TWO STEPS MULTI-TEMPORAL NON-LOCAL MEANS FOR SAR IMAGES

Xin Su[1], Charles-Alban Deledalle[2], Florence Tupin[1], Hong Sun[3]

1. Institut Telecom, Telecom ParisTech, LTCI, Paris, France
2. CNRS, Ceremade, Paris Dauphine University, Paris, France
3. School of Electronic Information, Wuhan University, Wuhan, P. R. China

ABSTRACT

This paper presents a denoising approach for multi-temporal Synthetic aperture radar (SAR) images based on Non-Local Means (NLM) method. To exploit redundancy existing in multi-temporal images, we develop a new strategy of NLM for multi-temporal data. Instead of directly overspreading the NLM operator from one image to temporal images, a two steps weighted average is proposed in this paper. The first step is a maximum likelihood estimate with binary weights on temporal pixels and the second step is iterative NL means on spatial pixels. Experiments in this paper illustrate that the proposed method can effectively exploit image redundancy and denoise multi-temporal images.

Index Terms— Image denoising, Non-Local Means (NLM), Multi-temporal SAR Images

1. INTRODUCTION

With the increase of the number of Synthetic aperture radar (SAR) systems, more SAR images on the same region are available. Those time series, or multi-temporal SAR images provide to people more comprehensive information including both spatial and temporal domains. However, speckle, universally existing in the coherent imagery systems, causes visually undesired effect to SAR images. Many papers have focused on SAR image denoising aiming to remove the speckle, even though the discussion about whether the speckle noise is non-informative or should be removed is still going on [1]. Nevertheless, for many tasks, SAR image denoising helps automatic tools or visual analysis.

Most existing image denoising techniques can be classified into spatial-domain methods or transform-domain methods. For instance, Buades *et al.* [2] presented an approach, called Non-Local Means (NLM), which estimates each pixel by a weighted average of the similar pixels. Typically, these weights were defined according to the distance between patches with these pixels in center, say Euclidean distance

when Gaussian noise is considered. An adaptive search window was used for the NLM in [3]. Deledalle *et al.* [4] derived the NLM to multiplicative speckle noise case (SAR images) by redefining the similarity between patches and iteratively refining weights. A successful spatial-domain denoising method called K-SVD was proposed in [5], which denoised the image with sparse representation on a patches-dictionary. As transform-domain method, Xie *et al.* [6] proposed a speckle reduction algorithm by fusing the wavelet Bayesian denoising technique with Markov random field based image regularization. The recently proposed BM3D [7] identified the similar patches in the image spatial domain and denoised patches in the 3D transform domain. An extension to SAR images is proposed in [8].

Noise reduction for time series data, such as video images, can use information not only in spatial but also temporal domain. Ji *et al.* [9] proposed a patch-based video denoising algorithm handling serious mixed noise in the video data. Since time series are similar to video data, denoising for multi-temporal SAR data can make good use of time series, while different noise statistical model and temporal coherency model should be considered. Based on the NLM framework, we employed the similarity measurement proposed in [10], dividing the denoising process into two steps, which deal with temporal and spatial information respectively (detailed in section 3). The main purpose of this paper is exploiting all the available information for temporally stable pixels while keeping the new information as much as possible.

2. RELATED WORKS

The key idea of the Non-Local Means (NLM) presented in [2] is to estimate actual pixel intensities with image redundancy. The way to exploit image redundancy is by searching similar pixels and estimating a weighted maximum likelihood with these pixels. Define the observed image as $\{y(i)\}$ and $\hat{u}(i)$ as the estimation of actual pixel $u(i)$ (all pixels indexed by i), the formulation expression of NLM can be any maximum likelihood estimator, but for the sake of simplicity we will consider intensity images, for which the formula holds $\hat{u}(i) = \sum_{j \in \Omega} w(i,j)y(j)$, where $w(i,j)$ denotes the weight of similar pixel $y(j)$ for pixel $y(i)$. The scope of $\sum$ is the

searching zone Ω.

A direct extension of NLM for multi-temporal data is to define a cube search window, by aggregating all the search windows of the different dates[1]. Let denote by $\mathbb{S} = \{I_{t1}, I_{t2}, ..., I_{tN}\}$ the multi-temporal images, this temporal NLM filter can be defined on the cube $C = \{\Omega_{t1}, \Omega_{t2}, ..., \Omega_{tN}\}$. Meanwhile, the formulation to estimate the *true* value $u_t(i)$ is:

$$\hat{u}_t(i) = \sum_{y_{t'}(j) \in C} w(i_t, j_{t'}) y_{t'}(j) \qquad (1)$$

where,

$$w(i_t, j_{t'}) = \frac{1}{Z} \exp\left[-\frac{1}{h_0} S(i_t, j_{t'}) - \frac{L}{h_1} R^{m-1}(i_t, j_{t'})\right]$$

$$S(i_t, j_{t'}) = \sum_{k \in K} \log\left[\frac{y_t(i, k) + y_{t'}(j, k)}{y_t^{1/2}(i, k) y_{t'}^{1/2}(j, k)}\right]$$

$$R^{m-1}(i_t, j_{t'}) = \sum_{k \in K} \frac{[\hat{u}_t^{m-1}(i, k) - \hat{u}_{t'}^{m-1}(j, k)]^2}{\hat{u}_t^{m-1}(i, k) \hat{u}_{t'}^{m-1}(j, k)}$$
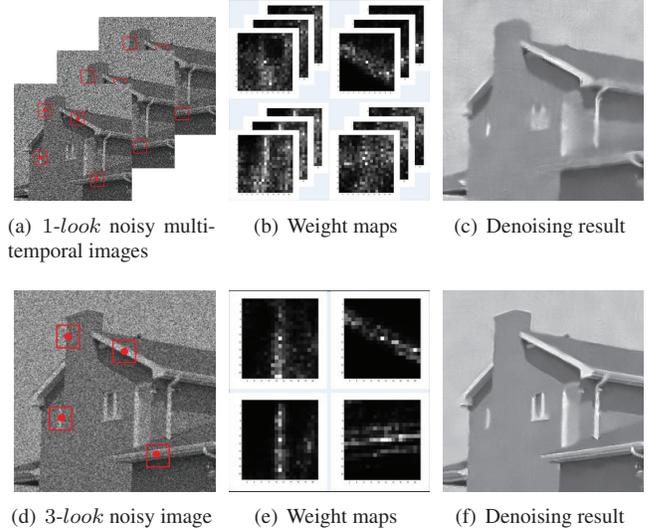
where t and $t' \in \{t1, t2, ..., tN\}$, $y_t(i)$ and $y_{t'}(j)$ are respectively the $i-$th pixel of I_t and the $j-$th pixel of $I_{t'}$. In order to fulfill the multiplicative Goodman's speckle noise distribution model [12] in SAR images, the similarity between pixels $y_t(i)$ and $y_{t'}(j)$ is defined as $S(i_t, j_{t'})$ [4, 10]. $R^{m-1}(i_t, j_{t'})$ is an iterative term which refines the weights by previous estimate results $\hat{I}_t^{m-1} = \{\hat{u}_t^{m-1}(i)\}$, and in the first iteration $m = 1$ and $\hat{I}_t^0 = \{\hat{u}_t^0(i) = 1\}$. Z is a normalization parameter, h_0 and h_1 control the decay of the weights, K is a patch with pixel i (or j) as center, $y_t(i, k)$ is the k-th neighbor of $y_t(i)$ in K. As such, a Temporal It-PPB is given by Eq.1 (a direct extension of It-PPB in [4]). However, unfortunately, the denoising results on the multi-temporal SAR images using the Temporal It-PPB demonstrate very limited improvement of the visual performance compared with the same operation on single image, see Fig.3.c and d.

3. METHODOLOGY

3.1. Analysis of Similarity between Noisy Patches

To analyze the limited improvement found in spatio-temporal case, we tested a synthetic set of multi-temporal SAR images $\mathbb{S} = \{I_{t1}, I_{t2}, I_{t3}\}$ (3 images with 1 equivalent number of *looks* speckle noise, Fig.1.a) using two kinds of strategies. 1) Denoise $\mathbb{S}$ directly using Eq.1. 2) Firstly get the mean values $\bar{\mathbb{S}} = \{\bar{I}_{t1 \sim 3}\}$ (with equal weights) of the multi-temporal SAR images $\mathbb{S}$ and then denoise the average image $\bar{\mathbb{S}}$ using Eq.1. The results are shown in Fig.1.c and f. It is interesting to see that the two approaches, using almost the same information provided by the input noisy images of $\mathbb{S}$, obtain significantly different results. Therefore, we illustrate the weights

[1]Contrary to [11] here, we consider a cube search window C

(a) 1-*look* noisy multi-temporal images

(b) Weight maps

(c) Denoising result

(d) 3-*look* noisy image

(e) Weight maps
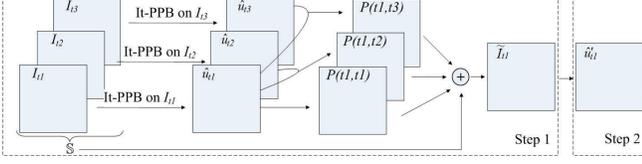
(f) Denoising result

Fig. 1. It-PPB in [4] results on temporal images and associated weight maps. (top: denoising of one date using the spatio-temporal cube; bottom: denoising of the temporal mean)

($w(i|t, j|t')$ in Eq.1) in normalized map, which can denote the similarity between patches, to explain the denoising process.

The two strategies use the same denoising method (Eq.1), but different input noisy data, the former is multi-temporal images $\mathbb{S}$ with 1-*look* speckle noise and the latter is single image $\bar{\mathbb{S}}$ with 3-*look* speckle noise (the average of 3 multi-temporal images $\mathbb{S}$). Four interesting pixels are chosen to illustrate their weight maps, in the middle in Fig.1. Compared with 3-*look* image $\bar{\mathbb{S}}$, the weight maps in 1-*look* image $I_{t1} \in \mathbb{S}$ indicate that more dissimilar pixels are incorrectly recognized, and have been given large weights, because of the higher noise. Although more pixels are available, they are not chosen more efficiently in the multi-temporal data set $\mathbb{S}$. Therefore, although 1-*look* image set $\mathbb{S}$ contains as much information as the average image $\bar{\mathbb{S}}$, it is hard to find out the usefull information in $\mathbb{S}$ to estimate the results.

3.2. Two Steps Patch-Based Denoising

Taking inspiration from the comparison experiment in Section 3.1, we propose a more efficient way to exploit similar pixels for estimation. In case of stable pixels (no change other time), the previous experiments illustrate the usefulness of temporal averaging. That is finding similar pixels in the equally weighted average image $\bar{\mathbb{S}}$ rather than in the original ones $\mathbb{S}$ (the premise is that there is no change taking place over time t). Because if pixels located at the same position but different time (like $y_t(i)$ 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 to estimate $u_t(i)$ as well as $u_{t'}(i)$. From a probability point of view, this equally weighted average can be consid-

Fig. 2. The sketch map of the proposed filter of one image I_{t1}. ($\mathbb{S} = \{I_{t1}, I_{t2}, I_{t3}\}$: the original multi-temporal images; $\hat{u}_{t1}$: the denoising result of I_{t1} using It-PPB [4]; $P(t1, t2)$: the temporal relation between I_{t1} and I_{t2}; $\tilde{I}_{t1}$: the temporal average image; $\hat{u}'_{t1}$: the final denoising result)

ered as a kind of estimate using prior information, relative to the estimate using likelihood information in [4]. Our aim is to create a first denoised image by combining stable pixels while keeping unchanged the pixels not in accordance with the other dates. This is done in Step 1. Then on this improved image, a denoising step is applied. The details of the proposed filter for multi-temporal images are illustrated as follows and in Fig.2.

1) Step 1: Temporal Average with Binary Weights

For the sake of simplicity, we use Eq.2 as a criterion to illustrate the temporal relation between $y_t(i)$ and $y_{t'}(i)$.

$$P_i(t, t') = \begin{cases} 1, & \text{if } \frac{[\hat{u}_t(i) - \hat{u}_{t'}(i)]^2}{\hat{u}_t(i)\hat{u}_{t'}(i)} > T_{SA} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where, $\hat{u}_t(i)$ and $\hat{u}_{t'}(i)$ are the denoising result respectively on $\{I_t\}$ and $\{I_{t'}\}$ using It-PPB [4], T_{SA} denotes the threshold. $P_i(t, t') = 1$ means pixel $y_t(i)$ and $y_{t'}(i)$ have high probability to be realizations coming from the same underlying reflectivity, and $P_i(t, t' = t) \equiv 1$. The following formula shows the temporal averaging process with binary weights.

$$\tilde{y}_t(i) = \frac{1}{Z} \sum_{t' \in [t1, tN]} P_i(t, t') y_{t'}(i) \quad (3)$$

where $Z = \sum_{t' \in [t1, tN]} P_i(t, t')$. The reason here we use It-PPB [4] denoising results $\hat{u}_t$ rather than the original images I_t is that the less noise makes the criterion more robust.

2) Step 2: Spatial Average

Denoising on the temporally weighted average image $\tilde{I}_t = \{\tilde{y}_t(i)\}$ is comparable to Eq.1. However, different pixels $\tilde{y}_t(i)$ may have different (equivalent) number of $looks$ depending on the number of averaged data. In this case the similarity $\tilde{S}(i_t, j_t)$ between pixels $\tilde{y}_t(i)$ and $\tilde{y}_t(j)$ has to be modified to take into account varying number of $looks$:

$$\tilde{S}(i_t, j_t) = \sum_{k \in K} \log \left[\frac{\left[\frac{\tilde{L}_{ti}\tilde{y}_t(i, k) + \tilde{L}_{tj}\tilde{y}_t(j, k)}{\tilde{L}_{ti} + \tilde{L}_{tj}} \right]^{\tilde{L}_{ti} + \tilde{L}_{tj}}}{\tilde{y}_t^{\tilde{L}_{ti}}(i, k)\tilde{y}_t^{\tilde{L}_{tj}}(j, k)} \right]$$

where $\tilde{L}_{ti}$ is the (equivalent) number of $looks$ of $\tilde{y}_t(i)$, and $\tilde{L}_{ti} = \sum_{t' \in [t1, tN]} L_{t'} P_i(t, t')$ in which $L_{t'}$ is the original (equivalent) number of $looks$ of image $I_{t'} = \{y_{t'}(i)\}$. And then, the estimate $\hat{u}'_t(i)$ using $\tilde{I}_t = \{\tilde{y}_t(i)\}$ will be:

$$\hat{u}'_t(i) = \sum_{\tilde{y}_t(j) \in \tilde{C}} \tilde{w}(i_t, j_t) \tilde{y}_t(j) \quad (4)$$

$$\tilde{w}(i_t, j_t) = \frac{1}{Z} \exp \left[-\frac{1}{h_0} \tilde{S}(i_t, j_t) - \frac{L}{h_1} R^{m-1}(i_t, j_t) \right]$$

and $R^{m-1}(i_t, j_t)$ plays the same role as in Eq.1.

4. EXPERIMENTS AND RESULTS

1) Data Sets

We firstly test the proposed filter on synthetic images corrupted by 1-$look$ multiplicative speckle noise, shown in Fig.3.a. In order to simulate the changes in the multi-temporal SAR images $\mathbb{S} = \{I_{t1}, I_{t2}, I_{t3}\}$, a dark $line$ and a bright $target$ are added to I_{t1}, labeled by red rectangles in Fig.3.a. I_{t2} and I_{t3} are corrupted by different speckle noise without the dark $line$ and the bright $target$ (Fig.3.b).

We also test a real multi-temporal SAR images set, single-$look$ TerraSAR X-bands data, 1m×1m spatial resolution, San Francisco, USA, which consists of 6 images sensed in 2007 and 2011. For image registration, we use the sensor parameters. Fig.3.f shows a patch of the latest image sensed in 2011-10-24, while the other 5 images are also used during the denoising process.

**2) Choice of the Parameters**

As suggested by Deledalle $et\ al$, the search window Ω and patch size K enlarge with the increase of the number of iteration, $\Omega \in \{3 \times 3, 7 \times 7, 11 \times 11, 21 \times 21\}$ and $K \in \{1 \times 1, 3 \times 3, 5 \times 5, 7 \times 7\}$, as well as $h_1 = 0.2K$. h_0 controlling the decay of weights depends on the distribution of similarity between two $pure$ noise patches, and is defined as the α-quantile of this $pure$ distribution ($\alpha = 0.92$). T_{SA} is also defined according to this $pure$ distribution, normally $0.98 < \alpha < 1$ in the α-quantile.

3) Results

The comparisons that have been tested here are It-PPB only on single image I_{t1} and Temporal It-PPB on the multi-temporal image set $\mathbb{S}$. Because of the limit of space and for the visual clarity, we just show parts of images to illustrate the performance. Fig.3 presents the denoised images. Using the proposed method, more details are kept in small objects than with other methods.

5. CONCLUSION

In this work, we have presented the limitation of spatio-temporal similarity. Inspired by this phenomenon, we have proposed a two steps denoising approach based on iterative
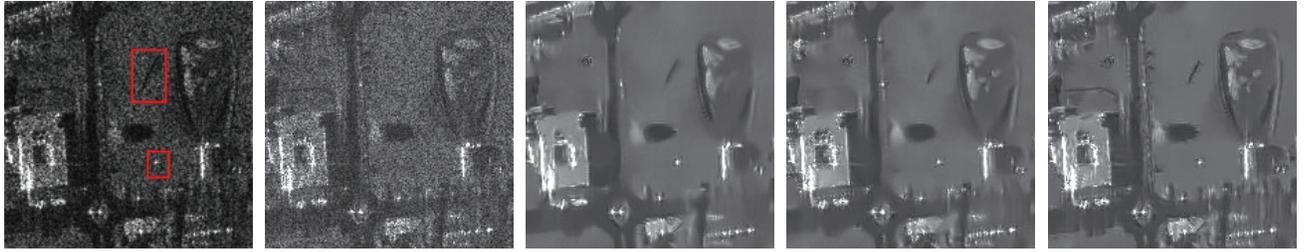
(a) I_{t1} in the noisy images set $\mathbb{S}$ (1-*look*)

(b) I_{t2} in the noisy images set $\mathbb{S}$ (1-*look*, I_{t3} is similar to image I_{t2})

(c) It-PPB [4] on the single image $\tilde{I}_{t1}$

(d) Temporal It-PPB (Eq.1) on the multi-temporal image set $\mathbb{S}$

(e) The proposed approach (Eq.4) on the temporally weighted average image $\tilde{I}_{t1}$
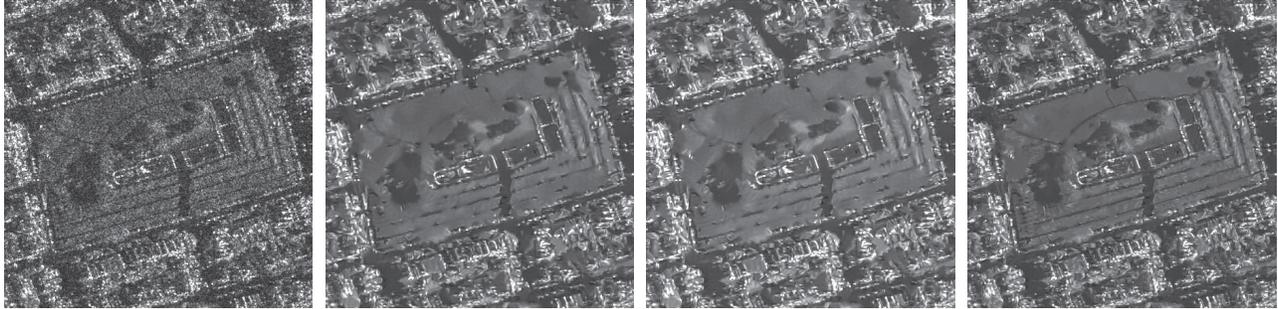
(f) I_{t1}, One of the noisy images (1-*look*)

(g) It-PPB [4] on the single image $\tilde{I}_{t1}$

(h) Temporal It-PPB (Eq.1) on the multi-temporal image set $\mathbb{S}$

(i) The proposed approach (Eq.4) on the temporally weighted average image $\tilde{I}_{t1}$

Fig. 3. Denoising results. (a-e: results of simulation images; f-i: results of real SAR images)

weighted NLM. The effective performance in the experiments shows that the proposed method exploits more available information for stable objects, meanwhile comparably keeping new objects. However, the main drawback of the proposed method is that the temporal average depends on the pre-denoising results (shown in Eq.2) and requires a very accurate registration. In this case the final denoising results are too sensitive to the temporal average, which will be the subject of future work.

6. REFERENCES

[1] Z.G. Xia and Y. Sheng, "Radar speckle: noise or information?," in *IGARSS'96.'Remote Sensing for a Sustainable Future.', International*. IEEE, 1996, vol. 1, pp. 48–50.

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

[3] C. Kervrann and J. Boulanger, "Optimal spatial adaptation for patch-based image denoising," *IEEE Trans. on Image Processing*, vol. 15, no. 10, pp. 2866–2878, 2006.

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

[5] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. on Image Processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[6] H. Xie, L.E. Pierce, and F.T. Ulaby, "SAR speckle reduction using wavelet denoising and Markov random field modeling," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 2196–2212, 2002.

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

[8] S. Parrilli, M. Poderico, C.V. Angelino, and L. Verdoliva, "A nonlocal sar image denoising algorithm based on llmmse wavelet shrinkage," *IEEE Trans. on Geoscience and Remote Sensing*, , no. 99, pp. 1–11, 2012.

[9] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using low rank matrix completion," in *CVPR, 2010 IEEE Conference on*. IEEE, 2010, pp. 1791–1798.

[10] C.A. Deledalle, L. Denis, and F. Tupin, "How to compare noisy patches? patch similarity beyond gaussian noise," *International Journal of Computer Vision*, vol. (DOI) 10.1007/s11263-012-0519-6, 2012.

[11] S. Hachicha, C.A. Deledalle, F. Chaabane, and F. Tupin, "Multi-temporal SAR classification according to change detection operators," in *Analysis of Multi-temporal Remote Sensing Images (Multi-Temp), 2011 6th International Workshop on the*. IEEE, 2011, pp. 133–136.

[12] J.W. Goodman, "Some fundamental properties of speckle," *JOSA*, vol. 66, no. 11, pp. 1145–1150, 1976.