A Speckle Filtering Method Based on Hypothesis Testing for Time-Series SAR Images

Jili Yuan, Xiaolei Lv, and Rui Li

1 Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China; yuanjili17@mails.ucas.ac.cn (J.Y.);
liru315@mails.ucas.ac.cn (R.L.)
2 Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China
3 School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: academism2017@sina.com; Tel.: +86-131-2666-4001

Received: 4 July 2018; Accepted: 29 August 2018; Published: 30 August 2018

Abstract: To improve the suppression effect for the speckle noise of synthetic aperture radar (SAR) images and the ability of spatiotemporal information preservation of the filtered image without losing the spatial resolution, a novel multitemporal filtering method based on hypothesis testing is proposed in this paper. A framework of a two-step similarity measure strategy is adopted to further enhance the filtering results. Firstly, bi-date analysis using a two-sample Kolmogorov-Smirnov (KS) test is conducted in step 1 to extract homogeneous patches for 3-D patch stacks generation. Subsequently, the similarity between patch stacks is compared by a sliding time-series likelihood ratio (STSLR) test algorithm in step 2, which utilizes the multi-dimensional data structure of the stacks to improve the accuracy of unchanged pixels detection. Finally, the filtered values are obtained by averaging the similar pixels in time-series. The experimental results and analysis of two multitemporal datasets acquired by TerraSAR-X show that the proposed method outperforms the other typical methods with regard to the overall filtering effect, especially in terms of the consistency between the filtered images and the original ones. Furthermore, the performance of the proposed method is also discussed by analyzing the results from step 1 and step 2.

Keywords: multitemporal SAR images filtering; hypothesis test; similarity measurement

1. Introduction

A synthetic aperture radar (SAR) is a kind of active microwave imaging radar capable of imaging all day without relying on the light source. In addition, it has the advantage of penetrating clouds and rain because of the microwave band, which is able to achieve all-weather imaging. These advantages make SAR images an important data source in remote sensing applications and they are widely used in forestry monitoring [1], geological prospecting [2], disaster assessment [3], and other military or civilian fields [4–6]. However, due to the principle that SAR imaging is based on the coherent superposition of radar echoes reflected by a large amount of randomly distributed scattering, it is inevitable that a multiplicative random noise called speckle will be generated in SAR images [7–9]. The speckle noise seriously reduces the visual quality of images and makes both manual and automatic interpretation difficult to implement effectively. Therefore, the suppression of speckle is an important part in SAR images processing, which is of great significance for improving the image quality and finally promoting the wider application of SAR images.

To effectively reduce and remove the influence of speckle noise, many filtering methods based on a single SAR image have been proposed. In general, they can be categorized into three methods:
(i) multi-look processing method; (ii) spatial filtering method; and (iii) transform domain filtering method. The first approach utilizes multiple sub-views generated by dividing the Doppler bandwidth of the synthetic aperture before imaging to form a multi-look image, or averages the adjacent individual pixels after imaging to obtain the filtered image. Since the multi-look method does not take into account the statistical characteristics of speckle noise, it is simple to implement. However, this technique loses a great deal of detail information and seriously reduces the resolution of the image while improving the quality of the image. With the aid of mature estimation theory, the spatial domain filtering method [10–14] has been widely studied. The second method relies on the statistical model of the speckles and scenes, and many classical algorithms have been proposed, such as Lee’s filtering [15], Kuan’s filtering [16], Frost filtering [17], and plenty of their enhanced methods [18,19]. Touzi made a comprehensive summary and analysis of these methods in [20] and proposed effective approaches for the steady-state and unsteady multiplicative speckle noise models. Most of the aforementioned filtering methods are pixel-based, and the adaptive spatial filtering is performed using a strategy that is adaptive to the coefficient variation (CV) in the local window of the pixel. Although such methods can smooth off the speckle noise and maintain the details of the image in homogeneous regions, they will reduce the resolution of the image within the sliding window in the heterogeneous areas, for example, the boundaries and edges are somewhat blurred. The last kind of method is based on the transform domain, due to the favorable performance of wavelet and other multi-scale transformations in suppressing the noise of common optical images, and many transform domain filtering methods for SAR images based on different transforms have been proposed successively. For instance, the Contourlet transform-based method [21], Shearlet transform-based method [22], wavelet-based method [23,24], etc. By adjusting the transform coefficients, a better denoised result can be obtained, and there is less loss of edge information.

Apart from the abovementioned methods, there is a kind of method called non-local mean (NLM) filtering proposed by Buades et al. [25] in 2005, which was a breakthrough compared to the traditional local denoising models. The NLM algorithm fully considers the similarity and redundancy of natural image content and applies it to the image filtering, which obtains the filtered value by considering all the points being in a similar context. Deledalle et al. [10] extended the NLM to SAR images by using the generalized likelihood ratio (GLR) and Kullback-Leibler (KL) divergence to measure the similarity between SAR patches. Parrilli et al. [26] proposed the SARBM3D method based on the idea of block matching and collaborative filtering, which refines the results by taking advantage of the local linear minimum mean square in the wavelet domain.

Although many methods have been put forward for single-date SAR image filtering, there is always a loss of spatial resolution or insufficient detail information due to the limitation of only using spatial information. With the rapid increase of massive SAR data and large amounts of historical archive data in recent years, the temporal resolution of SAR images is getting higher and higher, which provides a good environment for the proposal of multi-date filtering methods. Using multi-date images, we can take advantage of both temporal and spatial information to improve the filtering results without losing the details and edge information [27]. Expanding the spatial filtering method to 3-D accounts for a large part of the multi-date filtering method. By estimating the local statistics of the pixel block sequence located at the same position in the time dimension, Bruniquel and Lopes [28] presented an approach that exploited the Kuan’s filter to obtain the filtered time-series images. Coltuc et al. [29] transformed the time dimension of the multi-date images into the frequency domain by using a discrete cosine transform and then adaptively filtered each frequency component to reduce the noise interference, and finally the filtered images were obtained through an inverse transform. In [30,31], a method is proposed using the mean of all the spatial windows in the same position of time-series to reduce the variance by weighted summation and obtain the filtered pixels.

Most multitemporal filtering methods can realize the suppression for speckle noise without the loss of spatial resolution, but they assume that the spatial pixels remain unchanged over time, which makes them simply suitable for short-term multi-date images with homogeneous areas.
However, the ground objects may change frequently due to human or natural influences in most actual situations, so an algorithm was proposed to improve the application scope of traditional multi-date filtering methods by finding out the similar pixels in time-series [32]. In addition, Su et al. [33] and Chierchia et al. [34] have proposed two kinds of multitemporal SAR filtering approaches combined with non-local mean idea and time-series images, respectively, which achieved outstanding results in SAR image denoising. One of the most important parts in the above methods is the selection and design of the similarity measurement method, which determines the accuracy of the last-found unchanged points. Nevertheless, the similar measurement called the CV test used in [32] was followed by the spatial adaptive filtering method, and it is not well applied to high-dimensional data. To solve those problems, a multi-date filtering approach considering the data structure should be developed, indicating a better suitability for long-term time-series data with changeable areas.

In this paper, we propose a novel method called the sliding time-series likelihood ratio (STSLR) test that is appropriate for the similarity measurement of the 3-D case, which makes full use of the multi-dimensional structure and the characteristics of time-series data to enhance the accuracy of the similar points search and reduce the misdetection rate. Firstly, the similarity comparison between patches is conducted by bi-date analysis using a two-sample KS test in order to generate the 3-D patch stacks. Secondly, an STSLR test is used to further improve the detection rate of similar points. Finally, the statistical mean is calculated according to the similar pixels obtained along the time axis.

This paper is organized as follows: Section 2 presents the overall framework of the proposed multitemporal filtering method in detail. In Section 3, the experimental results and analysis based on simulated data and real SAR datasets are shown to verify the effectiveness of the proposed method. Section 4 discusses the results. Finally, the conclusions are presented in Section 5.

2. Materials and Methods

In comparison with traditional multitemporal speckle filtering methods, which are usually applied under the assumption that the pixels at the same coordinate position remain unchanged along the time axis, we propose a novel similarity measurement approach based on multitemporal data to extract similar pixels over time and finally improve the effect of the smoothness for speckle noise. In general, similarity measurements are used to describe the similarity between two vectors, and such measures are often converted to inverse distance metrics, which means a zero value for very similar vectors. Apart from that, image matching and retrieval in computer vision also utilizes the similarity measure to analyse the distance between the feature vectors that are extracted from respective images. In this study, a multitemporal filtering methodology is proposed taking advantage of the idea of similarity measurement. It searches the highly similar points of every pixel at the same spatial position in the time-series and then applies this to multitemporal smoothing.

In order to make use of the 3-D data structure and enhance the performance of filtering results, we use a framework of a two-step filtering strategy. A different method is employed according to the data form at the first and second step, respectively. The workflow of the overall algorithm is depicted in Figure 1, which illustrates the procedure and the role of each part in the proposed multitemporal filtering method.
Figure 1. The processing flow of the proposed filtering method.

2.1. Data Preprocessing

For the purpose of achieving multitemporal speckle reduction, some basic SAR image preprocessing operations should first be carried out to avoid errors as much as possible, which mainly includes radiometric calibration and co-registration. First, because of the imaging peculiarity of the SAR data, such as radiation distortion, it should be calibrated radiometrically [35] in order to eliminate the difference between the observed value by sensors and the radiance of targets, so that the values of the pixels can represent the real backscatter strength of the terrestrial surface. Calibration is important for the quantitative application of SAR images. Then, co-registration is another fundamental procedure for multitemporal filtering and change detection, which aims to ensure that the position of the same ground objects within images taken at different times is identical. Wrong registration will not only lead to a serious impact on the filtering result, but also a lot of pseudo change points, so the registration precision usually demanded is a sub-pixel level [36]. In this paper, we use the GAMMA [37] software developed by Charles Werner and Urs Wegmüller in Gümligen, Switzerland to realize the radiometric calibration and co-registration of SAR images.

2.2. Two-Step Similarity Measure Strategy

In this section, a two-step strategy [32] is adopted for optimizing the process of detection for similar points, and a specific similarity measure algorithm is applied to corresponding steps according to the analysis of the data structure. Lastly, the filtered image is obtained by using the results from a previous similarity comparison.

2.2.1. Framework of the Two-Step Strategy

Given n SAR images collection $I = \{ I_j; 1 \leq j \leq n \}$ which is co-registered in advance, $j$ is the time position in the sequence where the image is located and relates to the data acquisition time. Let $I_j(x, y)$ be the pixel value located at the coordinate $(x, y)$ in the $j$th image $I_j$, and let $I^w_j(x, y)$ be the
sliding boxcar window surrounding the central point, where \( w \) denotes the window size. Then, we can get a pixel vector \( \{ I_j(x, y); 1 \leq j \leq n \} \) and a set of patches \( \{ I^\mu_j(x, y); 1 \leq j \leq n \} \) that are extended according to the time dimension. With that, a two-step similarity analysis approach is used to enhance the filtering performance in consideration of the collection of images.

- **Step 1—Bi-date analysis**

  The similar patch extraction, which means finding unchanged pixels in the pixel vector, is conducted by a pairwise similarity comparison. Let us assume that \( M \) is the method that measures the similarity between any two dates, for example, \( j, k \) in the patch stack; then, the similarity degree can be obtained by

  \[
  S^w_{jk}(x, y) = M\left( I^w_j(x, y), I^w_k(x, y) \right), \quad 1 \leq j, k \leq n. \tag{1}
  \]

  Thus, for \( n \)-length time-series SAR images, the similar degree at the position \((x, y)\) can be constructed as a two-dimension matrix with the size of \( n \times n \) in accordance with the time order, which is called the similarity measure matrix and described as follows

  \[
  S^w(x, y)_{n \times n} = \begin{bmatrix}
  S^w_{1,1}(x, y) & S^w_{1,2}(x, y) & \cdots & S^w_{1,n}(x, y) \\
  S^w_{2,1}(x, y) & S^w_{2,2}(x, y) & \cdots & S^w_{2,n}(x, y) \\
  \vdots & \vdots & \ddots & \vdots \\
  S^w_{n,1}(x, y) & S^w_{n,2}(x, y) & \cdots & S^w_{n,n}(x, y)
  \end{bmatrix}, \tag{2}
  \]

  where whether the matrix \( S^w(x, y)_{n \times n} \) is symmetrical depends on the measure method. To distinguish the unchanged pixels from the changed ones in the pixel pile, a threshold \( \lambda \) needs to be introduced. Given the threshold, the similarity measure matrix will be transformed into a binary matrix that only contains the value of 0 and 1.

  \[
  D^w_{jk}(x, y) = \begin{cases}
  0, S^w_{jk}(x, y) \leq \lambda \\
  1, S^w_{jk}(x, y) > \lambda
  \end{cases}, \quad 1 \leq j, k \leq n, \tag{3}
  \]

  where 0 indicates that there are no significant changes between \( I^w_j(x, y) \) and \( I^w_k(x, y) \), and on the contrary, the pixel at time \( j \) is dissimilar to the one at time \( k \) if the result is equal to 1.

- **Step 2—Multi-date analysis**

  After above bi-date analysis, we can obtain a collection \( \Phi\left(I^w_j(x, y)\right) \) including all the similar patches of \( I^w_j(x, y) \). In order to obtain pixels with a higher similarity precisely, this step substitutes the set \( \Phi\left(I^w_j(x, y)\right) \) constructed by temporal neighborhoods for the previous patch form as the comparative object. Hence, the similarity degree between two sets is

  \[
  S^w_{jk}(x, y) = M'\left( \Phi\left(I^w_j(x, y)\right), \Phi(I^w_k(x, y)) \right), \quad 1 \leq j, k \leq n, \tag{4}
  \]

  and by comparing it with the threshold \( \lambda' \), the new similarity measure matrix is as follows

  \[
  D^w_{jk}(x, y) = \begin{cases}
  0, S^w_{jk}(x, y) \leq \lambda' \\
  1, S^w_{jk}(x, y) > \lambda'
  \end{cases}, \quad 1 \leq j, k \leq n, \tag{5}
  \]

  In this step, the main goal is to improve the result of step 1 by utilizing the information of multi-date neighborhoods. The matrix has the same pattern as that in step 1, and the greatest difference between the two steps is the similarity measurement method.
2.2.2. Two-Sample Kolmogorov-Smirnov Test for Homogeneous Patch Extraction

The Kolmogorov-Smirnov (KS) test is a form of hypothesis testing that inspects the equality of the distribution of the dataset. It has two forms of test for specific applications, which are respectively named as the one-sample KS test and two-sample KS test. The former is usually applied to the comparison between the sample and a proper reference probability distribution as a goodness-of-fit testing approach, and the latter is used to determine if the distributions of two sample data are significantly different. Here, we exploit the two-sample KS test as the similarity measurement approach in step 1, which is denoted by $M$ in Equation (1). For the two-sample case, the test statistic is defined as

$$D_{KS} = \sup_{-\infty < x < \infty} |F_{1,l_1}(x) - F_{2,l_2}(x)|,$$

(6)

where $F_{1,l_1}(x)$ and $F_{2,l_2}(x)$ represent the empirical distribution functions of scattering intensity in two SAR patches, respectively; $l_1, l_2$ are the pixel number of a patch related to the window size; and sup denotes the supremum function.

The KS statistic takes the maximum distance between the empirical distribution functions as a measure of dissimilarity, and the null hypothesis ($H_0$) can be described as follows: the two patches located at the same position while acquired at different dates are statistically homogenous. ($H_0$) is accepted at the significant level $\alpha$ if

$$D_{KS} \leq \sqrt{\frac{1}{2} \ln \left(\frac{\alpha}{2}\right)} \sqrt{\frac{l_1 + l_2}{l_1 l_2}}.$$

(7)

In other words, the two patches are similar when the test statistic does not exceed the critical value. Equation (7) plays the role of finding out the homogenous patch and the right side of it is equivalent to the adaptive threshold $\lambda$ in (3). Therefore, Equation (3) can be rewritten as

$$D^w_{j,k}(x, y) = \begin{cases} 
0, & S^w_{j,k}(x,y) \leq -\frac{1}{2} \ln \left(\frac{4}{\alpha}\right) \sqrt{\frac{l_1 + l_2}{l_1 l_2}} \Leftrightarrow I^w_{j,k}(x,y) \in \Phi \left(I^w(x,y)\right), \\
1, & S^w_{j,k}(x,y) > -\frac{1}{2} \ln \left(\frac{4}{\alpha}\right) \sqrt{\frac{l_1 + l_2}{l_1 l_2}} \Leftrightarrow I^w_{j,k}(x,y) \notin \Phi \left(I^w(x,y)\right) \end{cases}, \quad 1 \leq j, k \leq n,$$

(8)

where $\Phi \left(I^w(x,y)\right)$ denotes the $j$th patch stack that is constructed by the patches similar to $I^w_{j,k}(x,y)$, and after the KS test, we can obtain $n$ patch stacks that are the inputs for the STSLR test.

2.2.3. Sliding Time-Series Likelihood Ratio Test for Similarity Analysis of Patch Stacks

It has been previously mentioned that a collection of homogeneous patches is built with the bi-date analysis using a two-sample KS test; more precisely, the patch stack is generated at each date in the same spatial coordinate, which means $n$ patch stacks for $n$-length images. However, the two-sample KS test is not suitable for the multi-date analysis in step 2 of the two-step similarity measure framework due to the element form of the patch stack, which will decrease the accuracy and reliability of the ultimate similar pixels. Therefore, a novel method called the STSLR test is proposed to analyze the similarity between any two patch stacks and to refine the results from the last step, which is appropriate for the processing of a three-order tensor. The idea is inspired by the joint-scatterer interferometric SAR (JSInSAR) technology [38], where the similarity between two JS tensors located at different positions is judged by the time-series likelihood ratio (TSLR) test.

As shown in Figure 2, there are two cases in the process of multitemporal analysis. In the first case, two comparative patch stacks have the same length, and each stack can be represented by

$$P(\ell) = [x_1(\ell), x_2(\ell), \ldots, x_m(\ell)],$$

(9)
which is a three-order tensor that consists of \( m \) patches, where \( x_i(\ell) \) is the \( i \)-th patch centered at the \( \ell \)-th pixel, and \( m \) is the length of a patch stack obtained from the KS test. Before evaluating whether two equal-length patch stacks are statistically homogeneous, a measurement criterion between the pair of patches that can be extended to a higher dimension should first be established. Considering two patches in the same order denoted by \( x_{1,i} \) and \( x_{2,j} \), respectively, the pair \( (x_{1,i}, x_{2,j}) \) is similar when they are generated from the probability distribution function (PDF) with uniform parameters, which can be interpreted as hypothesis testing given by

\[
\begin{align*}
H_0 & : \theta_1 = \theta_2 = \theta_{12}, \\
H_1 & : \theta_1 \neq \theta_2,
\end{align*}
\]

where \( \theta_i \) is the parameter of PDF and \( 1 \leq i \leq m \). Based on the Neyman-Pearson lemma [39], the likelihood ratio test statistic is the optimal criterion for the measurement:

\[
\Lambda^{(i)}(x_{1,i}, x_{2,j}) = \frac{L(\theta_{12}|x_{1,i}, x_{2,j}; H_0)}{L(\theta_1, \theta_2|x_{1,i}, x_{2,j}; H_1)}
\]

where the similarity degree between \( x_{1,i} \) and \( x_{2,j} \) is proportional to the value of \( \Lambda(x_{1,i}, x_{2,j}) \). Drawing on the experience from the two-sample KS test, the definition of the TSLR test statistic [38] is as follows

\[
D_{TSLR} = \sup \left| -2 \log \Lambda^{(i)}(x_{1,i}, x_{2,j}) \right|,
\]

where \( -2 \log \Lambda^{(i)} \) represents i.i.d. random variables, and the \( D_{TSLR} \) is the supremum of the absolute value of the likelihood ratio test statistics \( -2 \log \Lambda^{(i)}(x_{1,i}, x_{2,j}) \), where each \( \Lambda^{(i)}(x_{1,i}, x_{2,j}) \) represents the similarity between a pair of patches.

![Figure 2](image_url). Two cases in multi-date analysis using the sliding time-series likelihood ratio (STSLR) test.

Generally, the parameter \( \theta_i \) of a PDF is unavailable, i.e., the hypothesis is not simple. In these situations, we usually use the estimated parameters obtained by the maximum likelihood estimation (MLE) method, which is called the GLR test. Considering that the research site in the experimental data is mainly an urban area and contains a rich variety of ground features, we use the lognormal distribution as the PDF of the likelihood function, which is a good statistical representation of clutter in the case of heterogeneous terrain.

We assume that the amplitudes of scatterers in two patches \( (x_{1,i}, x_{2,j}) \) follow a lognormal distribution with parameters \( (\mu, \eta) \), i.e., \( x_{1,i} \sim f_1(x; \mu_{1,i}, \eta_{1,i}) \) and \( x_{2,j} \sim f_1(x; \mu_{2,j}, \eta_{2,j}) \), \( l = 1, \ldots, \omega_0 \), where \( \omega_0 \) is the number of scatterers in each patch. Under \( H_0 \), the hypothesis test can be redescribed as: \( \eta_{1,i}^2 = \eta_{2,j}^2 = \eta_{12,l}^2 \). According to the MLE method, the estimated value of \( \eta_{12,l}^2 \) is as follows

\[
\hat{\eta}_{12,l}^2 = \frac{\hat{\eta}_{1,i}^2 + \hat{\eta}_{2,j}^2}{2},
\]

with the same length part in the longer stack. In the case of heterogeneous terrain, the parameter \( \theta_i \) is obtained from the two-sample KS test, the definition of the TSLR test statistic is as follows

\[
D_{TSLRD} = \sup \left| -2 \log \Lambda^{(i)}(x_{1,i}, x_{2,j}) \right|,
\]

where \( -2 \log \Lambda^{(i)} \) represents i.i.d. random variables, and the \( D_{TSLRD} \) is the supremum of the absolute value of the likelihood ratio test statistics \( -2 \log \Lambda^{(i)}(x_{1,i}, x_{2,j}) \), where each \( \Lambda^{(i)}(x_{1,i}, x_{2,j}) \) represents the similarity between a pair of patches.
where \( \hat{\eta}_1^2 \) and \( \hat{\eta}_2^2 \) are the MLEs of parameters \( \eta_1^2 \) and \( \eta_2^2 \), respectively. Therefore, the GLR becomes

\[
\Lambda^{(i)}(x_{1,j}, x_{2,j}) = \left( \frac{\hat{\eta}_1^2}{\tilde{\eta}_1^2} \right)^{w_2} \left( \frac{\hat{\eta}_2^2}{\tilde{\eta}_2^2} \right)^{w_0} \left( \frac{\hat{\eta}_1^2}{\tilde{\eta}_1^2} \right)^{w_0}.
\]  

(14)

Finally, the modified \( D_{TLSR} \) is given by

\[
D_{TLSR} = \sup \left| -2\log \left( \frac{\hat{\eta}_1^2}{\tilde{\eta}_1^2} \right)^{w_2} \left( \frac{\hat{\eta}_2^2}{\tilde{\eta}_2^2} \right)^{w_0} \left( \frac{\hat{\eta}_1^2}{\tilde{\eta}_1^2} \right)^{w_0} \right|.
\]  

(15)

According to the Wilks’s theorem [40], the null distribution of \(-2\log \Lambda^{(i)}\) trends towards a chi-square distribution with degrees of freedom equal to \( d \) as the sample size approaches infinity, so the cumulative distribution function (CDF) of \( D_{TLSR} \) can be derived as follows

\[
F(D_{TLSR} \leq t) = F\left( -2\log \Lambda^{(1)} \leq t, -2\log \Lambda^{(2)} \leq t, \ldots, -2\log \Lambda^{(m)} \leq t \right)
= F_{\chi^2}(t; d)^m,
\]

(16)

where \( m \) denotes the length of the patch stack, and the PDF is just the derivative of the CDF. Thus, the null hypothesis that two patch stacks \( P_1(\ell) \) and \( P_2(\ell) \) are considered to be similar is accepted if \( D_{TLSR} \leq C \), where \( C \) is the threshold determined by the equation \( F(D_{TLSR} \leq C) = 1 - \alpha \).

Nevertheless, the length between two patch stacks is different in most situations, where the TSLR test is no longer applicable. As a result, the sliding TSLR is proposed for the second case. Given two stacks \( h_1 \) and \( h_2 \), we use the shorter stack \( h_1 \) as the sliding object. Then, the problem will be transformed into an integration of case 1 by comparing \( h_1 \) with the same length part in the longer stack \( h_2 \), and the test statistic of STSLR is written as

\[
D_{STSLR} = \sup \left| D_{TLSR}^{(i)} \right|, 0 \leq i \leq (h_2 - h_1 + 1),
\]

(17)

where each \( D_{TLSR}^{(i)} \) represents the similarity between equal length stacks, and the STSLR test statistic is the maximum TSLR statistic. The distribution of \( D_{STSLR} \) is identical to the TSLR test statistic and the threshold \( C \) can be used to distinguish between the dissimilarity and similarity.

2.3. Multitemporal Speckle Filtering

After two-step similarity measurement, the pixels with a high similarity can be found in time-series more accurately. The filtering procedure depends on the result of multi-date analysis, which includes the information of unchanged pixels for every point. To obtain the final filtered images, we use the statistical mean to deal with the similar pixels

\[
\hat{I}_j(x, y) = \frac{1}{N_{\Phi_j(x,y)}} \sum_{k=1}^{N_{\Phi_j(x,y)}} \left( 1 - \hat{D}^w_j(x, y) \right) I_k(x, y),
\]

(18)

where \( N_{\Phi_j(x,y)} \) is the number of elements contained in the collection \( \Phi \left( I^w_j(x, y) \right) \).

3. Experiments and Results

In this section, the proposed algorithm will be compared with the other four speckle filtering methods in order to evaluate its feasibility and effectiveness, two of which are classical multitemporal filtering methods Quegan’s filter (QF) [31] and change detection matrix based filter (CDMF) [32],
and the remaining two are the state-of-the-art in single-temporal and multi-temporal speckle filtering approaches called fast adaptive nonlocal SAR despeckling (FANS) [41] and multitemporal SAR block-matching and 3D filtering (MSAR_BM3D) [34], respectively. Due to the lack of a noise-free reference image, the performance assessment in SAR image despeckling usually depends on the original or filtered images. To evaluate the performance of each method more comprehensively and accurately, we carried out the experiments on real multitemporal SAR images and synthetic SAR images by using the above approaches. The optimal parameters of each experiment were obtained by the trial-and-error method. The setting of parameters in each method is as follows: (i) the local window size of QF was set as $3 \times 3$; (ii) in the CDMF approach, the window size was $3 \times 3$ and $\eta = 0.95$; (iii) the parameters of the FANS and MSAR_BM3D were set as suggested in the reference paper; and (iv) the parameters of the proposed method were set to $\alpha_{KS} = 0.05$ and $\alpha_{STSLR} = 0.05$, where the window size was $3 \times 3$.

### 3.1. Comparison on Synthetic Multitemporal Synthetic Aperture Radar Images

In this simulated experiments, the synthetic multitemporal images are generated by multiplying the optical images with simulated speckle noise, and the noise obeys the Rayleigh distribution [42]. To evaluate the quantitative performance of each method, we adopted the indices of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) to measure the degree of similarity between the distorted image and reference image. Firstly, we generated two sets of eight SAR-like images with 1-Look speckle noise to simulate time-series SAR images. The first set of images did not change over time. Considering that the simulation of temporal change helps to better analyze the effects of each approach, we added three dark lines to the first image of the second set. In addition, we added eight images to evaluate the impact of the increased number of images on the simulation results. Finally, we added a group of experiments based on 4-Looks speckle noise to assess the performance of different approaches from multi-aspects. The filtering results of the local region of two sets with eight images and 1-Look speckle are shown in Figures 3 and 4. The 4-Looks results are shown in Figures 5 and 6. Tables 1 and 2 present the quantitative results based on a different number of looks of speckle noise, respectively.

![Figure 3](image-url)  
**Figure 3.** Filtering results of the image “Hill” corrupted by simulated speckle (first scene with unchanging and 1-Look noise): (a) Reference image (noise-free); (b) Noisy with single-look; (c) Fast adaptive nonlocal SAR despeckling (FANS); (d) Change detection matrix based filter (CDMF); (e) Quegan’s filter (QF); (f) Multitemporal SAR block-matching and 3D filtering (MSAR_BM3D); (g) Proposed.
that the exploitation of temporal information helps improve the filtering effect. Then, we can compare the proposed method to MSAR_BM3D in the use of temporal information, and our proposed method gains up to 3.11 dB and 0.099 on PSNR and SSIM, respectively, which indicates the robustness of the similarity measurement in this paper. The first dataset consists of 13 single-look SAR images acquired between 28 March 2012 and 8 March 2016. It covers a lake island, factory buildings, crops, etc. Figure 7c,d show the dataset and the optical image.

### Table 1

<table>
<thead>
<tr>
<th>Image Number</th>
<th>Noisy with single-look</th>
<th>Changing and 1-Look noise</th>
<th>Changing and 4-Looks noise</th>
<th>Unchanging and 1-Look noise</th>
<th>Unchanging and 4-Looks noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 (change)</td>
<td>19.54–0.573</td>
<td>26.07–0.830</td>
<td>22.21–0.719</td>
<td>24.69–0.827</td>
<td>26.94–0.905</td>
</tr>
<tr>
<td>8 (unchange)</td>
<td>19.41–0.504</td>
<td>26.52–0.791</td>
<td>22.89–0.678</td>
<td>25.74–0.802</td>
<td>31.78–0.924</td>
</tr>
<tr>
<td>16 (change)</td>
<td>19.54–0.573</td>
<td>26.07–0.830</td>
<td>23.69–0.795</td>
<td>25.32–0.857</td>
<td>26.74–0.912</td>
</tr>
<tr>
<td>16 (unchange)</td>
<td>19.41–0.504</td>
<td>26.52–0.791</td>
<td>24.98–0.764</td>
<td>26.91–0.841</td>
<td>33.45–0.946</td>
</tr>
<tr>
<td>16 (change)</td>
<td>22.08–0.454</td>
<td>29.77–0.666</td>
<td>27.66–0.569</td>
<td>29.22–0.677</td>
<td>30.76–0.779</td>
</tr>
<tr>
<td>16 (unchange)</td>
<td>22.21–0.454</td>
<td>29.87–0.666</td>
<td>29.08–0.569</td>
<td>29.59–0.677</td>
<td>31.10–0.779</td>
</tr>
</tbody>
</table>

### Figure 4
Filtering results of the image “Hill” corrupted by simulated speckle (first scene with changing and 1-Look noise): (a) Reference image (noise-free); (b) Noisy with single-look; (c) FANS; (d) CDMF; (e) QF; (f) MSAR_BM3D; (g) Proposed.

### Figure 5
Filtering results of the image “Hill” corrupted by simulated speckle (first scene with unchanging and 4-Looks noise): (a) Reference image (noise-free); (b) Noisy with single-look; (c) FANS; (d) CDMF; (e) QF; (f) MSAR_BM3D; (g) Proposed.

### Figure 6
Filtering results of the image “Hill” corrupted by simulated speckle (first scene with changing and 4-Looks noise): (a) Reference image (noise-free); (b) Noisy with single-look; (c) FANS; (d) CDMF; (e) QF; (f) MSAR_BM3D; (g) Proposed.
Table 1. The quantitative results of different methods on synthetic aperture radar (SAR) images (peak signal-to-noise ratio (PSNR) and structural similarity (SSIM), PSNR-SSIM).

<table>
<thead>
<tr>
<th>Image Number</th>
<th>Noisy (1-Look)</th>
<th>FANS</th>
<th>CDMF</th>
<th>QF</th>
<th>MSAR_BM 3D</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 (unchange)</td>
<td>14.07–0.240</td>
<td>22.51–0.517</td>
<td>17.01–0.397</td>
<td>19.50–0.554</td>
<td>28.07–0.845</td>
<td>21.35–0.658</td>
</tr>
<tr>
<td>8 (change)</td>
<td>14.21–0.299</td>
<td>22.09–0.563</td>
<td>16.63–0.446</td>
<td>19.33–0.599</td>
<td>23.26–0.774</td>
<td>20.72–0.680</td>
</tr>
<tr>
<td>16 (unchange)</td>
<td>14.07–0.240</td>
<td>22.51–0.517</td>
<td>18.58–0.508</td>
<td>20.37–0.612</td>
<td>29.57–0.887</td>
<td>24.46–0.757</td>
</tr>
<tr>
<td>16 (change)</td>
<td>14.21–0.299</td>
<td>22.09–0.563</td>
<td>18.02–0.522</td>
<td>19.99–0.654</td>
<td>23.82–0.817</td>
<td>22.33–0.768</td>
</tr>
</tbody>
</table>

Table 2. The quantitative results of different methods on synthetic SAR images (PSNR-SSIM).

<table>
<thead>
<tr>
<th>Image Number</th>
<th>Noisy (4-Looks)</th>
<th>FANS</th>
<th>CDMF</th>
<th>QF</th>
<th>MSAR_BM3D</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 (unchange)</td>
<td>19.41–0.504</td>
<td>26.52–0.791</td>
<td>22.89–0.678</td>
<td>25.74–0.802</td>
<td>31.78–0.924</td>
<td>28.24–0.857</td>
</tr>
<tr>
<td>8 (change)</td>
<td>19.54–0.573</td>
<td>26.07–0.830</td>
<td>22.21–0.719</td>
<td>24.69–0.827</td>
<td>26.94–0.905</td>
<td>26.62–0.878</td>
</tr>
<tr>
<td>16 (unchange)</td>
<td>19.41–0.504</td>
<td>26.52–0.791</td>
<td>24.98–0.764</td>
<td>26.91–0.841</td>
<td>33.45–0.946</td>
<td>31.10–0.915</td>
</tr>
<tr>
<td>16 (change)</td>
<td>19.54–0.573</td>
<td>26.07–0.830</td>
<td>23.69–0.795</td>
<td>25.32–0.857</td>
<td>26.74–0.912</td>
<td>27.84–0.917</td>
</tr>
</tbody>
</table>

From Table 1, we can see that the proposed method and MSAR_BM3D achieved the best results in the four groups of experiments. Take the unchanged case as an example. Since FANS is based on a single image, its indexes remain unchanged with the increase in the number of images, while the proposed method gains up to 3.11 dB and 0.099 on PSNR and SSIM, respectively, which indicates that the exploitation of temporal information helps improve the filtering effect. Then, we can compare the results of the first two sets in the table, and it can be observed that the PSNR of MSAR_BM3D drops dramatically when introducing the temporal change, close to 5 dB. This is mainly because the three dark lines added were smoothed out too much (see Figure 4f), resulting in the large difference between the filtered and reference image. However, the proposed method retains most of the temporal change, and the PSNR only decreased by 0.63 dB.

As for the case of 4-Looks speckle noise, it can be seen that the results shown in Figures 5 and 6 are much better than those of Figures 3 and 4, both in terms of quantitative indices and qualitative filtering effects due to noise reduction. Table 2 shows that the proposed method is better than the single-temporal method of FANS because of the exploitation of temporal information, and the denoising effect can be further improved with the increase of the number of images. In addition, the proposed method is more effective than MSAR_BM3D in the use of temporal information, and our method achieves the best performance in the changing case with 16 images.

3.2. Comparison on Real Multitemporal Synthetic Aperture Radar Images

3.2.1. Data Description

To assess the performance of the proposed method more accurately, two datasets of real SAR images acquired by TerraSAR-X from Tongzhou District of Beijing, China, are used as the experimental data. TerraSAR-X is a commercial SAR satellite operating in the X-band and launched in 2007, with a revisit period of 11 days. The highest spatial resolution of the spotlight and strip imaging modes is 1 and 3 m, respectively, and these modes are commonly used in earth observation tasks. As the sub-center of Beijing, Tongzhou District bears the future development of the capital and has been undergoing rapid changes in recent years. The complicated land cover changes in the demolition and reconstruction of urban buildings are conducive to verify the reliability and robustness of the similarity measurement in this paper. The first dataset consists of 13 single-look complex SAR images acquired from 22 January 2012 to 8 March 2016 with a size of 1000 × 1000 and a 2-m resolution, and the features included in this dataset are mainly residential areas, bridges, and rivers. The first SAR image of the time-series and the corresponding optical image are shown in Figure 7a,b. The second dataset has the same size and resolution as the first one, which contains 12 images acquired between 28 March 2012
and 8 March 2016. It covers a lake island, factory buildings, crops, etc. Figure 7c,d show the dataset and the optical image.

![Figure 7](image)

**Figure 7.** Two time-series datasets of TerraSAR-X remote sensing images in Tongzhou District of Beijing: (a) Dataset 1 (first scene); (b) Optical image; (c) Dataset 2 (1st scene); (d) Optical image.

### 3.2.2. Experimental Results and Analysis

The speckle filtering results of Dataset 1 are depicted in Figure 8. Considering the length of the article, we cannot list the filtering results of SAR images at all times. Therefore, the experimental results of a representative image in the time-series are given as the analysis object. Figure 8a is the original SAR image acquired on 22 January 2012 from Dataset 1, which is seriously corrupted by the speckle noise. Figure 8b is the filtering result using the FANS method. Although there is a great enhancement of the suppression effect for the speckle, the details information of the image displays a great loss, which is the defect of the non-temporal filter. For instance, the red ellipse in the upper right corner of the figure contains a bridge, and the strong scattering points on both sides of the bridge are the street lamps. It can be seen from Figure 8b that the lamps are too blurred compared to the other multitemporal filtering approaches. Figure 8c shows the result of QF, where the smoothness for the noise is not obvious due to the change of the ground objects. Figure 8d gives the CDMF multitemporal filtering result, which has a smoother effect and better image detail preservation ability compared to the previous methods. However, there are some imaginary shadows of the buildings in red box 2 because of the error detection of similar points, which changes the surface feature information of the original image. Figure 8e presents the result of the MSAR_BM3D approach, for which we can see that the suppression effect of speckle noise is further improved, taking advantage of the temporal and non-local information. However, the brightness of the image displays a certain decline and the building areas present are somewhat over-smoothed, which leads to a little unnaturalness, such as the residential buildings in yellow box 1. The proposed method is shown in Figure 8f, where the smoothing effect has been largely improved and the naturalness and details of the filtered image are well preserved because of the high precision for the detection of similar pixels in the temporal axis, which maintains a high degree of consistency with the original image.
The preservation capability of strong point targets is an important qualitative evaluation index for measuring the performance of different filtering methods. In order to better compare the effect of each method in this ability, we cropped the bridge region within the ellipse in Figure 8 as the analytic object. It can be seen from Figure 9 that the filtering results of FANS and MSAR_BM3D have different degrees of ambiguity, while the other three approaches, including that proposed in this paper, show a better preservation ability for strong point targets.

To further assess the experimental results of each method, we introduce two quantitative evaluation indices: the equivalent number of looks (ENL) and mean bias (MB). The ENL is used to measure the degree of speckle reduction, which is calculated as $ENL = \frac{\mu}{\sigma^2}$, where $\mu$ and $\sigma$ represent the mean and standard deviation of a local homogeneous area, respectively, which is shown as a green square with about 50 × 50 pixels in Figure 8. Then, mean bias is expressed as $MB = \frac{\mu_B - \mu_I}{\mu_I}$.
where $\mu_I^1, \mu_I^2$ are the statistical mean of the whole image before and after the filtering operation. According to the product model and the statistical properties of speckle noise, the mean bias should be equal to 0 in the ideal case [20,42,43], so this index can characterize the information preservation ability of the filtering method to the feature of the ground object. The closer the value is to 0, the better the filter method maintains the average of the image, i.e., the less loss of image detail information. In order to better compare the difference of mean deviation for each method, a negative logarithm is applied, which is revised as $MB = -\log \left( \frac{|\mu_I^2 - \mu_I^1|}{\mu_I^1} \right)$. Thus, the larger the MB, the better the retention of image information for the method.

The quantitative results of images at every time obtained by five different algorithms are listed in Table 3 and representing ENL and MB, respectively. From Table 3, we can see that the higher ENL is achieved by the proposed method, FANS, and MSAR_BM3D. However, in Table 4, the mean biases of FANS and MSAR_BM3D are relatively worse than the other three algorithms due to the use of a non-local operation, which decreases the brightness and some details of images. Our method performs much better than QF and CDMF in mean bias because of the high accuracy of searching for similar points in the time dimension. Considering that the subsequent application of time-series SAR images is mainly in classification and change detection, the maintenance of temporal and detail information of images is more important. Therefore, the proposed method achieves better results.

Table 3. The quantitative comparison of different filtering results for Dataset 1, using equivalent number of looks (ENL).

<table>
<thead>
<tr>
<th>No.</th>
<th>Origin</th>
<th>FANS</th>
<th>CDMF</th>
<th>QF</th>
<th>MSAR_BM3D</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.92</td>
<td>25.17</td>
<td>4.26</td>
<td>4.38</td>
<td>37.48</td>
<td>9.81</td>
</tr>
<tr>
<td>2</td>
<td>0.87</td>
<td>16.43</td>
<td>3.97</td>
<td>3.63</td>
<td>17.73</td>
<td>7.13</td>
</tr>
<tr>
<td>3</td>
<td>0.89</td>
<td>11.91</td>
<td>4.13</td>
<td>3.30</td>
<td>13.60</td>
<td>7.53</td>
</tr>
<tr>
<td>4</td>
<td>0.99</td>
<td>22.76</td>
<td>5.26</td>
<td>4.15</td>
<td>51.44</td>
<td>9.23</td>
</tr>
<tr>
<td>5</td>
<td>0.96</td>
<td>26.16</td>
<td>4.08</td>
<td>4.00</td>
<td>39.55</td>
<td>8.01</td>
</tr>
<tr>
<td>6</td>
<td>0.93</td>
<td>18.72</td>
<td>5.45</td>
<td>4.18</td>
<td>24.09</td>
<td>8.77</td>
</tr>
<tr>
<td>7</td>
<td>0.85</td>
<td>23.26</td>
<td>4.89</td>
<td>4.25</td>
<td>25.54</td>
<td>7.13</td>
</tr>
<tr>
<td>8</td>
<td>0.94</td>
<td>22.99</td>
<td>4.59</td>
<td>4.20</td>
<td>28.26</td>
<td>9.15</td>
</tr>
<tr>
<td>9</td>
<td>0.85</td>
<td>16.18</td>
<td>3.00</td>
<td>3.25</td>
<td>13.07</td>
<td>5.83</td>
</tr>
<tr>
<td>10</td>
<td>0.84</td>
<td>13.87</td>
<td>3.86</td>
<td>3.45</td>
<td>14.26</td>
<td>8.66</td>
</tr>
<tr>
<td>11</td>
<td>0.92</td>
<td>20.10</td>
<td>4.86</td>
<td>4.49</td>
<td>21.50</td>
<td>8.80</td>
</tr>
<tr>
<td>12</td>
<td>0.93</td>
<td>21.80</td>
<td>5.31</td>
<td>4.16</td>
<td>25.65</td>
<td>8.92</td>
</tr>
<tr>
<td>13</td>
<td>0.88</td>
<td>15.80</td>
<td>5.05</td>
<td>3.85</td>
<td>20.93</td>
<td>9.08</td>
</tr>
<tr>
<td>Mean</td>
<td>0.91</td>
<td>19.63</td>
<td>4.52</td>
<td>3.95</td>
<td>25.62</td>
<td>8.31</td>
</tr>
</tbody>
</table>

Figure 10a,b intuitively show the variation trends of the filtering effect on all images for the proposed framework and another four methods. The horizontal axis indicates the change of the image acquisition time, and the different colored lines represent the results of different methods with time.

Figure 11 shows the filtering results of Dataset 2. The acquisition time of the representative image is 28 March 2012, which is the first scene in the time-series and shown in Figure 11a. Analogously, the isolated targets have been almost eliminated by FANS in the red elliptical area (see Figure 11b), whereas the other multitemporal algorithms still retain the information of those isolated objects. Figure 11c–e shows the results of QF, CDMF, and MSAR_BM3D, respectively. It can be seen that the speckle suppression effect of QF improves little; in the meantime, the results of CDMF for the factory buildings surrounding the vegetation in the lake island also display some erroneous traces of construction, and the MSAR_BM3D result shows some degradation of the image brightness and detail information within the red box, in spite of the high smoothness. The result of the proposed method is shown in Figure 11f, which presents a better comprehensive filtering effect than the other methods.
Table 4. The quantitative comparison of different filtering results for Dataset 1, using mean bias (MB).

<table>
<thead>
<tr>
<th>No.</th>
<th>FANS</th>
<th>CDMF</th>
<th>QF</th>
<th>MSAR_BM3D</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.1328</td>
<td>4.1384</td>
<td>4.2081</td>
<td>2.2018</td>
<td>4.8555</td>
</tr>
<tr>
<td>2</td>
<td>2.1115</td>
<td>4.1545</td>
<td>4.3856</td>
<td>2.1586</td>
<td>6.2961</td>
</tr>
<tr>
<td>3</td>
<td>2.1245</td>
<td>7.6344</td>
<td>4.3708</td>
<td>2.1888</td>
<td>7.7011</td>
</tr>
<tr>
<td>4</td>
<td>2.1118</td>
<td>4.5191</td>
<td>4.3884</td>
<td>2.1449</td>
<td>6.1204</td>
</tr>
<tr>
<td>5</td>
<td>2.1135</td>
<td>4.6101</td>
<td>4.3931</td>
<td>2.1544</td>
<td>5.4897</td>
</tr>
<tr>
<td>6</td>
<td>2.1187</td>
<td>5.6915</td>
<td>4.4240</td>
<td>2.1743</td>
<td>6.3087</td>
</tr>
<tr>
<td>7</td>
<td>2.1257</td>
<td>4.5390</td>
<td>4.4226</td>
<td>2.1876</td>
<td>5.4535</td>
</tr>
<tr>
<td>8</td>
<td>2.1150</td>
<td>4.8211</td>
<td>4.4041</td>
<td>2.1601</td>
<td>9.2219</td>
</tr>
<tr>
<td>9</td>
<td>2.1229</td>
<td>4.9170</td>
<td>4.4199</td>
<td>2.1813</td>
<td>5.4119</td>
</tr>
<tr>
<td>10</td>
<td>2.1132</td>
<td>4.1384</td>
<td>4.2081</td>
<td>2.2018</td>
<td>4.8555</td>
</tr>
<tr>
<td>11</td>
<td>2.1392</td>
<td>4.4409</td>
<td>4.3403</td>
<td>2.2000</td>
<td>5.3605</td>
</tr>
<tr>
<td>12</td>
<td>2.1396</td>
<td>4.7689</td>
<td>4.3288</td>
<td>2.1972</td>
<td>5.6655</td>
</tr>
<tr>
<td>Mean</td>
<td>2.1234</td>
<td>4.8493</td>
<td>4.3747</td>
<td>2.1763</td>
<td>6.1698</td>
</tr>
</tbody>
</table>

Figure 10. Change trends of the indices for Dataset 1: (a) ENL; (b) Mean bias (negative logarithm).

Figure 11. Filtering results for Dataset 2: (a) Original image (28 March 2012); (b) FANS; (c) QF; (d) CDMF; (e) MSAR_BM3D; (f) the proposed approach.
In order to analyze the ability of each method to preserve the point targets in Dataset 2, we selected a region that contains the strong point. It can be observed from Figure 12 that FANS and MSAR\textunderscore BM3D have the problem of over-smoothing, which reduced the brightness of the point after filtering. CDMF failed to maintain the target and there was a slight loss in the result of QF, while the proposed method achieved the best preservation capability for strong point targets.

![Figure 12](image-url)  
Figure 12. Preservation of strong target points for Dataset 2. (a) Original image (28 March 2012); (b) FANS; (c) QF; (d) CDMF; (e) MSAR\textunderscore BM3D; (f) the proposed approach.

Tables 5 and 6 list the quantitative indices of different multitemporal filtering approaches in detail based on Dataset 2. It can be observed that the performance of the proposed approach is stable and achieves the best overall results.

Table 5. The quantitative comparison of different filtering results for Dataset 2 (ENL).

<table>
<thead>
<tr>
<th>No.</th>
<th>Origin</th>
<th>FANS</th>
<th>CDMF</th>
<th>QF</th>
<th>MSAR\textunderscore BM3D</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.96</td>
<td>20.39</td>
<td>4.68</td>
<td>4.44</td>
<td>25.07</td>
<td>9.50</td>
</tr>
<tr>
<td>2</td>
<td>0.92</td>
<td>29.39</td>
<td>4.59</td>
<td>4.33</td>
<td>30.30</td>
<td>9.10</td>
</tr>
<tr>
<td>3</td>
<td>0.85</td>
<td>19.72</td>
<td>3.98</td>
<td>4.54</td>
<td>17.26</td>
<td>8.95</td>
</tr>
<tr>
<td>4</td>
<td>0.93</td>
<td>25.84</td>
<td>4.51</td>
<td>4.38</td>
<td>29.92</td>
<td>9.48</td>
</tr>
<tr>
<td>5</td>
<td>0.94</td>
<td>23.69</td>
<td>5.16</td>
<td>4.28</td>
<td>28.17</td>
<td>9.55</td>
</tr>
<tr>
<td>6</td>
<td>0.98</td>
<td>28.73</td>
<td>4.62</td>
<td>4.70</td>
<td>42.29</td>
<td>10.02</td>
</tr>
<tr>
<td>7</td>
<td>0.98</td>
<td>34.98</td>
<td>4.34</td>
<td>4.83</td>
<td>40.38</td>
<td>10.02</td>
</tr>
<tr>
<td>8</td>
<td>0.95</td>
<td>28.97</td>
<td>4.35</td>
<td>4.64</td>
<td>28.24</td>
<td>9.63</td>
</tr>
<tr>
<td>9</td>
<td>0.91</td>
<td>22.53</td>
<td>4.71</td>
<td>4.20</td>
<td>35.07</td>
<td>8.88</td>
</tr>
<tr>
<td>10</td>
<td>0.86</td>
<td>20.16</td>
<td>5.19</td>
<td>4.35</td>
<td>20.65</td>
<td>8.43</td>
</tr>
<tr>
<td>11</td>
<td>0.83</td>
<td>20.08</td>
<td>4.02</td>
<td>3.74</td>
<td>28.46</td>
<td>8.17</td>
</tr>
<tr>
<td>12</td>
<td>0.89</td>
<td>19.74</td>
<td>4.55</td>
<td>3.83</td>
<td>22.55</td>
<td>8.64</td>
</tr>
<tr>
<td>Mean</td>
<td>0.92</td>
<td>24.52</td>
<td>4.56</td>
<td>4.36</td>
<td>29.03</td>
<td>9.20</td>
</tr>
</tbody>
</table>

Finally, Table 7 gives the computational time of each algorithm which is run on Intel i7 CPU at 2.80 GHz and 8 GB RAM. It can be seen from the table that the QF method is the fastest due to its simple calculation. In the remaining approach, the FANS and MSAR\textunderscore BM3D take less time because of the mix of MATLAB and C language and the use of a lookup table. Our method and CDMF are pure MATLAB codes which decrease the speed of runtime, but this time is acceptable for filtering the time-series SAR images, and the parallel processing and use of C language will be considered in future studies to improve the execution time.
Table 6. The quantitative comparison of different filtering results for Dataset 2 (MB).

<table>
<thead>
<tr>
<th>No.</th>
<th>FANS (MB)</th>
<th>CDMF (MB)</th>
<th>QF (MB)</th>
<th>MSAR_BM3D (MB)</th>
<th>Proposed (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.1144</td>
<td>4.9023</td>
<td>4.3143</td>
<td>2.1656</td>
<td>5.1611</td>
</tr>
<tr>
<td>2</td>
<td>2.1070</td>
<td>5.1058</td>
<td>4.3916</td>
<td>2.1392</td>
<td>5.8887</td>
</tr>
<tr>
<td>3</td>
<td>2.1035</td>
<td>4.0422</td>
<td>4.4278</td>
<td>2.1328</td>
<td>5.6099</td>
</tr>
<tr>
<td>4</td>
<td>2.1115</td>
<td>4.5933</td>
<td>4.4314</td>
<td>2.1448</td>
<td>5.6091</td>
</tr>
<tr>
<td>5</td>
<td>2.1159</td>
<td>4.7306</td>
<td>4.4091</td>
<td>2.1763</td>
<td>8.2699</td>
</tr>
<tr>
<td>6</td>
<td>2.1126</td>
<td>5.3516</td>
<td>4.4419</td>
<td>2.1627</td>
<td>6.3969</td>
</tr>
<tr>
<td>7</td>
<td>2.1056</td>
<td>4.5890</td>
<td>4.4622</td>
<td>2.1347</td>
<td>5.4423</td>
</tr>
<tr>
<td>8</td>
<td>2.1074</td>
<td>4.8968</td>
<td>4.4688</td>
<td>2.1425</td>
<td>5.2757</td>
</tr>
<tr>
<td>9</td>
<td>2.1125</td>
<td>4.7376</td>
<td>4.4370</td>
<td>2.1718</td>
<td>5.0765</td>
</tr>
<tr>
<td>10</td>
<td>2.1182</td>
<td>4.7993</td>
<td>4.3955</td>
<td>2.1888</td>
<td>5.6864</td>
</tr>
<tr>
<td>11</td>
<td>2.1155</td>
<td>4.2997</td>
<td>4.3081</td>
<td>2.1578</td>
<td>5.6087</td>
</tr>
<tr>
<td>12</td>
<td>2.1154</td>
<td>4.6592</td>
<td>4.2752</td>
<td>2.1669</td>
<td>4.7396</td>
</tr>
<tr>
<td>Mean</td>
<td>2.1116</td>
<td>4.7256</td>
<td>4.3969</td>
<td>2.1570</td>
<td>5.7304</td>
</tr>
</tbody>
</table>

Table 7. Computational time (in seconds) for filtering the 1000 × 1000 × 13 SAR images.

<table>
<thead>
<tr>
<th>FANS</th>
<th>CDMF</th>
<th>QF</th>
<th>MSAR_BM3D</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>149 s</td>
<td>654 s</td>
<td>48 s</td>
<td>260 s</td>
<td>836 s</td>
</tr>
</tbody>
</table>

4. Discussion

Compared with other multitemporal filtering approaches, the method proposed in this paper can remarkably improve the filtering effect, mainly owing to two factors: (i) Adopting the strategy of a two-step similarity measure, which improves the accuracy of finding unchanged points; and (ii) according to the data pattern, two more appropriate methods and criterion of similarity measurement based on hypothesis testing are applied to further enhance the success rate of similar points detection and to simultaneously reduce the commission error. For the aforementioned reasons, our method has great potential for the suppression of speckle noise and the preservation of spatiotemporal information in complex and changeable regions.

To illustrate the process of searching unchanged points, the analysis results of a pixel within the building areas are shown in Figure 13, where the buildings were constructed on the penultimate image in Dataset 1. Figure 13a is the similarity measure matrix obtained by the two-sample KS test from the first step, and Figure 13b shows the result of the second step using the STSLR test, which includes the similarity information of all pixels on the time-series in this position. It can be observed that the first matrix contains some erroneous results considered to be dissimilar, while in the second one, the wrong detection results are eliminated completely. Figure 13c shows the filtered results of all images based on the first and second similarity measure matrix, respectively, which indicates the advantages of the two-step strategy combining the proposed hypothesis testing method.

By suppressing the speckle noise, the multitemporal filtered SAR images can restore the structural features, including the contours, edges, and textures of real geographical information, considerably improve the image quality, and enable better interpretation of the image, especially for the high resolution SAR images, providing more reliable data for subsequent multitemporal applications such as forest monitoring, land cover change detection, and target recognition.
5. Conclusions

The traditional speckle suppression methods for SAR images either have the problem of resolution and detail information loss, or changes of the spatiotemporal information of the filtered images. To solve these problems, a novel multitemporal filtering method based on hypothesis testing was proposed, which can not only suppress the noise without the loss of spatial resolution, but also ensure the consistency of the images before and after filtering as far as possible. To combine the hypothesis testing approaches and improve the filtering results, the strategy of a two-step similarity measurement was conducted. In step 1, the similar patches for each date were extracted according to the bi-date analysis using the two-sample KS test method. Then, in step 2, a new test method called the STSLR test was proposed to process the patch stacks generated by the first step, which exploited the peculiarity of the multi-dimensional structure to further enhance the filtering results. By averaging the unchanged pixels, the final filtered images were obtained. To demonstrate the performance of the proposed method, we carried out the experiments using synthetic multitemporal SAR data and two real time-series datasets in Tongzhou District of Beijing, China, which contained different land cover

Figure 13. The similarity measurement matrix from the two-step strategy of Dataset 1: (a) Step 1; (b) Step 2; (c) filtering results of a pixel within the changeable building areas.
types and were acquired from 2012 to 2016 by TerraSAR-X. Our method showed the best performance in terms of both the qualitative and quantitative results. However, there are still some deficiencies in this work that need to be improved. For example, the computational efficiency of the algorithm can be accelerated by parallel processing. Apart from that, more experiments with different scenes and resolutions can be performed to validate the effectiveness more adequately. In future research, we will consider combining the non-local method and our approach to further improve the filtering effect.

Author Contributions: J.Y. and X.L. conceived and designed the proposed method; J.Y. carried out the experiments; J.Y. and X.L. analysed the data and results; R.L. contributed materials and preprocessed the datasets; J.Y. wrote the paper.

Funding: This research was funded by the Hundred Talents Program under Grant [Y53Z180390] and was partly funded by the National Natural Science Foundation of China under Grant 61331017.

Acknowledgments: This work was supported by the Hundred Talents Program under Grant [Y53Z180390] and was partly supported by the National Natural Science Foundation of China under Grant 61331017. The authors would like to thank Giovanni Chierchia et al. and Cozzolino Davide et al. very much for sharing the executable codes of MSAR_BM3D and FANS on their homepage.

Conflicts of Interest: The authors declare no conflict of interest.

References
