Change Detection in Multitemporal SAR Images Based on Slow Feature Analysis Combined With Improving Image Fusion Strategy

Weisong Li , *Student Member, IEEE*, Xiayang Xiao, *Student Member, IEEE*, Penghao Xiao, *Student Member, IEEE*, Haipeng Wang , *Senior Member, IEEE*, and Feng Xu , *Senior Member, IEEE*

Abstract—Change detection in multitemporal synthetic aperture radar (SAR) images has been an important research content in the field of remote sensing for a long time. In this article, based on the slow feature analysis (SFA) theory and the nonsubsampled contourlet transform (NSCT) algorithm, we propose a novel unsupervised change detection method called NSCT nonlocal means (NSCT-NLM). The powerful extraction to the changed information of SFA and the superior information fusion of NSCT are jointly adopted in this method. The main framework consists of the following parts. First, SFA and the log-ratio operator are used to generate difference images (DIs) independently. Then, the NSCT is used to fuse two DIs into a new higher quality DI. The newly fused DI combines the complementary information of the two kinds of original DI. Therefore, the contrast of the changed regions and unchanged regions is greatly enhanced, as well as the changed details are preserved more completely. Furthermore, an NLM filtering algorithm is employed to suppress the strong speckles in the fused DI. We use the fuzzy C-means algorithm to generate the final binary change map. The experiments are carried out on two public datasets and three real-world SAR datasets from different scenarios. The results demonstrate that the proposed method has higher detection accuracy by comparing with the reference methods.

Index Terms—Nonsubsampled contourlet transform (NSCT), slow feature analysis (SFA), synthetic aperture radar (SAR) images, unsupervised change detection.

I. Introduction

REMOTE sensing image change detection is a process of using multisource remote sensing images that cover the same geographical area at different times to determine the changes of objects or regions. It has been widely used in environmental monitoring [1]–[3], forest monitoring [4], [5], agricultural survey [6]–[8], disaster assessment [9]–[11], military monitoring, and so on. Compared with optical sensors and infrared sensors, synthetic aperture radar (SAR) has attracted considerable attention due to its unique advantages that it can work under nearly all weather conditions and any time [12]. SAR images have become an extremely important data source for change detection. It is of great realistic significance to research change detection in SAR images.

There are various change detection technologies used in SAR images. In general, it can be divided into two broad categories. One is utilizing the distinctive attributes of SAR images, such as coherence and polarization property. Coherent change detection, using phase differences, can detect subtle trace between repeat-pass images. In [13]–[15], InSAR coherence has been exploited for surface deformation and land subsidence. In [16]–[18], coherent information is made full use to detect the small and slight changes, such as the small changes in urban areas. The polarimetry information is also used for monitoring the changes of land cover. In [19]–[21], the PolSAR images are implemented in change detection to monitor crops and flood. The other is based on the digital image analysis and processing methods, such as image transformation and image fusion. Some deep learning methods have also been proposed recently, which show certain competitive power in detection accuracy. Essentially, they belong to the second kind.

Among the change detection methods, the supervised methods may have better performance by using a large amount of labeled data [22]. However, sufficient labeled data are always difficult to obtain; thus, they could not meet the needs of practical applications. On the contrary, unsupervised methods do not need the support of any prior information. Meanwhile, they can minimize the man-made errors. Therefore, in this article, we consider using an unsupervised method for SAR image change detection.

The general unsupervised methods for change detection in multitemporal SAR images usually include three steps: 1) image preprocessing; 2) generating difference images (DIs); and 3) analyzing DI [23]. The purpose of image preprocessing is to ensure that the multitemporal images have the consistent comparability. The processing of geometric correction, radiometric correction, and registration is necessary when we get the raw data at first time. Generating DI is the core of change detection. Among the algorithms for generating DI, both the difference operator [24]–[26] and the ratio operator [27] are classic algorithms. The difference operator is a simple and fast method, but it is very sensitive to noise, which could lead to serious false alarms. The log-ratio (LR) operator [28], [29] can suppress the speckle

noise by transforming the multiplicative noise into additive noise. However, when processing with logarithm, the contrast of changed and unchanged pixels is simultaneously decreased, which may reduce the accuracy of the final classification. The mean-ratio operator [30] and the neighborhood-based ratio (NR) operator [31], both using the neighborhood information, can extract the more accurate information of changed regions. However, it is easy to bring the fuzziness of boundary information. Some image transformation methods, such as principal component analysis (PCA) [32] and wavelet transform [33], [34], could not show excellent performance in SAR images due to the especially imaging mechanism. Some auxiliary denoising methods are also used in change detection, such as the low-rank model [35]. Although it may improve the signal-to-noise ratio to a certain extent, it cannot improve the separability between changed and unchanged pixels. It may work bad on those regions where the changes are not obvious.

Slow feature analysis (SFA), as an unsupervised feature learning algorithm [36], showed a superior performance in multispectral images [37]–[39] change detection. According to the principle of SFA, the differences of the unchanged pixels are suppressed so that the changed pixels can be better separated. In this article, we introduce it into SAR images and hope to obtain the DI, which has well separability as well as high contrast.

Recently, the strategy of image fusion has been used in change detection [40]. It showed a better result compared with those traditional methods. The main reason is that the information in changed regions can be enhanced, and the background information is suppressed effectively by using the appropriate fusion rules. The nonsubsampled contourlet transform (NSCT), as a multiscale and multidirectional decomposition method, has been widely applied in digital image processing [41]. It inherits the key merits of those traditional image fusion methods and has the special property of shift invariance. It is proved that the NSCT can overcome the problem of pseudo-Gibbs phenomena by upsampling the directional filter rather than downsampling the input signal in the process of decomposition. Based on this, this article uses the NSCT to fuse two independent DIs into a new DI, which the details can be better preserved, and the edge information will not be aliased.

The analysis of DI is the last step in change detection, which can be seen as a binary segmentation problem. The K-means clustering and the fuzzy C-means (FCM) clustering are common algorithms. The latter usually has a better flexibility and reliability in SAR images [42]. Therefore, in this article, we choose FCM to cluster the fused DI.

The purpose of our study is to build an unsupervised change detection framework in multitemporal SAR images with stronger stability and better generalization. The changed information in SAR images can be better extracted as the introduction of SFA. However, it could not always perform well because of the existence of the speckles. SFA is sensitive to noise, which may lead to the serious miss detection. In order to overcome this deficiency, a novel contourlet fusion method is, then, utilized. The new method is based on the NSCT, which combines the advantages of SFA and the LR operator. During the contourlet fusion, different fusion rules are implemented in different subbands. As for the low-frequency subbands, the weighted average algorithm is adopted. The regional characteristic measurement (RCM) algorithm is used in the high-frequency subbands. Considering that some strong speckles may be still preserved in the fused DI, the nonlocal-means (NLM) [43], [44] algorithm is implemented to have a further filtering. Finally, the changed and unchanged regions are classified by FCM.

The main contributions of this article are summarized as follows.

1) For the critical problem of low separability between the changed and unchanged regions in unsupervised change detection methods, the SFA algorithm is first used in SAR images. The feasibility of the SFA algorithm for SAR images is demonstrated through the theoretical analysis. In the SFA algorithm, the extraction of the changed information is considered as an optimization problem. The unchanged components can be extracted accurately by computing the generalized eigenvalue. Then, the changed components can be separated effectively by suppressing the unchanged components.

2) To improve the performance, we design a novel framework based on the image fusion strategy. The NSCT image fusion algorithm is employed, which could protect more details of texture and structure in the newly DI. The nonlocal-based despeckling method, which represents arguably the current state of the art, is then implemented in the fused DI. To verify the stability and generalization, the extensive experiments have been conducted on the datasets containing different changed types. Our proposed framework showed the better performance even in those datasets with strong speckles.

The rest of this article is organized as follows. Section II introduces the proposed change detection framework in detail. Section III presents our experiments on two public datasets and three real-world SAR datasets. Finally, Section IV concludes this article.

II. METHODOLOGY

Given two coregistered SAR intensity images $X = \{X(i,j),\ 1 \leq i \leq I,\ 1 \leq j \leq J\}$ and $Y = \{Y(i,j), 1 \leq i \leq I,\ 1 \leq j \leq J\}$, which have the same size $I \times J$. They are all obtained from the same geographical region at different times. The purpose of the change detection is to separate out changed regions from the background. The proposed change detection framework is shown in Fig. 1. It involves the following three main parts: 1) generating two DIs by using the SFA algorithm and the LR operator, respectively, as marked by the red boxes; 2) fusing the generated DI through the NSCT, as marked by the purple box; and 3) the NLM algorithm is adopted to suppress the speckle noise, as marked by the green box. The following sections will give the full details of how each part is realized.

A. SFA for Generating DI

SFA, as an unsupervised feature learning algorithm, can extract invariant and slowly varying features from input high-dimensional time-series signals. It is applied in change detection

SAR image X   SAR image Y

Slow feature analysis   Log-ratio operator

Nonsubsampled contourlet transform

Nonlocal-means filtering

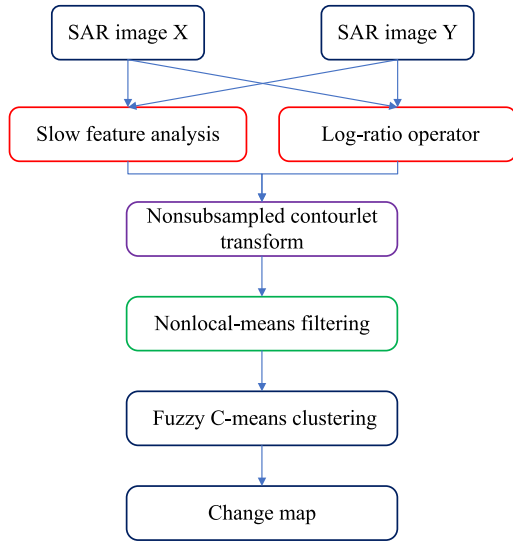Fuzzy C-means clustering

Change map

Fig. 1.   Proposed change detection framework.

in multispectral remote sensing images and shows superior performance. When using SFA in change detection, unchanged pixels are equivalent to slowly varying components and changed pixels correspond to fast varying components. The unchanged pixels can be suppressed, and the changed pixels can be separated effectively.

Fig. 2 shows the diagram of the SFA algorithm. The blue rectangles represent changed pixels, and the red triangles represent unchanged pixels. As for bitemporal SAR images, the difference value of the corresponding unchanged pixels is usually not equal to zero because of the difference in imaging and the interference of the speckles. The contrast between the changed and unchanged pixels is not obvious, and it is hard to find an appropriate classification method to separate them. SFA can transform the original difference space to the new slow feature difference space, as shown in the dashed frame. It minimizes the difference value of the unchanged pixels in background, so that the unchanged regions are gloomier, and the changed regions can be better detected.

Assume that the scattering vector pair of bitemporal SAR images is $x^i = [x_1^i, x_2^i, \ldots, x_N^i]$ and $y^i = [y_1^i, y_2^i, \ldots, y_N^i]$, where i indicates the number of pixels and N indicates the number of polarimetric channels that participate in computing. First, we need to preprocess the input image data. The normalized expressions are

$$\hat{x}_j^i = \frac{x_j^i - \mu_{x_j}}{\sigma_{x_j}} \tag{1}$$

$$\hat{y}_j^i = \frac{y_j^i - \mu_{y_j}}{\sigma_{y_j}} \tag{2}$$

where μ_{x_j} indicates the mean and σ_{x_j} indicates the variance of the image X in the jth polarimetric channel.

SFA is a feature space transformation algorithm, and we usually consider the linear case. Therefore, the transforming

function can be expressed as

$$g_j(\hat{x}^i) = w_j^T \hat{x}^i \tag{3}$$

where w is the mapping matrix.

Then, the optimized object function of SFA can be constructed as

$$\min \frac{1}{P} \sum_{i=1}^{P} (w_j^T \hat{x}^i - w_j^T \hat{y}^i)^2 \tag{4}$$

where P is the number of pixel pairs in the bitemporal images. The constraints of the optimized object function are composed of the following:

$$\frac{1}{2P} \left[\sum_{i=1}^{P} (w_j^T \hat{x}^i) + \sum_{i=1}^{P} (w_j^T \hat{y}^i) \right] = 0 \tag{5}$$

$$\frac{1}{2P} \left[\sum_{i=1}^{P} (w_j^T \hat{x}^i)^2 + \sum_{i=1}^{P} (w_j^T \hat{y}^i)^2 \right] = 1 \tag{6}$$

$$\frac{1}{2P} \left[\sum_{i=1}^{P} (w_j^T \hat{x}^i)(w_l^T \hat{x}^i) + \sum_{i=1}^{P} (w_j^T \hat{y}^i)(w_l^T \hat{y}^i) \right] = 0. \tag{7}$$

Equation (5) is the zero-mean constraint, which can simplify the solution process. The unit variance constraint (6) helps to avoid the constant solution. Constraint (7) guarantees that the different output components carry different information and avoids the redundancies of output information.

Construct two new matrices as

$$A = \frac{1}{P} \sum_{i=1}^{P} (\hat{x}^i - \hat{y}^i)(\hat{x}^i - \hat{y}^i)^T \tag{8}$$

$$B = \frac{1}{2P} \left[\sum_{i=1}^{P} (\hat{x}^i)(\hat{x}^i)^T + \sum_{i=1}^{P} (\hat{y}^i)(\hat{y}^i)^T \right]. \tag{9}$$

Then, (4), (6), and (7) can be rewritten as

$$\frac{1}{P} \sum_{i=1}^{P} (w_j^T \hat{x}^i - w_j^T \hat{y}^i)^2 = w_j^T A w_j \tag{10}$$

$$\frac{1}{2P} \left[\sum_{i=1}^{P} (w_j^T \hat{x}^i)^2 + \sum_{i=1}^{P} (w_j^T \hat{y}^i)^2 \right] = w_j^T B w_j \tag{11}$$

$$\frac{1}{2P} \left[\sum_{i=1}^{P} (w_j^T \hat{x}^i)(w_l^T \hat{x}^i) + \sum_{i=1}^{P} (w_j^T \hat{y}^i)(w_l^T \hat{y}^i) \right] = w_j^T B w_l. \tag{12}$$

With (8), (9), and (11), (10) can be written as

$$\frac{1}{P} \sum_{i=1}^{P} (w_j^T \hat{x}^i - w_j^T \hat{y}^i)^2 = \frac{w_j^T A w_j}{w_j^T B w_j}. \tag{13}$$

This optimization problem is equivalent to the generalized eigenvalue problem

$$AW = BW\Lambda \tag{14}$$

where W is the generalized eigenvector matrix and Λ is the diagonal matrix composed of the generalized eigenvalues.
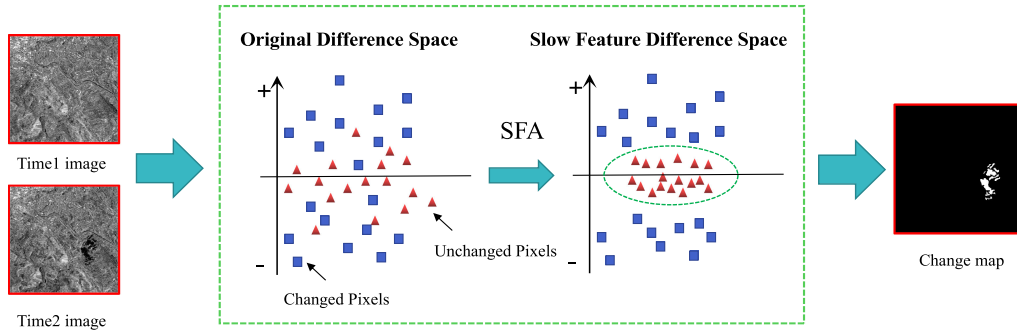
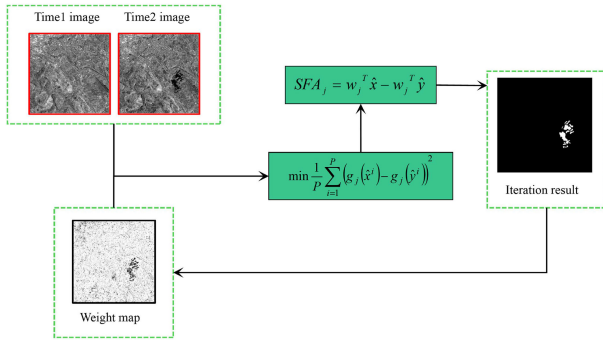Fig. 2. Schematic of SFA in SAR image change detection.

Fig. 3. Schematic of ISFA in SAR image change detection.

The generalized eigenvector matrix W is the mapping matrix that we hope to get, which can extract the slowly varying component from input data. Finally, the slow feature difference values are calculated as

$$\text{SFA}_j = w_j^T \hat{x} - w_j^T \hat{y}. \tag{15}$$

In order to have a further suppression to the difference values of the unchanged pixels, the adaptive iterative processing is usually carried out. During the iteration, the weights of the unchanged pixels become larger, and the weights of the changed pixels get close to zero. The unchanged pixels become more important to the calculation, and the changed pixels have almost no influence. The diagram of iteration slow feature analysis (ISFA) is shown in Fig. 3. As can be seen in the weight map, there are higher values in unchanged regions. It is obvious that unchanged regions obtain the higher weights. The derivation process of ISFA is as follows.

According to (5) and (6), the slow feature difference value satisfies an approximately Gaussian distribution. We assume that the polarimetric channels are independent of each other. Therefore, the square sum of the slow feature difference values from different polarimetric channels approximately follows a chi-squared distribution with N degrees of freedom. The chi-squared distance can be expressed as

$$T_i = \sum_{j=1}^{N} \frac{\left(\text{SFA}_j^i\right)^2}{\lambda_{\text{SFA}_j}} \quad \in \chi^2(N) \tag{16}$$

where λ indicates the eigenvalue computed from (14).

Apply the same weight (such as 1) to the pixels at the beginning of the iteration. Then, the weight is updated as

$$v^i = P\left\{\chi^2(N) > T^i\right\} \tag{17}$$

where v^i indicates the weight of the pixel i. This weight also represents the probability of no change. Then, the weighted mean value and the variance of the image X are updated as

$$\mu_{x_j} = \frac{\sum_{i=1}^{P} v^i x_j^i}{\sum_{i=1}^{P} v^i} \tag{18}$$

$$\sigma_{x_j}^2 = \frac{\sum_{i=1}^{P} v^i \left(x_j^i - \mu_{x_j}\right)^2}{\sum_{i=1}^{P} v^i}. \tag{19}$$

The mean value μ_y and the variance σ_y^2 of the image Y are updated at the same time. The input data should be normalized again using (1) and (2) with the weighted mean value and variance. We also add the weights in the matrices A and B

$$\hat{A} = \frac{\sum_{i=1}^{P} v^i \left(\hat{x}^i - \hat{y}^i\right)\left(\hat{x}^i - \hat{y}^i\right)^T}{\sum_{i=1}^{P} v^i} \tag{20}$$

$$\hat{B} = \frac{1}{2\sum_{i=1}^{P} v^i} \left[\sum_{i=1}^{P} v^i \left(\hat{x}^i\right)\left(\hat{x}^i\right) + \sum_{i=1}^{P} v^i \left(\hat{y}^i\right)\left(\hat{y}^i\right)\right]. \tag{21}$$

Then, the new mapping matrix W can be obtained. Finally, we use this new W to update (16) and (17). The iteration would not stop until the convergence condition is reached. We usually use $\Delta\lambda$ as the convergence indicator. In this article, the convergence threshold is set to be 10^{-6}.

From the above theoretical analysis, we can find that ISFA has a powerful ability of extracting change information. However, it should be noticed that the ability of resisting noise is weak. Considering that the LR operator can protect the spatial information effective but bad for extracting change information, we hope to find a method to combine the advantages of both and realize the advantage complementation. Obviously, image fusion is a good choice.

B. Image Fusion via NSCT

Images fusion is an effective way to obtain the DI with high quality. WT, as a multiscale image decomposition method, has obtained a wide range of applications. However, it can only
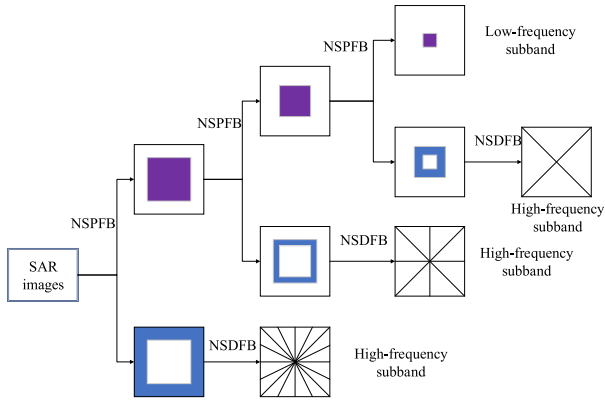
Fig. 4.   Diagram of multiscale and multidirectional decomposition of the NSCT.

Fig. 5.   Flowchart of the proposed image fusion method.

decompose in vertical, horizontal, and diagonal directions. The information from other directions is not sufficiently utilized. That means it could not satisfy the anisotropic characteristic of SAR images, which may not retain the image edge details effectively and could lead to edge fuzzy. Besides, downsamplers are usually adopted in image decomposition, which could generate pseudo-Gibbs phenomena and lead to the image aliasing [45].

Aim at the aforementioned problems, we introduce the NSCT into DI fusion. Fig. 4 shows the diagram of the NSCT, where the blue regions in the rectangles represent the high-frequency subbands and the purple regions represent the low-frequency subbands. It can be divided into two main parts: nonsub-sampled pyramid filter bank (NSPFB) and nonsubsampled directional filter bank (NSDFB). NSPFB is employed to the multiscale image decomposition. NSDFB is used to filter the high-frequency subbands in multidirection. When using NSPFB and NSDFB, signal downsampling is replaced by the filter upsampling, which overcomes the problem of pseudo-Gibbs phenomena.

Assume that the vector $L = [m_1, m_2, \ldots, m_l]$ includes all decomposition parameters, where l means the number of multiscale decomposition layers. m_l-level directional decomposition is implemented in the lth layer. Therefore, the total number of subbands is $1 + \sum_{i=1}^{l} 2^{m_i}$, i.e., one low-frequency subband and $\sum_{i=1}^{l} 2^{m_i}$ high-frequency subbands.

Fusion rules are the core of image fusion. After image decomposition, low-frequency subband coefficients contain a majority of information of images. They are an approximate expression of the original image and determine the general outline. High-frequency subband coefficients contain the details of texture and structure [46]. They reflect the edge mutation characteristics. Based on this, different rules are applied to different subband coefficients. Fig. 5 shows the flowchart of contourlet fusion.

For low-frequency subband coefficients, a weighted average method is adopted, which can be described as

$$D_F^L = \alpha D_{\text{ISFA}}^L + \beta D_{\text{LR}}^L \tag{22}$$

$$\alpha = \min\left(\left(\overline{D_{\text{ISFA}}^L}, \overline{D_{\text{LR}}^L}\right) / \left(\overline{D_{\text{ISFA}}^L} + \overline{D_{\text{LR}}^L}\right)\right) \tag{23}$$
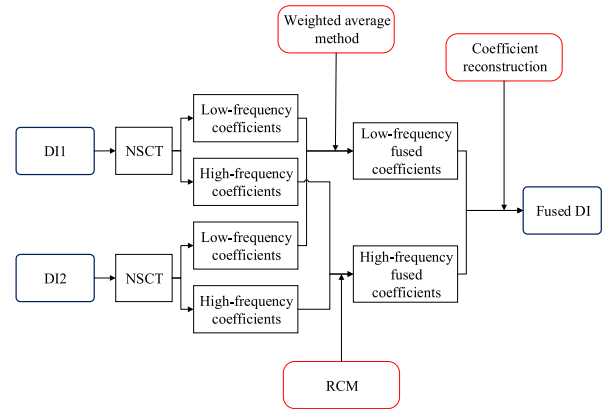
$$\beta = 1 - \alpha \tag{24}$$

where D_{ISFA}^L and D_{LR}^L represent the low-frequency subband coefficients from the DI generated by ISFA and LR, respectively. D_F^L indicates the updated low-frequency subband coefficients after fusion. $\overline{D_{\text{ISFA}}^L}$ and $\overline{D_{\text{LR}}^L}$ represent the mean values of the low-frequency subband coefficients, and α and β are weights. The LR has a better antinoise ability than ISFA. Therefore, the high weight is assigned to D_{LR}^L, and the background could be smoother and cleaner.

For high-frequency subband coefficients, it is more reasonable to determine the fused coefficients by utilizing neighborhood information. Here, an RCM fusion rule is given. First, define the local energy of the center pixel (i, j)

$$E_{l,k}(i,j) = \sum_{m=-1}^{1} \sum_{n=-1}^{1} w(m,n)[D_{l,k}(i+m, j+n)]^2 \tag{25}$$

$$w(m,n) = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{26}$$

where w is the weighted coefficient matrix of a 3×3 local region. $D_{l,k}(i,j)$ denotes the high-frequency subband coefficient at the point (i, j) in the lth layer and the kth direction.

Then, the matching degree of the local region $M(i,j)$ can be calculated as

$$M(i.j) = \frac{2\sum_{m=-1}^{1} \sum_{n=-1}^{1} w(m,n) \cdot D_{l,k}^{\text{ISFA,LR}}(i+m, j+n)}{E_{l,k}^{\text{ISFA}}(i,j) + E_{l,k}^{\text{LR}}(i,j)} \tag{27}$$

$$D_{l,k}^{\text{ISFA,LR}}(i,j) = D_{l,k}^{\text{ISFA}}(i,j) + D_{l,k}^{\text{LR}}(i,j). \tag{28}$$

Set a threshold parameter $T(T \in [0,1])$, and the matching degree is then compared with T. If the value of the matching degree is less than T, the specific rule can be expressed as

$$D_F^H(i,j) = \begin{cases} D_{\text{ISFA}}^H(i,j), & E_{l,k}^{\text{ISFA}} \geq E_{l,k}^{\text{LR}} \\ D_{\text{LR}}^H(i,j), & E_{l,k}^{\text{ISFA}} < E_{l,k}^{\text{LR}} \end{cases} \tag{29}$$

TABLE I
PROCESS OF GENERATING THE FUSED DI USING NSCT

Algorithm 1 NSCT for images fusion

Input: DI1 and DI2 generated by ISFA and LR;
Output: The fused DI3;
Main loop:
1: Initialize l, k, T and w.
2: Decompose the DI1 and DI2, obtain D_{ISFA} and D_{LR}.
3: **for** $i = 1 : l$ **do**
4:     Calculate the D_F^L according to (22)-(24);
5:     **for** $j = 1 : k$ **do**
6:         Calculate the M_R according to (27);
7:         **if** $M_R < T$ **then**
8:             Calculate the D_F^H according to (29);
9:         **else**
10:            Calculate the D_F^H according to (30)-(31);
11:        **end**
12:    **end**
13: **end**
14: **return** D_F
15: Obtain DI3 after coefficient reconstruction.

If the value of the matching degree is larger than T, the specific rule can be expressed as

$$D_F^H(i,j) = \begin{cases} \rho_{\max}D_{ISFA}^H(i,j) + \rho_{\min}D_{LR}^H(i,j), \\ \quad E_{l,k}^{ISFA} \geq E_{l,k}^{LR} \\ \rho_{\min}D_{ISFA}^H(i,j) + \rho_{\max}D_{LR}^H(i,j), \\ \quad E_{l,k}^{ISFA} < E_{l,k}^{LR} \end{cases} \quad (30)$$

$$\begin{cases} \rho_{\min} = \frac{1}{2} - \frac{1}{2}\left(\frac{1-M(i,j)}{1-T}\right) \\ \rho_{\max} = 1 - \rho_{\min} \end{cases} \quad (31)$$

where ρ represents the weights of the high-frequency subband coefficients. The specific calculation procedure of the image fusion algorithm is represented in Table I.

For low-frequency subband coefficients, we adopt a weighted average method to better reflect the general outline of the original DI. For high-frequency subband coefficients, RCM is implemented, so that more edges and textures can be preserved. If the matching degree is less than T, which means there are significant differences between the two DI in the same local region, then we select the coefficients with higher energy as the fused coefficients. If the matching degree is larger than T, which means there is a minimal difference in between, then the weighted average method is applied. Finally, the novel fused DI can be obtained after coefficient reconstruction.

C. Image Denoising via NLM

Although the NSCT has a certain suppression to the speckle noise, some strong speckles could be still preserved and seriously affect the final classification results. To solve this problem, the speckle suppression algorithm is then adopted.

SAR speckle reduction technologies have been developed rapidly during the past decades. A large number of models have been proposed, including spatial adaptive methods [47]–[49], homomorphic approaches [50], wavelet-based methods [51], [52], and nonlocal-based methods [53]–[55]. Among these methods, the NLM method represents arguably the current state of the art. The core idea is to compute the best value of the current
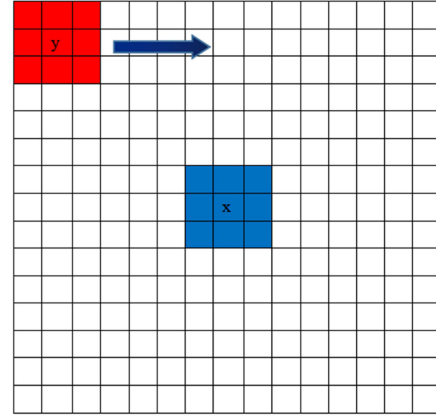

Fig. 6. Schematic of the NLM filtering algorithm.

pixel through a suitable measure of similarity based on the local patch. Both local and nonlocal information in fused DI are made full use during the algorithm execution. The strong speckles can be filtered better and details could be greatly protected.

Theoretically, all the pixels get involved in the process of relevant calculation. However, NLM is usually carried out in a small window in practical application. As shown in Fig. 6, the searching window centered on the pixel x is 15×15, the red sliding window centered on the pixel y is 3×3, and the blue neighborhood window of x is also 3×3. As the red window moves in the searching window, the similarity of the red and blue window is calculated.

The updated gray value of the center pixel x can be calculated by

$$\tilde{\mu}(x) = \sum_{y \in U} w(x,y)v(y) \quad (32)$$

$$\sum_{j \in U} w(x,y) = 1 \quad (33)$$

where $v(y)$ is the gray value of pixel y and U denotes the range of the searching window. The weight $w(x,y)$ represents the similarity of pixels x and y, which is defined as

$$w(x,y) = \frac{1}{Z(x)} \exp\left(-\frac{d(x,y)}{h^2}\right) \quad (34)$$

$$Z(x) = \sum_y \exp\left(-\frac{d(x,y)}{h^2}\right) \quad (35)$$

$$d(x,y) = G_\rho \|v(N_x) - v(N_y)\|^2 \quad (36)$$

where $d(x,y)$ represents the Euclidean distance between the two neighborhoods centered on x and y. G_ρ indicates a normalized Gaussian weighting function with zero mean and ρ standard deviation. h is an important smoothing parameter, which is used to control the degree of filtering. The larger the value of h, the smoother the images. Meanwhile, the more details may be lost. Therefore, a suitable compromise value of h is usually selected in practice.

TABLE II
BASIC INFORMATION OF SAR IMAGE DATASETS

Data sets	Bern	Sardinia	Sendai	Gulf of Mexico	Poyang Lake
Satellite	ERS-2	Landsat-5	TerraSAR-X	ALOS-1	ALOS-1
Acquisition Date	04/1999, 05/1999	09/1995, 07/1996	10/2010, 03/2011	03/2010, 06/2010	01/2009 to 01/2010
Polarization	HH	HH	HH	HH	HH
Ground range resolution	30m	30m	1.94m	4.68m	4.68m
Azimuth range resolution	26.3m	30m	3.30m	4.68m	4.68m
Image number	2	2	3	2	6
Image size	301×301	412×300	1001×1001	1038×731	445×468
Changed area	Lake	River	Coastal areas	Ocean	Lake

III. EXPERIMENT

To evaluate the performance of the proposed method, in this section, experiments are carried out on two public [56] and three self-made SAR image datasets. For public datasets, the ground-truth maps are offered by their authors. For self-made datasets, the reference change samples in ground truth are manually marked pixel by pixel. The basic information of these datasets is shown in Table II.

A. Experiment Settings

In the experiments, some change detection methods are introduced for comparison: NR operator [28], ISFA [34], wavelet fusion (WF) algorithm [37], neighborhood-based ratio and extreme learning machine (NR-ELM) [57], PCANet [58], and convolutional wavelet neural network (CWNN) [59]. The first three methods are unsupervised, and the last three methods are deep learning based. Some parameters in these methods are set to the recommended value. For NR, the neighborhood size is set to be 3×3. For ISFA, the convergence threshold is set to be 10^{-6}. For NR-ELM, the neighborhood size of the NR operator and the feature extraction are set to be 3×3 and 5×5, respectively. To control the variables, for three unsupervised comparison methods, the same NLM algorithms are implemented to suppress the speckle noise. Three deep-learning-based methods are implemented by the authors' open-source codes.

Some quantitative evaluation indexes are adopted, including false negatives (FN), false positives (FP), the overall error (OE), F1 score (F1), and the Kappa coefficient (Kappa). It should be noted that the Kappa coefficient and the F1 score are the two most convincing indexes in change detection.

B. Ablation Experiment

In order to verify the effectiveness of the proposed framework, the controlled experiment was first designed on the public Bern dataset. Two original images of this dataset were acquired in April and May 1999, respectively. The images size is 301×301. Fig. 7(a) and (b) shows the two original images, and Fig. 7(c) shows the ground-truth map.

Fig. 8 and Table III show the results of this test. As can be seen from Fig. 8(a), plenty of speckles are preserved, and some of the unchanged regions are wrongly classified in the final change map when using single ISFA. The result using single LR operator is also unsatisfying. The classification accuracy has been improved greatly when using the NSCT to obtain the
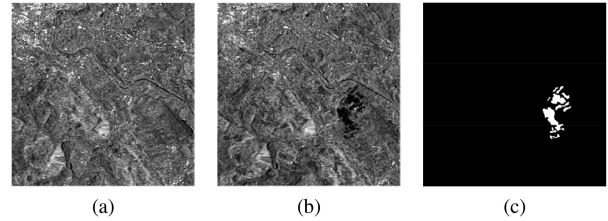

Fig. 7. Bern dataset. (a) Image acquired in April 1999. (b) Image acquired in May 1999. (c) Ground-truth map.
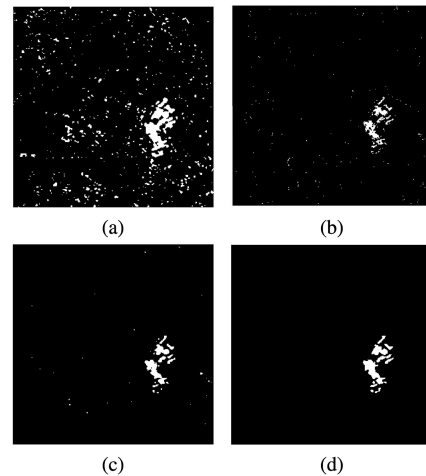

Fig. 8. Change detection results of the Bern dataset using (a) single ISFA, (b) single LR operator, (c) NSCT without NLM, and (d) proposed NSCT-NLM.

TABLE III
RESULTS OF THE CONTROLLED EXPERIMENT ON THE BERN DATASET

Settings	FN	FP	OE	F1	Kappa
Baseline:ISFA	85	2084	2169	0.4966	0.4871
LR	432	231	663	0.6856	0.6820
+NSCT	272	90	362	0.8299	0.8279
+NLM	195	86	281	0.8723	0.8708

better DI. The Kappa coefficient is nearly at 0.83, which is a very high value without despeckling. Fig. 8(d) shows the result when adding the despeckling algorithm into the framework. There are hardly speckles preserved in the final change map, and the value of Kappa coefficient rises almost 4 percentage points to 0.87. The improvement of the classification accuracy is significant as the implementation of NLM.
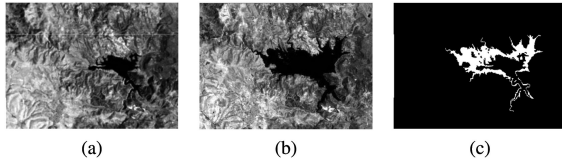
Fig. 9. Sardinia dataset. (a) Image acquired in September 1995. (b) Image acquired in July 1996. (c) Ground-truth map.
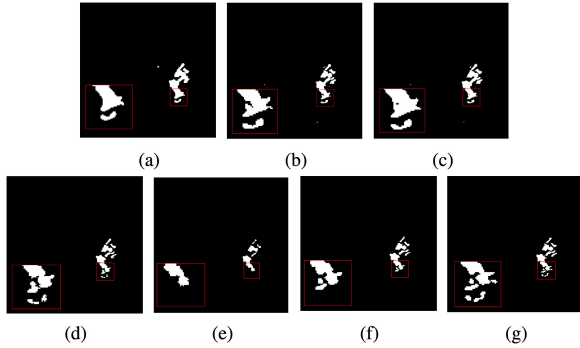

Fig. 10. Change detection results of the Bern dataset using (a) NR-NLM, (b) ISFA-NLM, (c) WF-NLM, (d) NR-ELM, (e) PCANet, (f) CWNN, and (g) proposed NSCT-NLM.

TABLE IV
RESULTS OF QUANTITATIVE EVALUATION ON THE BERN DATASET

Methods	FN	FP	OE	F1	Kappa
NR-NLM	**73**	442	515	0.8078	0.8049
ISFA-NLM	115	292	407	0.8361	0.8341
WF-NLM	**95**	342	437	0.8291	0.8267
NR-ELM	188	122	316	**0.8596**	**0.8578**
PCANet	483	**23**	506	0.7265	0.7238
CWNN	230	**85**	**315**	0.8545	0.8528
NSCT-NLM	195	86	**281**	**0.8723**	**0.8708**

C. Experiment on Two Public Datasets

In this section, we conducted the proposed framework on two public datasets: the first one is the Bern dataset, as mentioned in Section III-B, and the second one is the Sardinia dataset, acquired in September 1995 and July 1996, respectively. The image size is 412×300. Fig. 9(a) and (b) shows the original images, and Fig. 9(c) shows the ground-truth map of this dataset.

The final binary change maps of the Bern dataset are given in Fig. 10, and we zoom in on key areas for better visuals. The results of quantitative evaluation are given in Table IV, and the optimal two values of each evaluation index are shown in bold. As can be seen from Fig. 10, the clean and clear changed regions are extracted for three unsupervised comparison methods because of the implement of NLM. Among all the methods, the proposed NSCT-NLM shows the best performance. The reason is mainly due to the strong capability of detail preservation, such as the regions marked by the red rectangle. The results of CWNN and NR-ELM are also satisfactory. NR-NLM, ISFA-NLM, and WF-NLM are all working bad on edge protecting, which bring high values of OE. PCANet shows the worst performance as a large number of changed pixels are undetected.
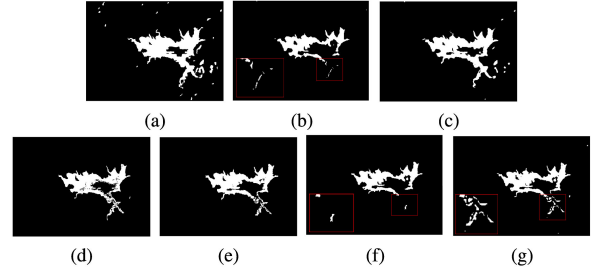

Fig. 11. Change detection results of the Sardinia dataset using (a) NR-NLM, (b) ISFA-NLM, (c) WF-NLM, (d) NR-ELM, (e) PCANet, (f) CWNN, and (g) proposed NSCT-NLM.

TABLE V
RESULTS OF QUANTITATIVE EVALUATION ON THE SARDINIA DATASET

Methods	FN	FP	OE	F1	Kappa
NR-NLM	**236**	5692	5928	0.7137	0.6895
ISFA-NLM	1410	**589**	**1999**	0.8615	0.8529
WF-NLM	**359**	3953	4312	0.7712	0.7531
NR-ELM	1004	2316	3320	0.7996	0.7853
PCANet	717	1960	2677	0.8377	0.8262
CWNN	1384	**609**	**1993**	**0.8623**	**0.8538**
NSCT-NLM	785	1325	2110	**0.8664**	**0.8573**

The final binary change maps of the Sardinia dataset are given in Fig. 11, and the results of quantitative evaluation are given in Table V. As shown in Fig. 11(a) and (c), some isolated noise points are still preserved in the binary change maps of NR-NLM and WF-NLM, which bring the low kappa values as well as F1 scores. The classification accuracy of ISFA-NLM, CWNN, and NSCT-NLM is very close, as the Kappa values are all over than 0.85 and F1 scores are all over than 0.86. As a whole, NSCT-NLM slightly outperforms the other two methods. We can find that NSCT-NLM has a lower FN as 785, which means fewer changed pixels are left out, such as the regions marked by the red rectangle in Fig. 10(b), (f), and (g).

According to the above tests, we can see that the proposed method shows satisfactory performance. In order to verify the generalization ability, the further experiments are carried out on three self-made SAR datasets, which contain different changed types.

D. Experiment on the Sendai Dataset

The first dataset is the Sendai dataset. On March 11, 2011, a strong earthquake and subsequent tsunami attacked the coastal areas of Sendai, Japan. The tsunami caused devastating damage and large swath of land was drowned [60]. This dataset has three images, acquired on October 20, 2010, March 13, 2011, and March 23, 2011, respectively. As the row images are too big to process, one part of the changed regions is selected with the size of 1001×1001. Fig. 12(a)–(c) shows the original images, and Fig. 12(d) and (e) shows the two ground-truth maps.

Fig. 13 shows the DI of the first two images generated by NR, ISFA, WF, and the proposed NSCT. For the NSCT, we set the value of the threshold parameter h to be 0.7. As can be seen from Fig. 13, the DI generated by NR has low values of the
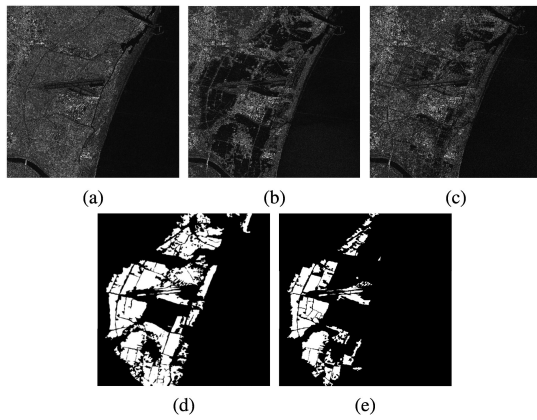
Fig. 12. Sendai dataset. (a) Image acquired in October 20, 2010. (b) Image acquired in March 13, 2011. (c) Image acquired in March 23, 2011. (d) Ground-truth map of the first two images. (e) Ground-truth map of the last two images.
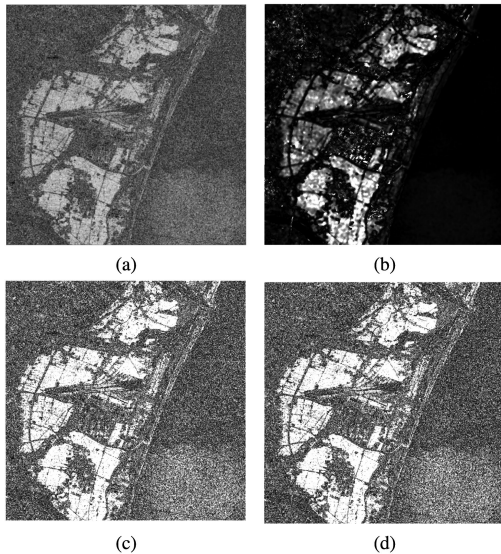

Fig. 13. Different DI generated by (a) NR, (b) ISFA, (c) WF, and (d) proposed NSCT.
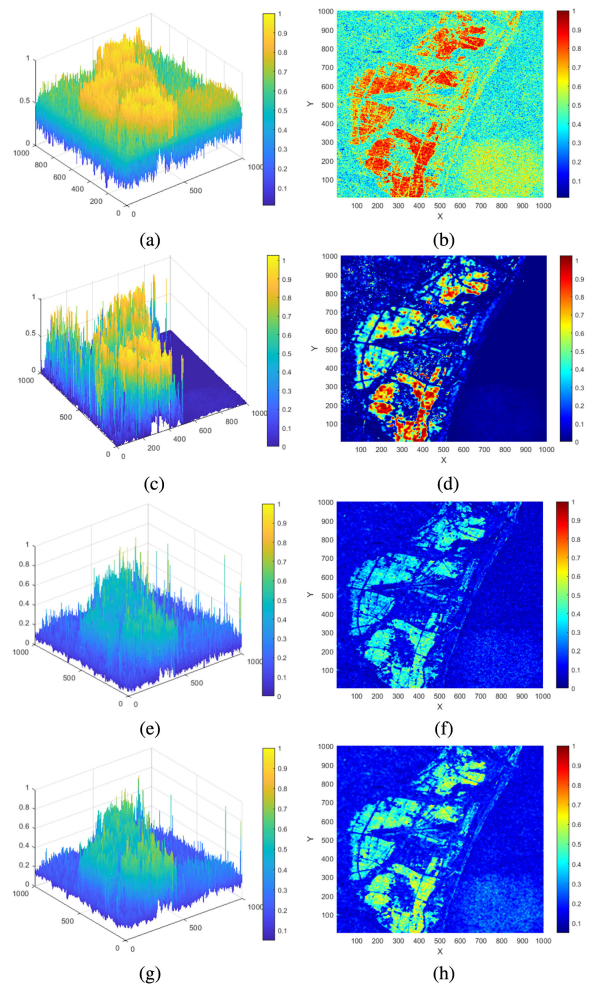

Fig. 14. Pseudocolor maps and 3-D maps of DI. (a), (c), (e), and (g) are the pseudocolor maps using the NR, ISFA, WF, and the proposed NSCT, respectively. (b), (d), (f), and (h) are the 3-D maps using the NR, ISFA, WF, and the proposed NSCT, respectively.

overall pixels, and the edges of the change regions are a bit fuzzy. Intuitively, there is a better separation between the changed and unchanged regions for ISFA. Visually, the difference between WF and NSCT is difficult to distinguish.

In order to have a better comparison, Fig. 14 shows the DI in the forms of the 3-D map and the pseudocolor map. The chroma values represent the normalized gray values of the pixels. As shown in Fig. 14(a) and (b), though the changed pixels have high chroma values, there is a low contrast between the changed and unchanged pixels. That means NR reduces the difference of changed regions and background when utilizing neighborhood information. On the contrary, ISFA maximizes the contrast, but the chromaticity range of changed regions is getting wider, which is approximately between 0.5 and 0.9. Meanwhile, there are some strong speckles preserved in the background. That is easy to generate the miss detection and false detection. The results of WF and NSCT are somewhat similar, but the separability of NSCT is higher, and the information in the background is better suppressed. The chromaticity range of

changed regions in the NSCT is approximately at 0.5–0.7. That means the changed regions could be better extracted in the final classification.

The selection of the parameter values will have great influence to the results when using the NLM algorithm. Theoretically, the larger the size of the searching window is, the better the result of filtering is. However, it will be very time consuming if the size is set too large. To obtain the best size of the searching window and sliding window, a large number of repetitive experiments were carried out on the Sendai dataset. Fig. 15(a) and (b) shows the curves of time cost and Kappa coefficient with different sizes of the searching window and sliding window, respectively. As can be seen from the changes of the curves, the time cost will increase dramatically with the increasing of the searching window size. However, the Kappa coefficient is stagnant and even decreases slightly. In other words, there are no good returns received by the large size of the searching window. It can easily lead to over-smoothing when using the large window in mid-low resolution data. In addition, the size of the sliding window is often set to be 3×3 or 5×5 on experience. From the curves, there are no significant difference of the Kappa coefficient between them,
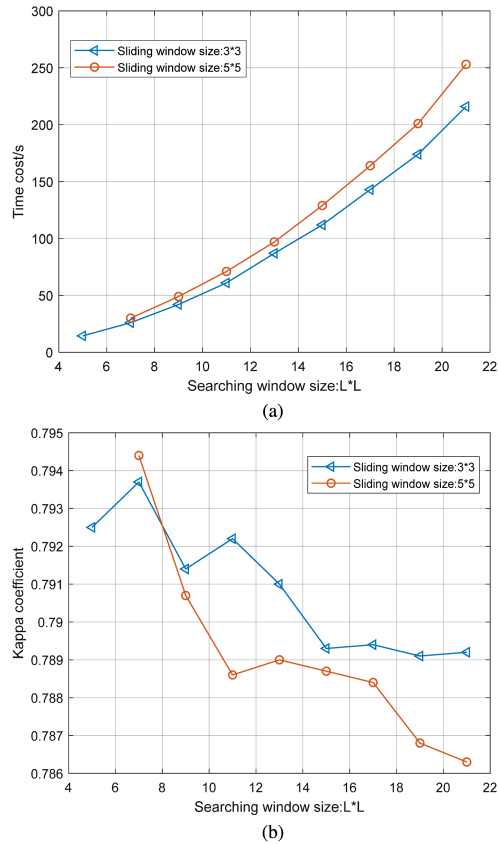
(a)

(b)

Fig. 15. Line chart of time cost and Kappa coefficient with different sizes of the searching window and sliding window. (a) Relationship between the time cost and window sizes. (b) Relationship between the Kappa coefficient and window sizes.
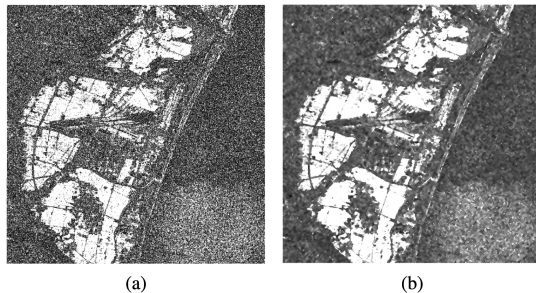

Fig. 16. Denoising result using NLM. (a) Fused DI generated by the NSCT. (b) Novel DI after denoising using NLM with h is 0.4.

Fig. 17. Change detection results of the first two images of the Sendai dataset using (a) NR-NLM, (b) ISFA-NLM, (c) WF-NLM, (d) NR-ELM, (e) PCANet, (f) CWNN, and (g) proposed NSCT-NLM.
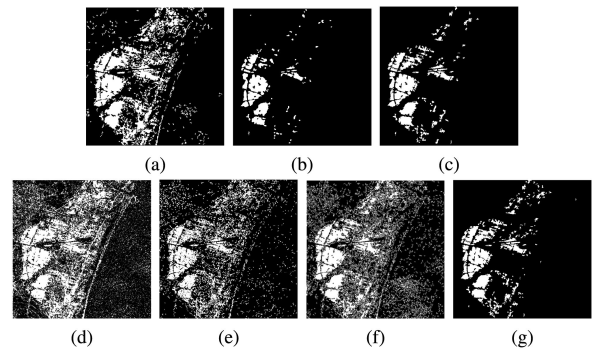
Fig. 18. Change detection results of the last two images of the Sendai dataset using (a) NR-NLM, (b) ISFA-NLM, (c) WF-NLM, (d) NR-ELM, (e) PCANet, (f) CWNN, and (g) proposed NSCT-NLM.

TABLE VI
RESULTS OF QUANTITATIVE EVALUATION ON THE SENDAI DATASET I

Methods	FN	FP	OE	F1	Kappa
NR-NLM	42814	54271	97085	0.7977	0.7340
ISFA-NLM	85143	10749	95892	0.7566	0.6997
WF-NLM	65257	**8077**	73334	0.8216	0.7767
NR-ELM	78446	38679	117125	0.7695	0.6918
PCANet	**16480**	51548	**68028**	**0.8430**	**0.7995**
CWNN	**30969**	129222	160191	0.7173	0.6105
NSCT-NLM	59522	**9400**	68922	**0.8352**	**0.7925**

while the size of 5×5 is more time-consuming. To improve the algorithm efficiency, in this article, the size of the searching window is set to be 5×5, and the size of the sliding window is set to be 3×3. Fig. 16 shows the despeckling results of the Sendai datasets using the above parameter values. After filtering, the speckle noise is suppressed significantly.

The final results are shown in Figs. 17 and 18. Fig. 17 shows the changed regions submerged by tsunami, and Fig. 18 shows the changed regions where the flood receded. The results of quantitative evaluation are given in Tables VI and VII, respectively. As can be seen from Fig. 17, there are hardly isolated noise points for those unsupervised methods. A lot of speckles
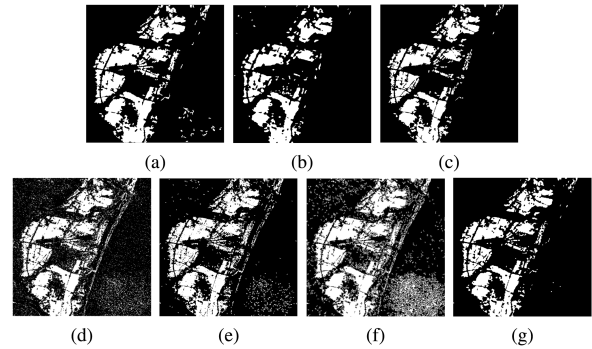
are preserved in the final change maps generated by NR-ELM and CWNN, while PCANet works better on suppressing the speckles. As shown in Table VI, PCANet and the proposed NSCT-NLM have the optimal two values of F1 score and Kappa coefficient. PCANet has the lowest FN as 16 480, which means it has a higher recall. NSCT-NLM has the lower FP as 9400, which means it has a higher precision ratio. Both of them are satisfactory. WF-NLM and NR-NLM take second place, while ISFA-NLM, NR-ELM, and CWNN show the worse performance. For ISFA, a larger number of changed pixels is left out, which leads to a bad result.

In Fig. 18, unsupervised methods show better than those deep-learning-based methods. As indicated in Table VII, WF and NSCT-NLM, as the image fusion methods, obtain the higher F1 scores and Kappa values. NSCT-NLM gets the lowest OE,

TABLE VII
RESULTS OF QUANTITATIVE EVALUATION ON SENDAI DATASET II

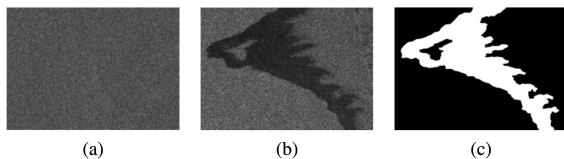| Methods	FN	FP	OE	F1	Kappa
NR-NLM	**17447**	104142	121589	0.6171	0.5531
ISFA-NLM	62412	**5309**	67721	0.6103	0.5776
WF-NLM	44060	**12650**	56710	**0.7230**	**0.6935**
NR-ELM	39048	79221	118269	0.5637	0.4972
PCANet	121284	30074	151358	0.5301	0.4486
CWNN	**19484**	203383	222867	0.4627	0.3555
NSCT-NLM	37563	14471	**52034**	**0.7496**	**0.7210**

Fig. 19. Gulf of Mexico dataset. (a) Image acquired in March 2010. (b) Image acquired in June 2010. (c) Ground-truth map.

TABLE VIII
RESULTS OF QUANTITATIVE EVALUATION ON THE GULF OF MEXICO DATASET

Methods	FN	FP	OE	F1	Kappa
NR-NLM	32599	7150	39749	0.8937	0.8591
ISFA-NLM	**5878**	23919	29797	0.9286	0.9016
WF-NLM	19083	**3997**	**23080**	**0.9399**	**0.9196**
NR-ELM	97973	33301	131274	0.6077	0.5020
PCANet	21171	54883	76054	0.7920	0.7267
CWNN	73459	5806	79265	0.7610	0.6977
NSCT-NLM	**13499**	5930	**19429**	**0.9504**	**0.9332**

and it shows a better balance in the recall and precision ratio. The results of three deep-learning-based methods are chaotic, and the changed information is submerged in noise. All of them have low F1 score and Kappa coefficient. The reason is mainly due to the wrong choice of training samples. In these methods, the positive and negative samples are selected by using the clustering methods. The confidence degree of samples will be lower when the clustering precision decreases due to the interference of the speckles. The wrong training samples will lead to the wrong classification of the DI.

E. Experiment on the Gulf of Mexico Dataset

The second dataset is the Gulf of Mexico dataset. On April 20, 2010, a serious oil spill happened in the north of Gulf of Mexico. After several months, the spilled oil has rapidly spread to the vast part of the Gulf of Mexico. Two images of this dataset were captured in March and June 2010, respectively. The size of them is 1038×731. Fig. 19(a) and (b) shows the original images and Fig. 19(c) shows the ground-truth map.

The final binary change maps are given in Fig. 20. The results of quantitative evaluation are shown in Table VIII. The classification accuracy of the Gulf of Mexico dataset, in general, is higher than that of the Sendai dataset. It is mainly because there is hardly strong electromagnetic scattering in the surface of sea. Therefore, the background of the Gulf of Mexico dataset is smooth enough. On the contrary, the background of the Sendai
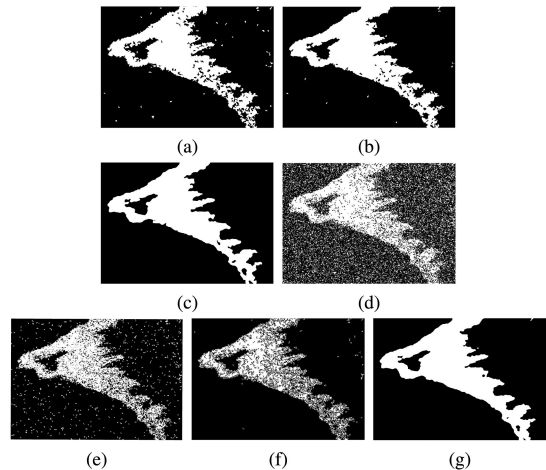

Fig. 20. Change detection results of the Gulf of Mexico dataset using (a) NR-NLM, (b) ISFA-NLM, (c) WF-NLM, (d) NR-ELM, (e) PCANet, (f) CWNN, and (g) proposed NSCT-NLM.

dataset mainly consists of building areas, which are easily to generate a large number of strong scattering points. For the Gulf of Mexico dataset, the unsupervised methods perform much better than those deep learning based. From the perspective of numerical, unsupervised methods are still keeping a higher classification accuracy, especially the proposed NSCT-NLM, which obtains the best values of F1 score and Kappa coefficient. Three deep-learning-based methods show the bad performance under the different evaluation indexes. It demonstrates that their generalization ability is insufficient.

F. Experiment on the Poyang Lake Dataset

The third dataset is from Poyang Lake, Jiang Xi province, China. Six images of this multitemporal dataset, which are captured from January 2009 to January 2010, present the dynamic change tendency of Poyang Lake in one year. In summer and autumn of 2009, the water surface area of Poyang Lake increased greatly because of the persistent heavy rainfall in Yangtze River Valley. In spring and winter, there is hardly heavy rainfall, and the water surface area was shrunk. As shown in Fig. 21, a typical changed region of Poyang Lake is selected in this experiment with the size of 445×468. Fig. 21(a)–(f) shows the six images and Fig. 21(g)–(k) shows five ground-truth maps of neighboring two images.

All the comparison methods are implemented in every neighboring two images. In order to save space, we exhibit only five sets of visible experiment results of NSCT-NLM, as shown in Fig. 22. The values of the Kappa coefficient for every method are listed in Table IX. In order to be expressed vividly, the line chart of the Kappa coefficient is presented as shown in Fig. 23. In Fig. 24, we present the values of FN, FP, and OE of different methods in forms of histograms. Generally, the classification results in the Poyang dataset is not as good as in the first two datasets. The reason mainly lies in the strong speckles. The background regions of this dataset are rougher, and the changes in some small areas are not obvious. Among all the methods, NSCT-NLM shows a best performance. It keeps the lowest value
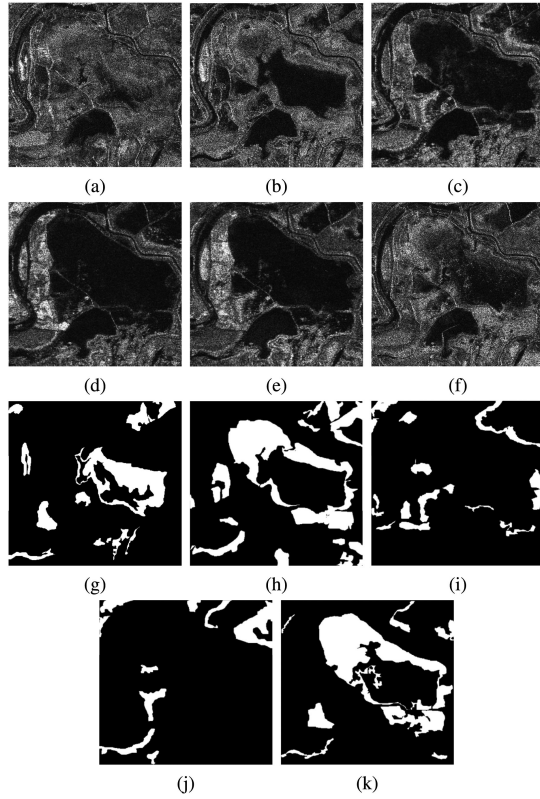
Fig. 21. Poyang Lake dataset. (a) Image acquired on January 15, 2009. (b) Image acquired on March 2, 2009. (c) Image acquired on July 18, 2009. (d) Image acquired on September 2, 2009. (e) Image acquired on October 18, 2009. (f) Image acquired on January 18, 2010. (g) Ground-truth map of the first two images. (h) Ground-truth map of (b) and (c). (i) Ground-truth map of (c) and (d). (j) Ground-truth map of (d) and (e). (k) Ground-truth map of the last two images.

TABLE IX
FIVE GROUPS OF KAPPA VALUES FOR DIFFERENT METHODS

Group number	1	2	3	4	5
NR-NLM	0.5485	0.6674	0.3574	**0.6287**	0.6645
ISFA-NLM	0.4691	0.5853	0.4525	0.3189	0.3914
WF-NLM	**0.6317**	**0.6942**	**0.5594**	0.6129	0.6811
NR-ELM	0.5655	0.6400	0.3666	0.3520	0.6526
PCANet	0.6152	0.6788	0.4979	0.4700	0.6947
CWNN	0.5249	0.4703	0.1960	0.2381	**0.7106**
NSCT-NLM	**0.6331**	**0.7115**	**0.6013**	**0.6572**	**0.7118**

of OE and the highest value of Kappa coefficient all the time. WF-NLM, NR-NLM, and PCANet also work well, while ISFA, CWNN, and NR-ELM are instable.

From the above experiments, we can see that the proposed NSCT-NLM has well generalization ability and strong robustness. In summary, NSCT-NLM is a superior method for multitemporal SAR image change detection.

G. Parameter Analysis

There are two main classes of parameters in our algorithm: the parameters used in generating the fused DI and the parameters used in filtering the fused DI. The former mainly includes the convergence threshold σ in ISFA and the matching threshold T in contourlet fusion. The latter mainly includes the searching
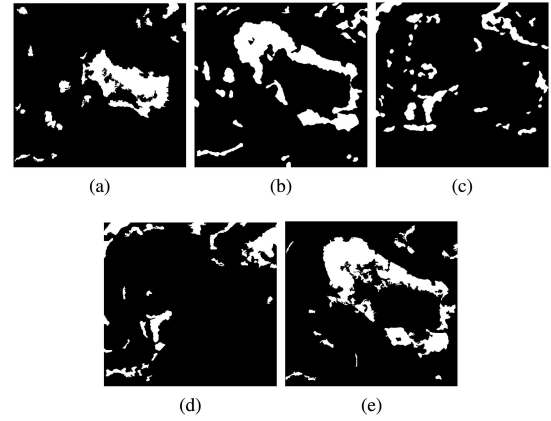

Fig. 22. Change detection results of the Poyang Lake dataset using NSCT-NLM. (a) Detection results of the first two images. (b) Detection results of the second and third images. (c) Detection results of the third and fourth images. (d) Detection results of the fourth and fifth images. (e) Detection results of the last two images.
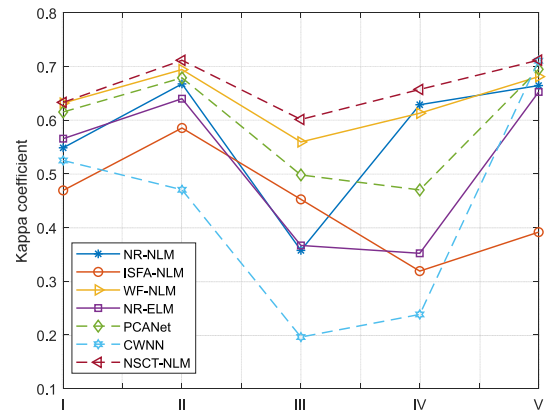

Fig. 23. Line chart of the Kappa coefficient for different methods.

window size L, the sliding window size l, and the smoothing parameter h. Based on experience, the changes of σ would not make an appreciable difference to the classification results, so σ is set to be a small value as 10^{-6}. The L and l usually have an important impact on filtering performance. To balance the filtering accuracy and computing time, we select two compromise values as 5×5 and 3×3, as described in Section III-D. The selection of T and h, theoretically, could affect the final classification results significantly. To evaluate their influence on classification accuracy, we analyze the relationships between Kappa coefficients and their values using the above three SAR datasets.

First, we vary the values of the matching threshold T from 0.5 to 0.95 with step 0.05, and other parameters remain the same. As shown in Fig. 25, the Kappa value first increases and then reduces as T increases. However, T has different impacts on different datasets. As a whole, these trends are not obvious in Sendai datasets as well as Poyang Lake datasets I and V. The reason is that the separability between changed and unchanged regions is very high in these datasets. The energy in the same local areas of the different DI is very close; thus, the changes of fused high-frequency subband coefficients are slight whether using the comparison strategy as (29) or the weighted average strategy as
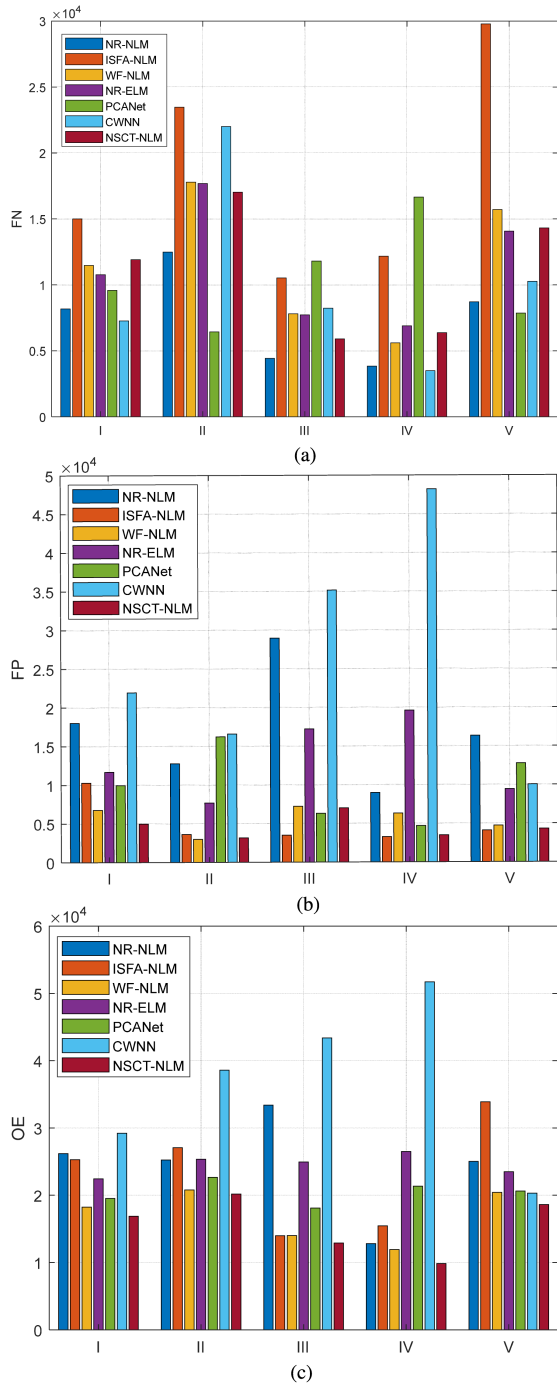
Fig. 24. Values of three quantitative evaluation indexes. (a) Values of FN for different methods. (b) Values of FP for different methods. (c) Values of OE for different methods.
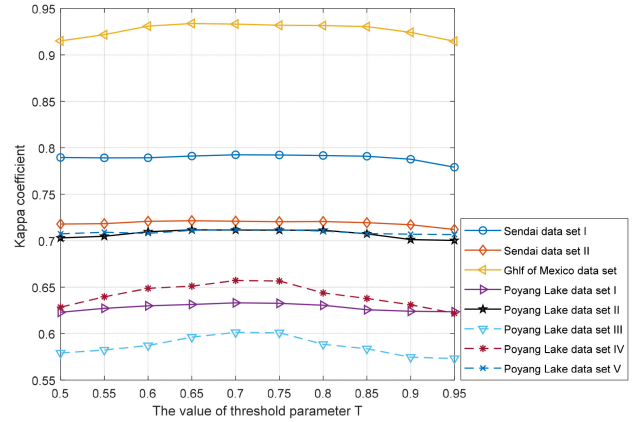

Fig. 25. Relationship between the threshold parameter T and the Kappa coefficient in different datasets.
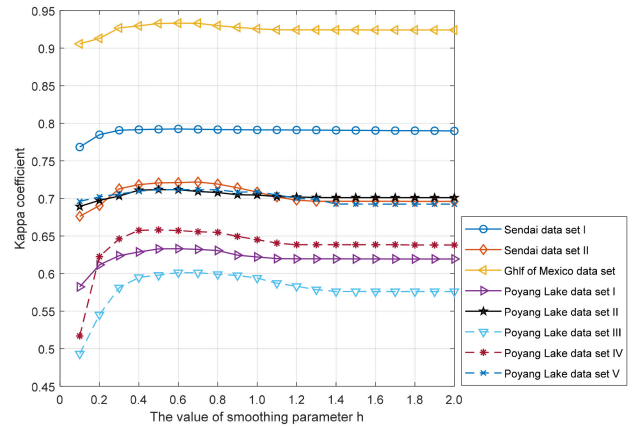

Fig. 26. Relationship between the smoothing parameter h and the Kappa coefficient in different datasets using NSCT-NLM.

Fig. 26, we present the results using NSCT-NLM with different values of h in different datasets. The h is set from 0 to 2.0, and other parameters remain the same. As the h increases, the Kappa values increase rapidly and then reduce and converge to a constant value gradually. Fig. 27 show the results in three different datasets using four unsupervised methods. There are similar change tendencies for different methods. When h is small, the changes will lead to an obvious improvement of Kappa values because of the deepening of filtering. However, the Kappa values hardly change as h becomes larger, and they no longer improve significantly as the fused DI is too smooth after a strong filtering. When h is between 0.4 and 0.8, the peak values of the Kappa coefficient are obtained.

Overall, the parameter h takes more important role in the final results. According to the above analysis, the small value of h is more suitable for the proposed method. Especially, when h is less than 1, the high values of the Kappa coefficient are obtained. Meanwhile, it should be noted that the value of h shall not set to be less than 0.3 or the performance of the proposed method will drop dramatically. In other words, the proposed method is strong robust when h is between in 0.3 and 1, no matter how it changes in this range.

(30). On the contrary, the Kappa value changes significantly in those datasets such as Poyang Lake datasets III and IV. We can find that the peak values of the Kappa coefficient in all of these datasets are obtained when T is between 0.65 and 0.75. To get the best performance of the proposed method, we set T to be 0.7.

The smoothing parameter h controls the decay of the exponential function and, therefore, decides the degree of filtering. In
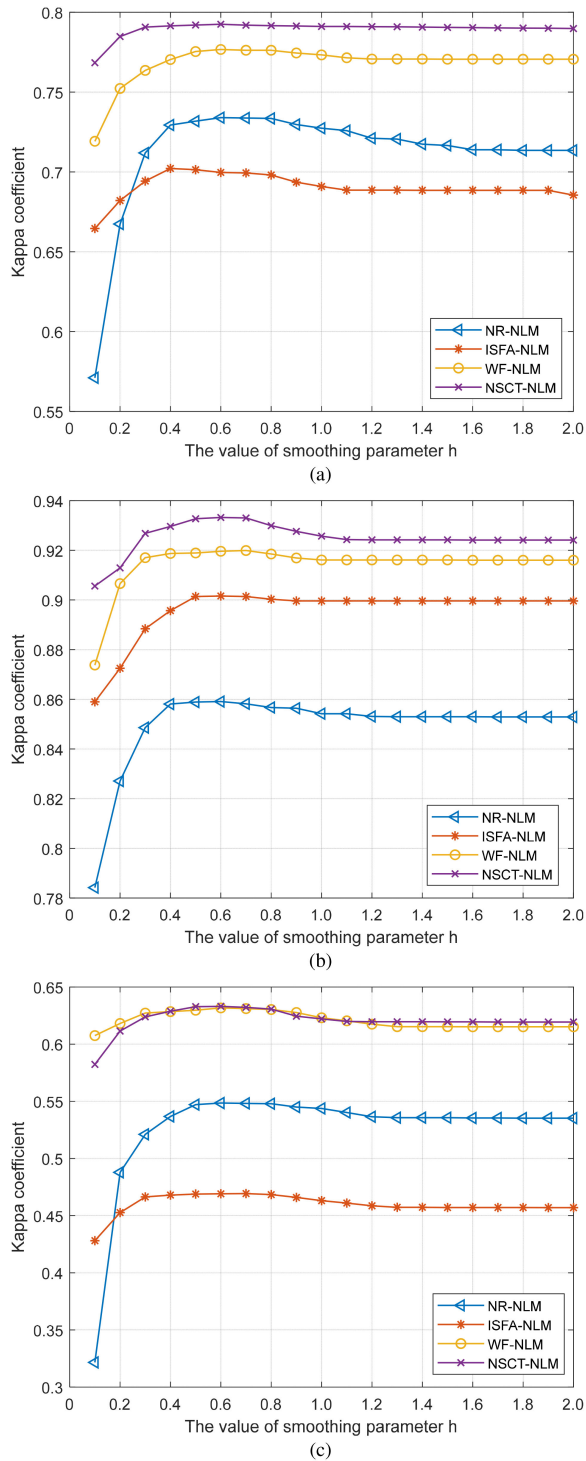
Fig. 27. Relationship between the threshold parameter h and the Kappa coefficient of unsupervised methods in (a) Sendai dataset I, (b) Gulf of Mexico dataset, and (c) Poyang Lake dataset I.

IV. CONCLUSION

In this article, we proposed a new multitemporal SAR image change detection method called NSCT-NLM. We first introduce the SFA algorithm into SAR image change detection. The key idea is to highlight the changed pixels by suppressing the difference between unchanged pixels. However, it is vulnerable to the influence of noise. To overcome the deficiencies of SFA, a novel

contourlet fusion method is then adopted. It should be noticed that fusion rules are not just based on individual pixels, the local region information is also involved. In this novel fusion method, the powerful feature extraction of SFA and the better antinoise of LR are both made full use. The integration of SFA and LR can make up for each other and improve the quality of the fused DI. Meanwhile, the introduction of NLM eliminates the interference from the strong speckles. The experiments are first conducted on two public SAR datasets. The results demonstrate that the proposed method is more effective than those comparison methods, including other unsupervised methods and even deep-learning-based methods. Then, we perform the additional experiments on three self-made datasets, which contain the different types of changes. The qualitative and quantitative results both show that our proposed method has well generalization ability and strong robustness. Especially, it has a greater potential on those SAR images with low imaging quality.

REFERENCES

[1] T. L. Evans, M. Costa, K. Telmer, and T. S. F. Silva, "Using ALOS/PALSAR and RADARSAT-2 to map land cover and seasonal inundation in the Brazilian pantanal," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 3, no. 4, pp. 560–575, Dec. 2010.

[2] I. Ali, C. Schuster, M. Zebisch, M. Förster, B. Kleinschmit, and C. Notarnicola, "First results of monitoring nature conservation sites in alpine region by using very high resolution (VHR) X-band SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 5, pp. 2265–2274, Oct. 2013.

[3] R. Fallourd *et al.*, "Monitoring temperate glacier displacement by multi-temporal terraSAR-X images and continuous GPS measurements," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 2, pp. 372–386, Jun. 2011.

[4] M. Hughes, S. Kaylor, and H. Daniel, "Patch-based forest change detection from landsat time series," *Forests*, vol. 8, no. 5, pp. 166–166, 2017.

[5] J. Reiche *et al.*, "Feature level fusion of multi-temporal ALOS PALSAR and landsat data for mapping and monitoring of tropical deforestation and forest degradation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 5, pp. 2159–2173, Oct. 2013.

[6] S. Yang, S. Shen, B. Li, T. Le Toan, and W. He, "Rice mapping and monitoring using ENVISAT ASAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 108–112, Jan. 2008.

[7] K. Miyaoka, M. Maki, J. Susaki, K. Homma, K. Noda, and K. Oki, "Rice-planted area mapping using small sets of multi-temporal SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1507–1511, Nov. 2013.

[8] A. Balenzano, F. Mattia, G. Satalino, and M. W. J. Davidson, "Dense temporal series of C- and L-band SAR data for soil moisture retrieval over agricultural crops," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 2, pp. 439–450, Jun. 2011.

[9] D. Brunner, G. Lemoine, and L. Bruzzone, "Earthquake damage assessment of buildings using VHR optical and SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2403–2420, May 2010.

[10] L. Shi, W. Sun, J. Yang, P. Li, and L. Lu, "Building collapse assessment by the use of postearthquake Chinese VHR airborne SAR," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 10, pp. 2021–2025, Oct. 2015.

[11] W. Sun, L. Shi, J. Yang, and P. Li, "Building collapse assessment in urban areas using texture information from postevent SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 8, pp. 3792–3808, Aug. 2016.

[12] F. Xu, H. Wang, and Q. Y. Jin, "Deep learning as applied in SAR target recognition and terrain classification," *J. Radars*, vol. 6, no. 2, pp. 136–148, 2017.

[13] F. Hu, F. J. van Leijen, L. Chang, J. Wu, and R. F. Hanssen, "Combined detection of surface changes and deformation anomalies using amplitude-augmented recursive inSAR time series," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5210816.

[14] B. Zhang, R. Wang, Y. Deng, P. Ma, H. Lin, and J. Wang, "Mapping the Yellow River delta land subsidence with multitemporal SAR interferometry by exploiting both persistent and distributed scatterers," *ISPRS J. Photogramm., Remote Sens.*, vol. 148, pp. 157–173, 2019.

[15] H. Yu, H. Lee, T. Yuan, and N. Cao, "A novel method for deformation estimation based on multibaseline inSAR phase unwrapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5231–5243, Sep. 2018.

[16] D. E. Wahl, D. A. Yocky, C. V. Jakowatz, and K. M. Simonson, "A new maximum-likelihood change estimator for two-pass SAR coherent change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2460–2469, Apr. 2016.

[17] A. V. Monti-Guarnieri, M. A. Brovelli, M. Manzoni, M. M. d'Alessandro, M. E. Molinari, and D. Oxoli, "Coherent change detection for multipass SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6811–6822, Nov. 2018.

[18] Y. Yang, X. Cong, K. Long, Y. Luo, W. Xie, and Q. Wan, "MRF model-based joint interrupted SAR imaging and coherent change detection via variational Bayesian inference," *Signal Process.*, vol. 151, pp. 144–154, 2018.

[19] C. Silva-Perez, A. Marino, J. M. Lopez-Sanchez, and I. Cameron, "Multi-temporal polarimetric SAR change detection for crop monitoring and crop type classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 12361–12374, 2021.

[20] A. Alonso-González, C. López-Martínez, K. P. Papathanassiou, and I. Hajnsek, "Polarimetric SAR time series change analysis over agricultural areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 10, pp. 7317–7330, Oct. 2020.

[21] S. Mahdavi, B. Salehi, W. Huang, M. Amani, and B. Brisco, "A PolSAR change detection index based on neighborhood information for flood mapping," *Remote Sens.*, vol. 11, no. 16, 2019, Art. no. 1854.

[22] Y. Zhang, S. Wang, C. Wang, J. Li, and H. Zhang, "SAR image change detection using saliency extraction and shearlet transform," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 12, pp. 4701–4710, Dec. 2018.

[23] L. Bruzzone and D. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 452–466, Apr. 2002.

[24] D. M. Muchoney and B. R. Haack, "Change detection for monitoring forest defoliation," *Photogramm. Eng. Remote Sens.*, vol. 60, no. 10, pp. 1243–1251, 1994.

[25] T. L. Sohl, "Change analysis in the United Arab Emirates: An investigation of techniques," *Photogramm. Eng. Remote Sens.*, vol. 65, no. 4, pp. 475–484, 1999.

[26] E. Rignot and J. van Zyl, "Change detection techniques for ERS-1 SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 4, pp. 896–906, Jul. 1993.

[27] A. Singh, "Review article digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, 1989.

[28] J. R. Dekker, "Speckle filtering in satellite SAR change detection imagery," *Int. J. Remote Sens.*, vol. 19, no. 6, pp. 1133–1146, 1998.

[29] 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.

[30] 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.

[31] M. Gong, Y. Cao, and Q. Wu, "A neighborhood-based ratio approach for change detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 307–311, Mar. 2012.

[32] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and k-means clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 772–776, Oct. 2009.

[33] S. Cui and M. Datcu, "Statistical wavelet subband modeling for multitemporal SAR change detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1095–1109, Aug. 2012.

[34] Z. Li, W. Shi, H. Zhang, and M. Hao, "Change detection based on Gabor wavelet features for very high resolution remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 783–787, May 2017.

[35] Y. Sun, L. Lei, D. Guan, X. Li, and G. Kuang, "SAR image change detection based on nonlocal low-rank model and two-level clustering," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 293–306, 2020.

[36] L. Wiskott and T. J. Sejnowski, "Slow feature analysis: Unsupervised learning of invariances," *Neural Comput.*, vol. 14, no. 4, pp. 715–770, 2002.

[37] C. Wu, B. Du, and L. Zhang, "Slow feature analysis for change detection in multispectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2858–2874, May 2014.

[38] C. Wu, L. Zhang, and B. Du, "Kernel slow feature analysis for scene change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 4, pp. 2367–2384, Apr. 2017.

[39] B. Du, L. Ru, C. Wu, and L. Zhang, "Unsupervised deep slow feature analysis for change detection in multi-temporal remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9976–9992, Dec. 2019.

[40] M. Gong, Z. Zhou, and J. Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2141–2151, Apr. 2012.

[41] A. Da Cunha, J. Zhou, and M. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, Oct. 2006.

[42] Y. Li, C. Peng, Y. Chen, L. Jiao, L. Zhou, and R. Shang, "A deep learning method for change detection in synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5751–5763, Aug. 2019.

[43] B. C. A. Buades and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.

[44] O. Yousif and Y. Ban, "Improving urban change detection from multitemporal SAR images using PCA-NLM," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2032–2041, Apr. 2013.

[45] B. Hou, Q. Wei, Y. Zheng, and S. Wang, "Unsupervised change detection in SAR image based on Gauss-log ratio image fusion and compressed projection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 8, pp. 3297–3317, Aug. 2014.

[46] Y. Yang, S. Tong, S. Huang, and P. Lin, "Multifocus image fusion based on NSCT and focused area detection," *IEEE Sens. J.*, vol. 15, no. 5, pp. 2824–2838, May 2015.

[47] J.-S. 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.

[48] V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. 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.

[49] D. T. Kuan, A. A. Sawchuk, T. C. 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.

[50] S. Solbo and T. Eltoft, "Homomorphic wavelet-based statistical despeckling of SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 4, pp. 711–721, Apr. 2004.

[51] F. Argenti and L. Alparone, "Speckle removal from SAR images in the undecimated wavelet domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2363–2374, Nov. 2002.

[52] T. Bianchi, F. Argenti, and L. Alparone, "Segmentation-based MAP despeckling of SAR images in the undecimated wavelet domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 9, pp. 2728–2742, Sep. 2008.

[53] C.-A. 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.

[54] C.-A. Deledalle, L. Denis, F. Tupin, A. Reigber, and M. Jäer, "NL-SAR: A unified nonlocal framework for resolution-preserving (Pol)(In)SAR denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2021–2038, Apr. 2015.

[55] C.-A. Deledalle, L. Denis, G. Poggi, F. Tupin, and L. Verdoliva, "Exploiting patch similarity for SAR image processing: The nonlocal paradigm," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 69–78, Jul. 2014.

[56] Z. Liu, Z. Chen, and L. Li, "An automatic high confidence sets selection strategy for SAR images change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4003505.

[57] F. Gao, C. Xie, Q. Zhi, J. Dong, and L. Bo, "Change detection from synthetic aperture radar images based on neighborhood-based ratio and extreme learning machine," *J. Appl. Remote Sens.*, vol. 10, no. 4, 2016, Art. no. 046019.

[58] F. Gao, J. Dong, B. Li, and Q. Xu, "Automatic change detection in synthetic aperture radar images based on PCANet," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1792–1796, Dec. 2016.

[59] F. Gao, X. Wang, Y. Gao, J. Dong, and S. Wang, "Sea ice change detection in SAR images based on convolutional-wavelet neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 8, pp. 1240–1244, Aug. 2019.

[60] S. Stramondo, "The Tohoku-Oki earthquake: A summary of scientific outcomes from remote sensing," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 895–897, Jul. 2013.

Weisong Li (Student Member, IEEE) received the B.S. degree in electronic information engineering from Shandong University, Weihai, China, in 2017, and the M.S. degree in electronic and communication engineering from the University of Chinese Academy of Sciences, Beijing, China, in 2020. He is currently working toward the Ph.D. degree with the Key Laboratory of Information Science of Electromagnetic Waves (Ministry of Education), Fudan University, Shanghai, China.

His research interests include synthetic aperture radar image processing, change detection, and deep learning applications.

Xiayang Xiao (Student Member, IEEE) received the B.Sc. degree in automation from Henan Polytechnic University, Henan, China, in 2014, and the M.S. degree in radar signal proceesing from Beijing Forestry University, Beijing, China, in 2017. He is currently working toward the Ph.D. degree in electronics and information with the Key Laboratory of Information Science of Electromagnetic Waves (Ministry of Education), Fudan University, Shanghai, China.

His research interests include synthetic aperture radar target detection and recognition and deep learning in remote sensing applications.

Penghao Xiao (Student Member, IEEE) received the M.S. degree in electronics and communication engineering from Shandong University of Science and Technology, Qingdao, China, in 2020. He is currently working toward the Ph.D. degree with the Key Laboratory of Information Science of Electromagnetic Waves (Ministry of Education), Fudan University, Shanghai, China.

His research interests include embedded artificial intelligence, model compression, and acceleration of deep learning.

Haipeng Wang (Senior Member, IEEE) received the B.S. and M.S. degrees in mechanical and electronic engineering from Harbin Institute of Technology, Harbin, China, in 2001 and 2003, respectively, and the Ph.D. degree in environmental systems engineering from the Kochi University of Technology, Kochi, Japan, in 2006.

He was a Visiting Researcher with the Graduate School of Information, Production and Systems, Waseda University, Fukuoka, Japan, in 2008. He is currently a Professor with the Key Laboratory of Electromagnetic Wave Information Science (Ministry of Education), Department of Communication Science and Engineering, School of Information Science and Engineering, Fudan University, Shanghai, China. His research interests include signal processing, synthetic aperture radar (SAR) image processing and analysis, speckle statistics, and applications to forestry and oceanography, machine learning, and its applications to SAR images.

Dr. Wang was the recipient of the Dean Prize of School of Information Science and Engineering, Fudan University in 2009 and 2017. Since 2011, he has been a Member of the Technical Program Committee of IEEE Geoscience and Remote Sensing Symposium. He is an Associate Editor for IEEE Geoscience and Remote Sensing Letters.

Feng Xu (Senior Member, IEEE) received the B.E. (Hons.) degree in information engineering from Southeast University, Nanjing, China, in 2003, and the Ph.D. (Hons.) degree in electronic engineering from Fudan University, Shanghai, China, in 2008.

From 2008 to 2010, he was a Postdoctoral Fellow with the National Oceanic and Atmospheric Administration Center for Satellite Application and Research, Camp Springs, MD, USA. From 2010 to 2013, he was a Research Scientist with Intelligent Automation Inc., Rockville, MD, USA. Since 2013, he has been a Professor with the School of Information Science and Technology, Fudan University, Shanghai, China. He is also the Vice-Dean of the School of Information Science and Technology and the Director of the Key Laboratory for Information Science of Electromagnetic Waves (Ministry of Education). He has authored or coauthored more than 70 articles in peer-reviewed journals and three books, among many conference papers and patents. His research interests include electromagnetic scattering theory, synthetic aperture radar information retrieval, and intelligent radar systems.

Dr. Xu was a recipient of the Early Career Award of the IEEE Geoscience and Remote Sensing Society (GRSS) in 2014 and the SUMMA Graduate Fellowship in the advanced electromagnetics area in 2007. Among other honors, he received the second-class National Nature Science Award of China in 2011. He is the Founding Chair of the IEEE GRSS Shanghai Chapter and an AdCom Member of the IEEE GRSS. He was an Associate Editor for IEEE Geoscience and Remote Sensing Letters.