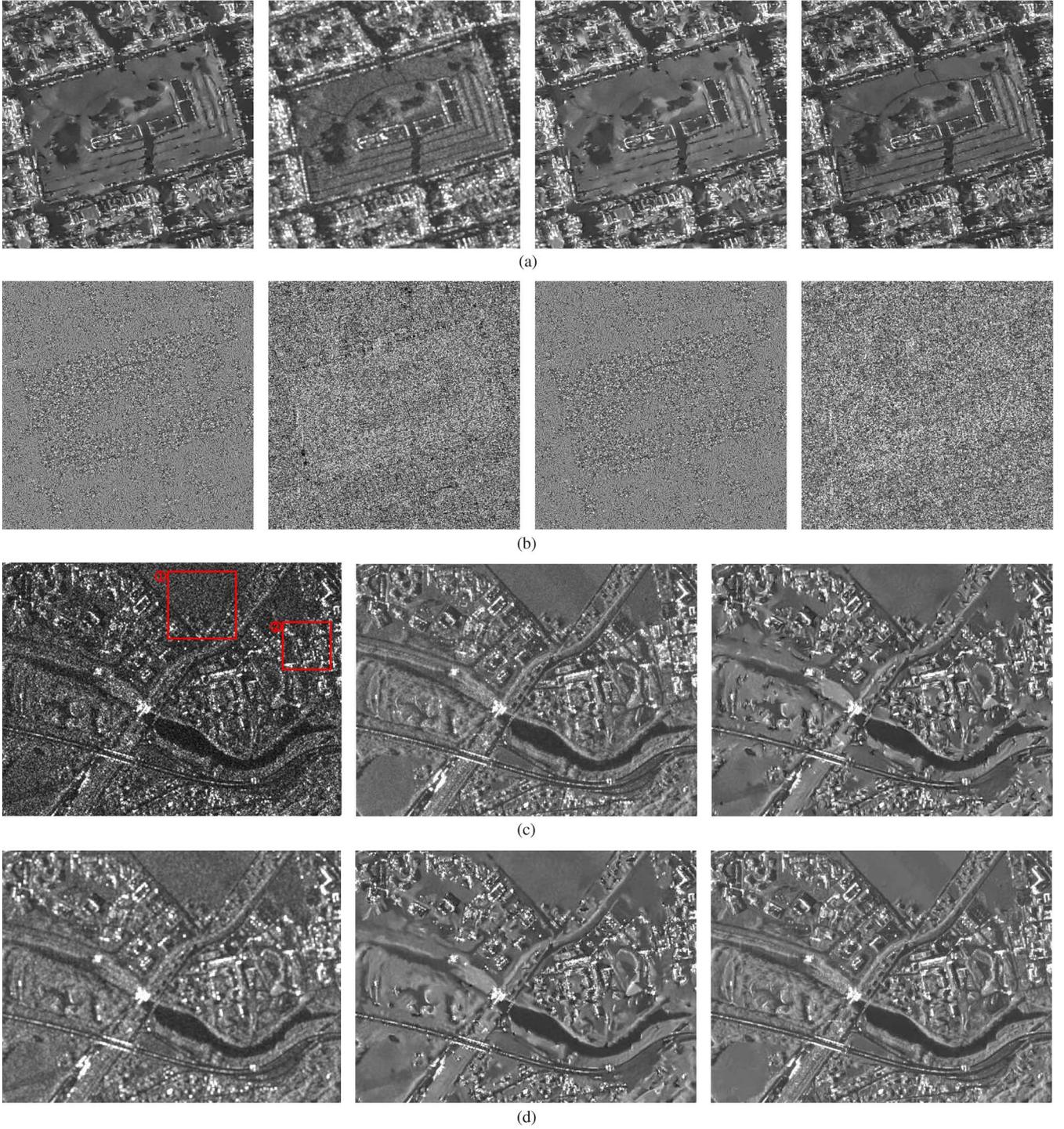


Fig. 13. Denoising results of well-registered *San Francisco* IGARSS and *Saint-Gervais-les-Bains*. (a) *San Francisco* denoising results. From left to right: 1-PPB on a single image, 3D-ANF, T-PPB, and 2S-PPB. (b) Ratio map of *San Francisco* images. From left to right: 1-PPB on a single image, 3D-ANF, T-PPB, and 2S-PPB. (c) *Saint-Gervais-les-Bains* images. From left to right: noisy image, denoising result by multilooking, and denoising results by 1-PPB on a single image. (d) *Saint-Gervais-les-Bains* images. From left to right: 3D-ANF, denoising result by T-PPB on temporal images, and denoising results by 2S-PPB on temporal images. Stable Region #1 in *Saint-Gervais-les-Bains* and changed Region #2 in *Saint-Gervais-les-Bains*.

as follows:

Substitute the Gamma distribution $p(y, u, L)$ in the GLR criterion with $y \leq 0$, i.e.,

$$D_{\text{KL},1} = \int_{-\infty}^{+\infty} p(y, u_1, L_1) \ln \left[\frac{p(y, u_1, L_1)}{p(y, u_2, L_2)} \right] dy. \quad (19)$$

$$D_{\text{KL},1} = \int_0^{+\infty} p(y, u_1, L_1) \ln[A \cdot B \cdot C] dy \quad (20)$$

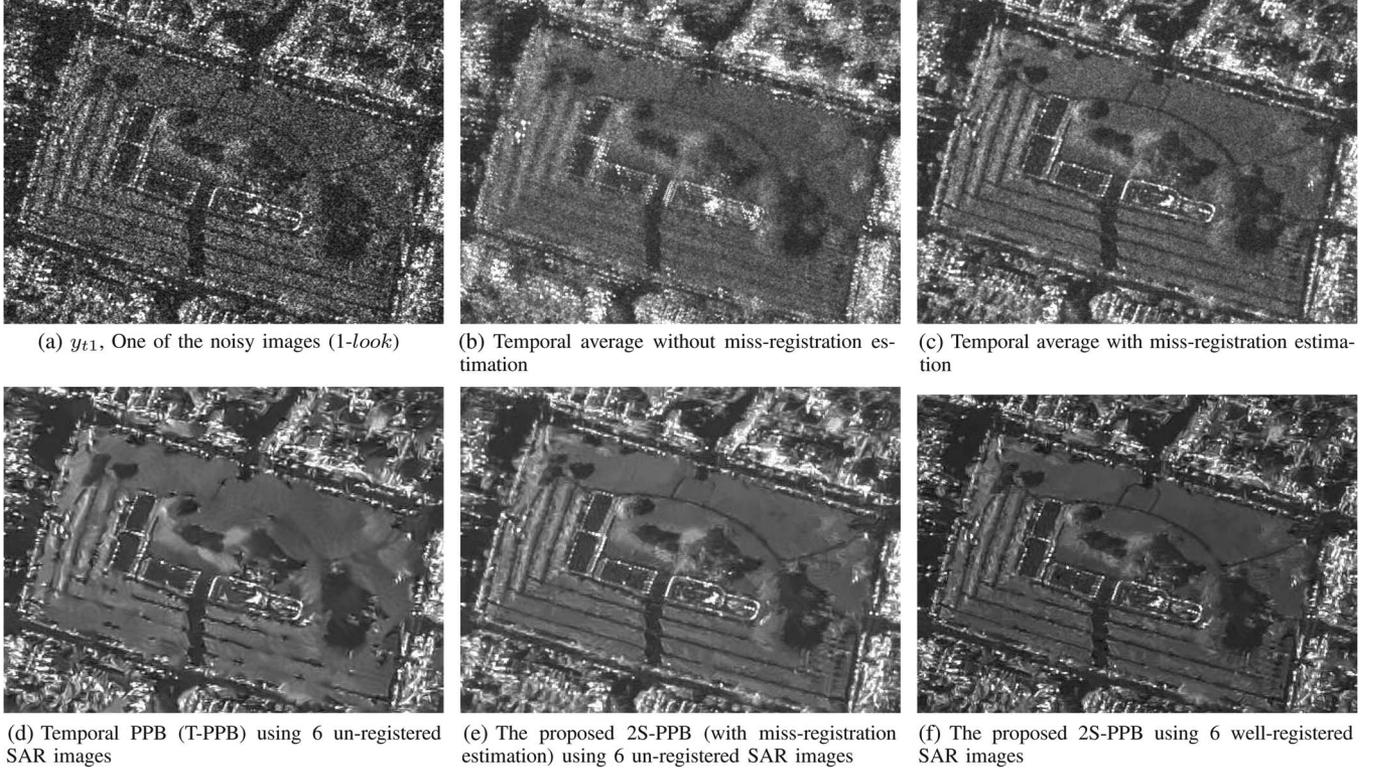


Fig. 14. Denoising results of unregistered *San Francisco* IGARSS. (a) y_{t1} , one of the noisy images (1-look). (b) Temporal average of the multitemporal SAR images, which shows that the temporal images are unregistered. (c) Temporal average of the multitemporal SAR images after miss-registration estimation. (d) Denoising results by T-PPB on unregistered temporal images. (e) Denoising results by 2S-PPB on unregistered temporal images. (f) Denoising results by 2S-PPB on well-registered temporal images.

where

$$A = \frac{\frac{1}{\Gamma(L_1)} \frac{L_1}{u_1}}{\frac{1}{\Gamma(L_2)} \frac{L_2}{u_2}}, B = \left(\frac{L_1 y}{u_1}\right)^{L_1-1} \left(\frac{L_2 y}{u_2}\right)^{1-L_2}$$

$$C = \exp\left[\frac{L_2 y}{u_2} - \frac{L_1 y}{u_1}\right]. \quad (21)$$

Then

$$D_{KL,1} = \ln A \int_0^{+\infty} p(y, u_1, L_1) dy + \int_0^{+\infty} p(y, u_1, L_1) \ln B dy$$

$$+ \int_0^{+\infty} p(y, u_1, L_1) \ln C dy. \quad (22)$$

Since

$$\int_0^{+\infty} p(y, u, L) dy = 1$$

$$\int_0^{+\infty} p(y, u, L) \ln y dy = E(\ln y) = \psi(L) + \ln(u)$$

$$\int_0^{+\infty} p(y, u, L) y dy = E(y) = u \quad (23)$$

thus

$$D_{KL,1} = \ln A + (L_1 - 1) \ln \frac{L_1}{u_1} + (1 - L_2) \ln \frac{L_2}{u_2}$$

$$+ (L_1 - L_2) [\psi(L_1) + \ln(u_1)] + L_2 \frac{u_1}{u_2} - L_1 \quad (24)$$

where $\psi(L)$ is the digamma function.

With

$$D_{KL,2} = \int_{-\infty}^{+\infty} p(y, u_2, L_2) \ln \left[\frac{p(y, u_2, L_2)}{p(y, u_1, L_1)} \right] dy \quad (25)$$

the KL divergence criterion is

$$S_{KL}(u_1, u_2) = D_{KL,1} + D_{KL,2}$$

$$= L_1 \frac{u_2}{u_1} + L_2 \frac{u_1}{u_2} - L_1 - L_2$$

$$+ (L_1 - L_2) [\psi(L_1) - \psi(L_2)]$$

$$+ (L_1 - L_2) [\ln(u_1) - \ln(u_2)]. \quad (26)$$

When $L_1 = L_2 = L$, we have

$$S_{KL}(u_1, u_2) = D_{KL,1} + D_{KL,2} = L \left[\frac{u_1}{u_2} + \frac{u_2}{u_1} - 2 \right] \quad (27)$$

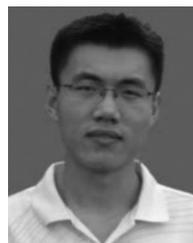
which is the same as the weights formula in [24].

ACKNOWLEDGMENT

The authors would like to thank J.-M. Nicolas for his help on image registration and the IGARSS Data Fusion Contest for providing the multitemporal SAR images. They would also like to thank the German Aerospace Agency (DLR) for the TerraSAR-X images (project MTH0232) and the Centre National d'Etudes Spatiales and the Office National d'Etudes et de Recherches Aérospatiales for the 100-look SAR images.

REFERENCES

- [1] J. Bruniquel and A. Lopes, "Multi-variate optimal speckle reduction in SAR imagery," *Int. J. Remote Sens.*, vol. 18, no. 3, pp. 603–627, Feb. 1997.
- [2] D. Coltuc, E. Trouvé, F. Bujor, N. Classeau, and J. Rudant, "Time-space filtering of multitemporal SAR images," in *Proc. IEEE IGARSS*, 2000, vol. 7, pp. 2909–2911.
- [3] M. Ciuc, P. Bolon, E. Trouve, V. Buzuloiu, and J. Rudant, "Adaptive-neighborhood speckle removal in multitemporal synthetic aperture radar images," *Appl. Opt.*, vol. 40, no. 32, pp. 5954–5966, Nov. 2001.
- [4] D. Sen, M. Swamy, and M. Ahmad, "A homomorphic system to reduce speckle in videos," in *Proc. IEEE IGARSS*, 2005, vol. 6, pp. 4287–4290.
- [5] 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, Dec. 2005.
- [6] J. Inglada and G. Mercier, "A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1432–1445, May 2007.
- [7] S. Hachicha, C. Deledalle, F. Chaabane, and F. Tupin, "Multi-temporal SAR classification according to change detection operators," in *Proc. 6th Int. Workshop Anal. Multi-Temp.*, 2011, pp. 133–136.
- [8] 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.
- [9] J. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 2, pp. 165–168, Mar. 1980.
- [10] J. Lee, "Refined filtering of image noise using local statistics," *Comput. Graph. Image Process.*, vol. 15, no. 4, pp. 380–389, Apr. 1981.
- [11] D. Kuan, A. Sawchuk, T. Strand, and P. Chavel, "Adaptive noise smoothing filter for images with signal-dependent noise," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-7, no. 2, pp. 165–177, Mar. 1985.
- [12] D. Kuan, A. Sawchuk, T. Strand, and P. Chavel, "Adaptive restoration of images with speckle," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 35, no. 3, pp. 373–383, Mar. 1987.
- [13] V. Frost, J. Stiles, K. Shanmugan, and J. 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.
- [14] R. Touzi, "A review of speckle filtering in the context of estimation theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2392–2404, Nov. 2002.
- [15] H. Guo, J. Odegard, M. Lang, R. Gopinath, I. Selesnick, and C. Burrus, "Wavelet based speckle reduction with application to SAR based ATDR," in *Proc. ICIP*, 1994, vol. 1, pp. 75–79.
- [16] L. Gagnon and A. Jouan, "Speckle filtering of SAR images: A comparative study between complex-wavelet-based and standard filters," in *Proc. SPIE*, 1997, pp. 80–91.
- [17] A. Achim, P. Tsakalides, and A. Bezerianos, "SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 8, pp. 1773–1784, Aug. 2003.
- [18] J. Ranjani and S. Thiruvengadam, "Dual-tree complex wavelet transform based SAR despeckling using interscale dependence," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2723–2731, Jun. 2010.
- [19] D. Gleich, M. Kseneman, and M. Datcu, "Despeckling of TerraSAR-X data using second-generation wavelets," *IEEE Trans. Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 68–72, Jan. 2010.
- [20] A. Buades, B. Coll, and J. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. CVPR*, 2005, vol. 2, pp. 60–65.
- [21] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [22] V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola, "From local kernel to nonlocal multiple-model image denoising," *Int. J. Comput. Vis.*, vol. 86, no. 1, pp. 1–32, Jan. 2010.
- [23] A. Foi, V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1395–1411, May 2007.
- [24] C. 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] J. Goodman, "Some fundamental properties of speckle," *J. Opt. Soc. Amer.*, vol. 66, no. 11, pp. 1145–1150, Nov. 1976.
- [26] S. Parrilli, M. Poderico, C. 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.
- [27] E. Trouvé, Y. Chambenoit, N. Classeau, and P. Bolon, "Statistical and operational performance assessment of multitemporal SAR image filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2519–2530, Nov. 2003.
- [28] S. Quegan and J. Yu, "Filtering of multichannel SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2373–2379, Nov. 2001.
- [29] V. Duval, J.-F. Aujol, and Y. Gousseau, "On the parameter choice for the Non-Local Means, Centre pour la Communication Scientifique Directe, Villeurbanne, France, Tech. Rep. hal-00468856, HAL. [Online]. Available: <http://hal.archives-ouvertes.fr/hal-00468856>
- [30] C. A. Deledalle, L. Denis, and F. Tupin (2012). How to compare noisy patches? Patch similarity beyond Gaussian noise. *Int. J. Comput. Vis.* [Online]. 99(1), pp. 86–102. Available: <http://dx.doi.org/10.1007/s11263-012-0519-6>
- [31] 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.
- [32] V. Duval, J. Aujol, and Y. Gousseau, "A bias-variance approach for the nonlocal means," *SIAM J. Imag. Sci.*, vol. 4, no. 2, pp. 760–788, 2011.
- [33] A. Buades, B. Coll, and J. Morel, "Nonlocal image and movie denoising," *Int. J. Comput. Vis.*, vol. 76, no. 2, pp. 123–139, Feb. 2008.
- [34] A. Gray, N. Short, K. Mattar, and K. Jezek, "Velocities and flux of the Filchner ice shelf and its tributaries determined from speckle tracking interferometry," *Can. J. Remote Sens.*, vol. 27, no. 3, pp. 193–206, 2001.
- [35] T. Strozzi, A. Luckman, T. Murray, U. Wegmuller, and C. L. Werner, "Glacier motion estimation using SAR offset-tracking procedures," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2384–2391, Nov. 2002.



Xin Su is currently working toward the Ph.D. degree at Télécom ParisTech, Paris, France.

His main research interests include multitemporal synthetic aperture radar image analysis, signal processing, image processing, pattern recognition, and machine learning.



Charles-Alban Deledalle received the B.S. degree in engineering from the École Pour l'Informatique et les Techniques Avancées, Kremlin Bicêtre Cedex, France; the M.S. degree in science and technology from the University Pierre et Marie Curie (Paris 6), Paris, France, in 2008; and the Ph.D. degree at Télécom ParisTech, Paris, in 2011.

He is currently a Research Associate with the Institut de Mathématiques de Bordeaux, Centre National de la Recherche Scientifique—Université Bordeaux I, Talence, France. His main research interests include image denoising, analysis, and risk estimation for inverse problems.

Dr. Deledalle received the Best Student Paper Award at the 17th IEEE International Conference on Image Processing (ICIP 2010) and a Ph.D. award in Signal, Image and Vision (Club EEA/GdR ISIS/GRETSI) for his thesis.



Florence Tupin (SM'07) received the B.S. degree in engineering and the Ph.D. degree in signal and image processing from the Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1994 and 1997, respectively.

She is currently a Professor with Télécom ParisTech, Paris. From 1997 to 1998, she was with SAGEM, engaged in fingerprint recognition. Since November 1998, she has been an Associate Professor and then Professor of signal processing and computer vision with the Department of Image and Signal

Processing, Télécom ParisTech. Her research interests include image analysis and interpretation, 3-D reconstruction, Markov random fields, and synthetic aperture radar, specially for urban remote sensing applications.

Dr. Tupin was the Chair of the Urban Remote Sensing Joint Event held in Paris in 2007. Since 2003, she has been a member of the technical committees Pattern Recognition in Remote Sensing of the International Association for Pattern Recognition and URBAN, the biennial GRSS/ISPRS Joint Workshops on Remote Sensing and Data Fusion over Urban Areas from 2004 to 2006. From 2005 to 2007, she was an Associate Editor of the French journal *Signal Processing*. She currently serves as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.



Hong Sun (M'01) received the B.S., M.S., and Ph.D. degrees in communications and electrical systems from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 1981, 1984, and 1995, respectively.

From 1984 to 2000, she held teaching and research positions, including being a Professor with HUST. She was a Visiting Scholar with the Conservatoire National des Arts et Métiers, Paris, France, in 1997 and a Visiting Professor with the École Nationale Supérieure des Télécommunications (ENST), Paris,

in 1998, 1999, 2000, 2001, and 2007. Since 2001, she has been with the School of Electronic Information, Wuhan University, Wuhan, where she is currently a Professor and the Head of the Signal Processing Laboratory. Her research interest includes digital signal processing theory and its applications, including works on image interpretation, communication signal processing, and speech signal processing. Since 1998, when she was with ENST, her research focus has been on synthetic aperture radar (SAR) and polarimetric SAR image processing and application techniques.

Dr. Sun is a member of the IEEE Group on Signal Processing and the Vice Chairman of the Signal Processing Society of the Chinese Institute of Electronics.