

# Moving Target Shadow Detection Method Based on Improved ViBe in VideoSAR Images

Zhitao Wu , Hongtu Xie , *Member, IEEE*, Ting Gao , Yuanjie Zhang, and Haozong Liu 

**Abstract**—The video synthetic aperture radar (VideoSAR) has the high frame-rate imaging capability, which can perform the ground moving target detection. The moving target often leaves the shadow in the sequence of the VideoSAR images, which can indirectly detect the moving target by detecting its shadow. In view of the high false alarm rate and high missed detection rate of existing algorithms, this article proposes a moving target shadow detection method based on the improved ViBe in the VideoSAR images. First, the shadow information and local contrast information of the VideoSAR image are extracted by using the super-pixel segmentation technology, which can achieve the background reconstruction. Besides, the ViBe algorithm is improved by modifying the definition of the distance in the image pixel classification, adding the image pixel screening, and performing the morphological filtering, which is more suitable for the moving target shadow detection in VideoSAR images. Finally, by jointly extracting the regions of interest from multiple frames' detection results, most of the false alarms are suppressed, which can obtain the excellent performance in detecting the moving target shadow. The experimental results tested on the VideoSAR data released by the Sandia National Laboratories are shown to validate the correctness and effectiveness of the proposed algorithm, for which the detection rate of moving target shadows reached 97.66%, with only 8 false alarms.

**Index Terms**—Background reconstruction, improved ViBe, moving target shadow detection, video synthetic aperture radar (VideoSAR) images.

## I. INTRODUCTION

**S**YNTHETIC aperture radar (SAR) is a kind of radar that can work all day and all weather [1], [2], [3], [4], [5], [6]. Ground moving targets indication (GMTI) based on the

Manuscript received 1 July 2024; revised 23 July 2024; accepted 8 August 2024. Date of publication 13 August 2024; date of current version 26 August 2024. This work was supported in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2023A1515011588, in part by the Shenzhen Science and Technology Program under Grant 202206193000001 and Grant 20220815171723002, in part by the National Natural Science Foundation of China under Grant 62203465 and Grant 62201614, in part by the Fundamental Research Funds for the Central Universities, Sun Yat-sen University under Grant 24pnp202, in part by the China Postdoctoral Science Foundation under Grant 2024M753740, and in part by the Science and Technology Planning Project of Key Laboratory of Advanced IntelliSense Technology, Guangdong Science and Technology Department under Grant 2023B1212060024. (*Corresponding authors: Hongtu Xie; Ting Gao.*)

Zhitao Wu, Hongtu Xie, Yuanjie Zhang, and Haozong Liu are with the School of Electronics and Communication Engineering, Shenzhen Campus of Sun Yat-Sen University, Shenzhen 518107, China (e-mail: wuzht7@mail2.sysu.edu.cn; xiehongtu@mail.sysu.edu.cn; zhangyj357@mail2.sysu.edu.cn; liuhz6@mail2.sysu.edu.cn).

Ting Gao is with the Air Early Warning System Application Laboratory, Air Force Early Warning Academy, Wuhan 430019, China (e-mail: 80944993@qq.com).

Digital Object Identifier 10.1109/JSTARS.2024.3443080

SAR image has unique advantages, and the detection of moving targets is a research hotspot in the field of SAR-GMTI.

Traditional moving target detection is mainly performed in the signal domain. The commonly used methods can be mainly categorized into single-channel methods and multichannel methods. Multichannel moving target detection methods mainly include the space-time adaptive processing [7], displaced phase center antenna [8], and along-track interferometry [9], which have the advantage of detecting slow moving targets that are overwhelmed by the stationary clutter, but multichannel radar systems are more costly and have problems such as the channel mismatch. The single-channel moving target detection method is mainly the Doppler filtering method, and the clutter can be suppressed by the Doppler filtering process. However, the single-channel moving target detection method is difficult to detect moving targets submerged in the clutter.

In recent years, indirect GMTI techniques based on the ground moving target shadow detection have gradually attracted researchers' attention. Due to the shadow effect of moving targets on the electromagnetic wave, moving targets leave the shadows in SAR images [10]. These shadows can reflect the true position of moving targets, enabling the indirect detection of moving targets through the shadow detection [11]. Raynal et al. [10] from the Sandia National Laboratories have elaborated on the generation principles and characteristics of shadows in video synthetic aperture radar (VideoSAR) images. The energy of shadows in SAR images is relatively weak, and the direct detection would introduce a large number of the false alarms. In the VideoSAR image shadow detection, the most commonly used method is the background subtraction [12], [13], [14]. First, the background image is estimated using multiple frames, and then the difference between the background image and each frame is calculated to obtain shadow detection results. Tian et al. [15] used the track-before-detect method to address the moving target shadow detection in VideoSAR images, effectively improving the detection rate of moving target shadow and achieving the good false alarm suppression. Zhong et al. [16] combined the moving target shadow information with the echo energy information to detect moving targets, obtaining the good detection performance. Liu et al. [17] initialized the background model of the ViBe algorithm with the reconstructed background, effectively solving the "ghost" phenomenon when detecting moving target shadows with the ViBe algorithm. Using the deep learning for the moving target shadow detection is also a popular direction [18], [19], [20], [21], [22], [23]. Ding and Zhang et al. realized the moving target shadow detection in VideoSAR

images using deep convolutional neural networks [18], [19], [20], [21].

Currently, there are some shortcomings in moving target shadow detection methods in the VideoSAR images. First, when there are many moving target shadows in the scene, the estimated background image is not accurate. Using inaccurate background images for shadow detection results in a high incidence of missed alarms and false alarms, ultimately leading to suboptimal detection performance. In addition, detection methods based on individual pixels ignore the structural information of the moving target shadow. Although moving target shadow detection methods based on deep convolutional neural networks have the excellent performance, they require a large amount of VideoSAR sample data for training and suffer from the poor generalization. To address these issues, this article proposes a moving target shadow detection algorithm based on the improved ViBe algorithm in VideoSAR images. This method first utilizes the super-pixel segmentation technology to extract the shadow information and local contrast information of SAR images, achieving the background reconstruction. Then, the reconstructed background image is used to initialize the background model of the improved ViBe algorithm. Finally, moving target shadows in VideoSAR videos are detected, and false alarms are suppressed using the false alarm suppression method.

The rest of this article is organized as follows. Section II introduces the methods of the background reconstruction, including the shadow detection, removal of stationary target shadows, and shadow pixels filling. Section III presents the processing flow and main ideas of the proposed moving target shadow detection method based on the improved ViBe algorithm in VideoSAR images. Section IV provides the experimental results of the moving target shadow detection in VideoSAR images. Finally, Section V concludes this article.

## II. BACKGROUND RECONSTRUCTION

The ViBe algorithm possesses the ability to initialize its model using a single frame image, the outstanding performance in moving target detection, and a rapid background model updating strategy. However, if there are moving targets in the initial frame used for the background modeling in the ViBe algorithm, the “ghost” phenomenon may appear in the subsequent detection results, persisting for several frames or even dozens of frames. To mitigate the occurrence of the “ghost” phenomenon, it is necessary to perform the background reconstruction on the initial frame, removing the shadows of moving targets [17]. The flowchart of the background reconstruction is illustrated in Fig. 1, which primarily consists of three steps, i.e., the shadow detection, removal of stationary target shadows, and shadow pixels filling.

### A. Shadow Detection

Image binarization is the most convenient method to obtain the region of interest (ROI). In order to obtain the shadow information in SAR images, binarization algorithms are first used to extract the darker regions in the images. Among binarization algorithms, the maximum between-class variance

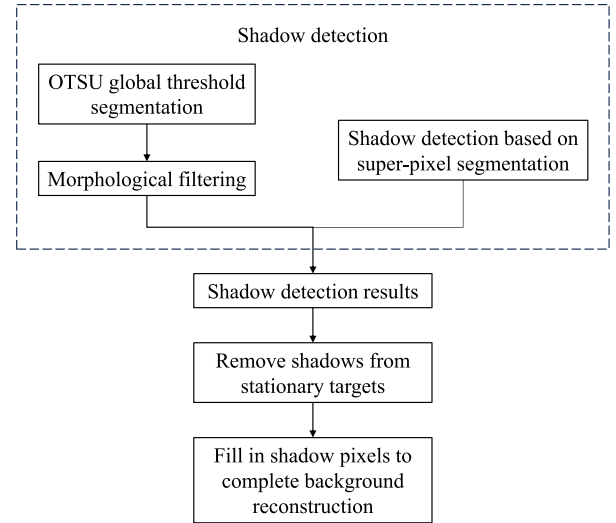


Fig. 1. Flowchart of the background reconstruction.

algorithm (OTSU algorithm) is widely used due to its high efficiency and good automation. Its main principle is to divide the pixels in the image into foreground and background using a grayscale threshold, and the selected threshold needs to ensure the maximum variance between the two classes of pixels. After obtaining the binary image, there will be a large number of false alarms in the image. Some of these false alarm regions are clearly not moving target shadows based on features such as the area, aspect ratio, and rectangularity. By using the morphological processing and connected component analysis, these regions can be effectively removed. Binarization and morphological filtering of SAR image are shown in Fig. 2.

After the binarization and morphological filtering, the ROIs with the darker appearance can be obtained. However, these darker regions may not necessarily be shadows. They could also be surfaces with low radar scattering coefficients, such as the road. Genuine shadow regions tend to be darker than their surroundings, whereas false shadow regions (such as the road with low radar scattering coefficients) exhibit the minimal difference in intensity compared to their surroundings. Exploiting this characteristic, a shadow detection method based on super-pixel segmentation can be employed to distinguish them.

Super-pixel segmentation is an image preprocessing technique that aggregates adjacent pixels with similar colors and textures into pixel blocks [24], [25], [26], [27]. Simple linear iterative clustering is a commonly used super-pixel segmentation algorithm, which not only produces excellent super-pixel segmentation results but also has the low computational complexity. If we represent a pixel using the three-dimensional coordinates  $[x, y, I]$ , composed of pixel positions and pixel intensities, then the distance  $D_{ij}$  between any two pixels can be expressed as

$$\begin{aligned}
 D_{ij} &= \sqrt{d_c^2 + \left(\frac{W^2}{S^2}\right) d_s^2} \\
 &= \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + \left(\frac{W^2}{S^2}\right) \cdot (I_i - I_j)^2} \quad (1)
 \end{aligned}$$

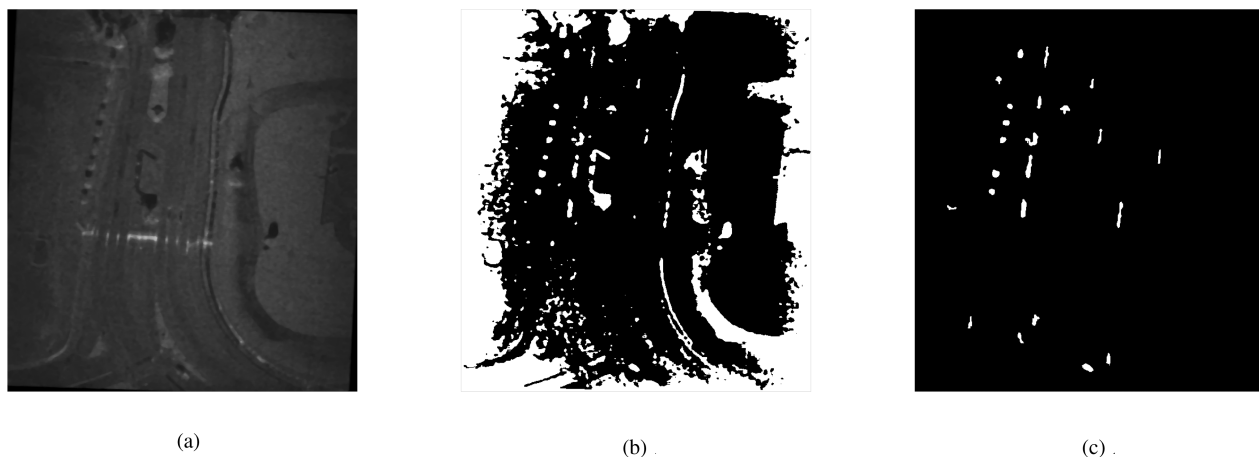


Fig. 2. Binarization and morphological filtering. (a) Original SAR image. (b) After binarization. (c) After morphological filtering.

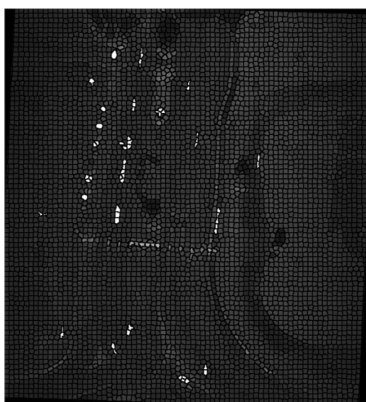


Fig. 3. Super-pixel segmentation.

where  $d_c$  represents the intensity space distance,  $d_s$  represents the spatial position distance,  $S$  represents the expected size of super-pixel, and  $W$  represents the maximum distance of the pixel intensity. When  $W$  is small, it indicates that the intensity space distance is more important, and the super-pixel blocks can better preserve the edge structure of the image. When  $W$  is large, it suggests that the spatial position distance is more important, and the generated super-pixels are more regular. Super-pixel segmentation is performed on the original image marked with ROIs, as shown in Fig. 3.

After performing the super-pixel segmentation, it is necessary to construct a shadow detection window for each super-pixel block, as illustrated in Fig. 4. The shadow detection window consists of a protection window and a reference window. The purpose of the protection window is to avoid including pixels from shadow areas in the reference window. The reference window is used to calculate the average intensity of background clutter, and then compare the average intensity of the reference window with the average intensity of the target super-pixel  $s$ . When the ratio of the average intensity of the target super-pixel  $s$  to the average intensity of the reference window is less than a preset threshold, the target super-pixel  $s$  is determined to be a shadow region. Conversely, if this ratio is greater than

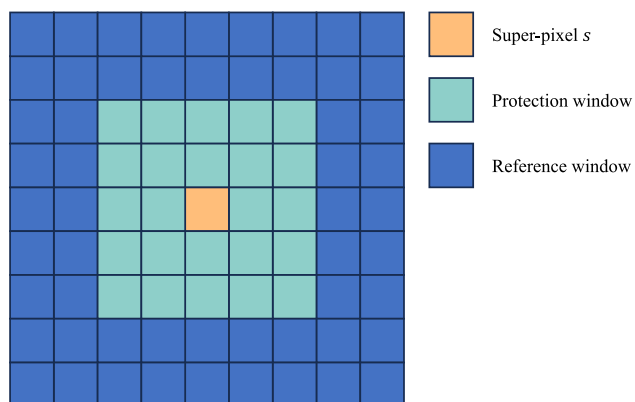


Fig. 4. Schematic of the shadow detection window.

the threshold, the target super-pixel  $s$  is determined to be a background region. The constructed detector can be expressed as

$$\begin{cases} H_0 : \frac{m_0}{m_1} > \chi \\ H_1 : \frac{m_0}{m_1} < \chi \end{cases} \quad (2)$$

where  $H_0$  assumes the target super-pixel  $s$  is a background region,  $H_1$  assumes the target super-pixel  $s$  is a shadow region,  $\chi$  represents the preset threshold.  $m_0$  and  $m_1$  represent the average intensity of the target super-pixel and the reference window, respectively. Fig. 4 illustrates the case where both the protection window and the reference window are two-level neighborhoods.

Using the shadow detection method based on the super-pixel segmentation, the shadow information for the entire image can be obtained. Next, this shadow information is used to filter the ROIs. If a region does not contain any shadow pixels, then it is reasonable to assume that this region is likely not a shadow region. Based on this premise, nonshadow regions are excluded from the ROIs. The specific filtering method involves determining whether the ROI contains shadow pixels. If it does, the current region is retained. Otherwise, it is removed. The process of filtering the ROIs using the shadow information is

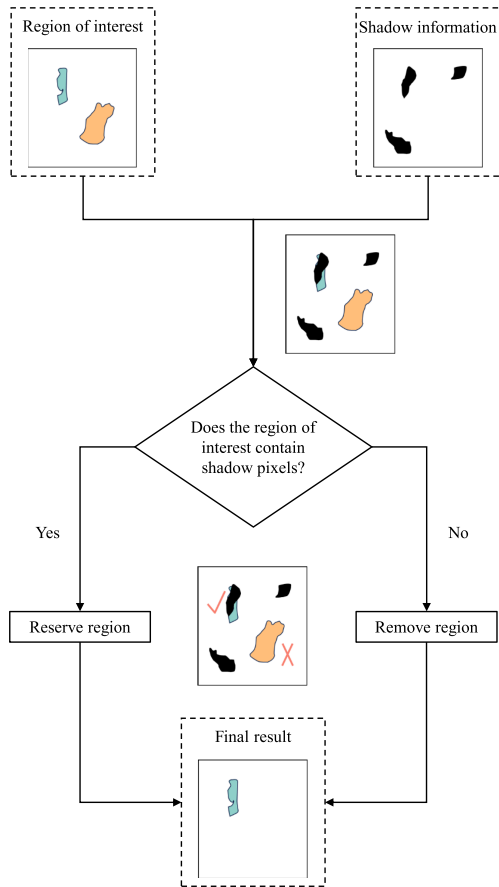


Fig. 5. Flowchart for filtering ROIs based on the shadow information.

illustrated in Fig. 5. After using the shadow information for filtering, nonshadow dark regions can be effectively removed.

### B. Removal of Stationary Target Shadows

In the ROIs obtained through binarization, morphological filtering, and shadow information-based filtering, shadows from both moving and stationary targets are included. When filling shadows for the background reconstruction, it is necessary to fill only the shadows of moving targets, not those of stationary targets. Therefore, it is essential to exclude the shadows of stationary targets and retain only the shadows of moving targets. Local contrast is an effective feature for distinguishing between stationary and moving target shadows [17]. As shown in Fig. 6, stationary targets do not deviate from their true positions in SAR images. Thus, the shadow neighborhoods of stationary targets contain imaging results of the stationary target itself, exhibiting the higher local contrast. On the other hand, moving targets deviate from their original positions due to their motion in SAR images, but their shadows remain near the true positions of the moving targets. Therefore, the shadow neighborhoods of moving targets do not contain imaging results of the moving target itself, resulting in the lower local contrast. By leveraging the difference in local contrast features between stationary and moving target shadows, it is possible to effectively distinguish the shadow regions between stationary and moving targets.

To calculate the local contrast, it is necessary to construct a local contrast detection window. As shown in Fig. 7, the traditional local contrast detection window uses the minimum bounding rectangle of the ROI as the central region of the detection window. Then, it constructs a complete detection window by building an eight-neighborhood of the same size as the central region. This type of the detection window ignores the edge structures of shadows and targets. The central region not only contains shadow pixels but also background clutter pixels. Similarly, for the neighboring regions containing targets, besides containing target pixels, they also include background clutter pixels. Moreover, situations may arise where a target spans two or even multiple neighborhoods. These issues can significantly impact the calculation of local contrast. To address this, this article proposes a redesigned local contrast detection window based on super-pixel segmentation. Super-pixel segmentation effectively preserves the edge structure information of shadows and targets in the image. It performs the super-pixel segmentation on the original image marked with ROIs. Then, it takes the super-pixel containing the ROI as the central region of the detection window, and directly adjacent regions to the central region as the neighborhood to complete the construction of the local contrast detection window. The local contrast detection window is illustrated in Fig. 8. By leveraging the ability of super-pixel segmentation to preserve the edge structure information in the image, the central region and neighboring regions including targets contain as few background clutter pixels as possible, resulting in the more accurate local contrast results. The formula for calculating local contrast  $C$  is given by

$$C = \frac{\max(G_{ni})}{G_c}, \quad i = 1, 2, \dots, k \quad (3)$$

where  $G_c$  represents the average pixel intensity of the central region,  $G_{ni}$  represents the average pixel intensity of the  $i$ th neighborhood, and  $k$  denotes the total number of neighborhoods. Since targets may appear in various orientations within shadows, when calculating the local contrast, it is advisable to select the neighborhood with the highest average pixel intensity for computation.

Computing the local contrast for all super-pixels in the image and binarizing them based on a preset threshold can obtain the local contrast information of the image. The acquired local contrast information is utilized to filter the ROIs. If a region does not contain pixels with the high local contrast, it is reasonable to assume that this region likely represents the shadow of a moving target. Based on this assumption, it can filter out the shadows of stationary targets from the ROIs, and then retain only the shadows of moving targets. The specific filtering method involves determining whether a region contains pixels with the high local contrast. If a region does not contain pixels with the high local contrast, the current region is retained. Otherwise, it is removed. The process of filtering ROIs based on the local contrast information is illustrated in Fig. 9. Utilizing local contrast information for filtering can effectively remove shadows of stationary targets.

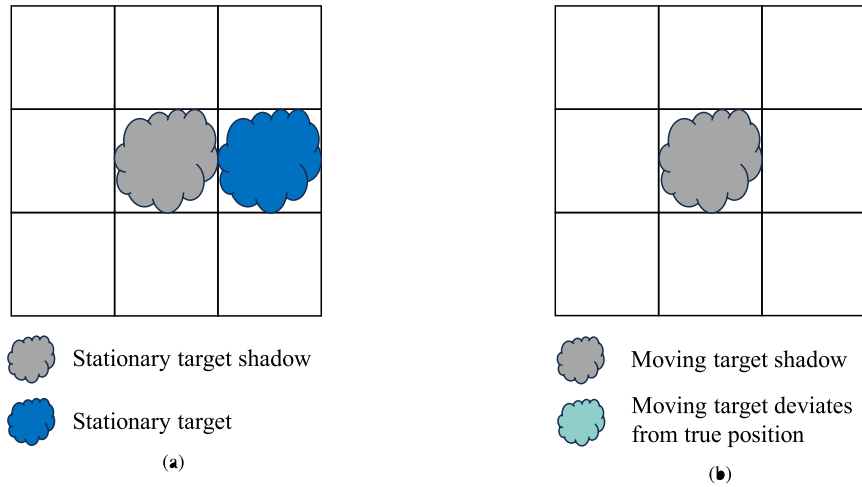


Fig. 6. Stationary and moving targets and their shadows. (a) Stationary target and its shadow. (b) Moving target and its shadow.

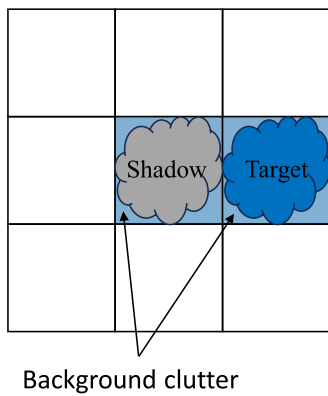


Fig. 7. Traditional local contrast detection window.

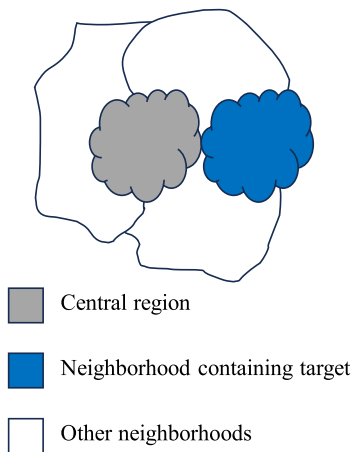


Fig. 8. Local contrast detection window based on the super-pixel.

C. Shadow Pixels Filling

After the shadow detection and removal of stationary shadow, the shadow regions of moving targets can be obtained. Next, it is necessary to fill in the shadow regions of the moving targets to complete the background reconstruction. According to the

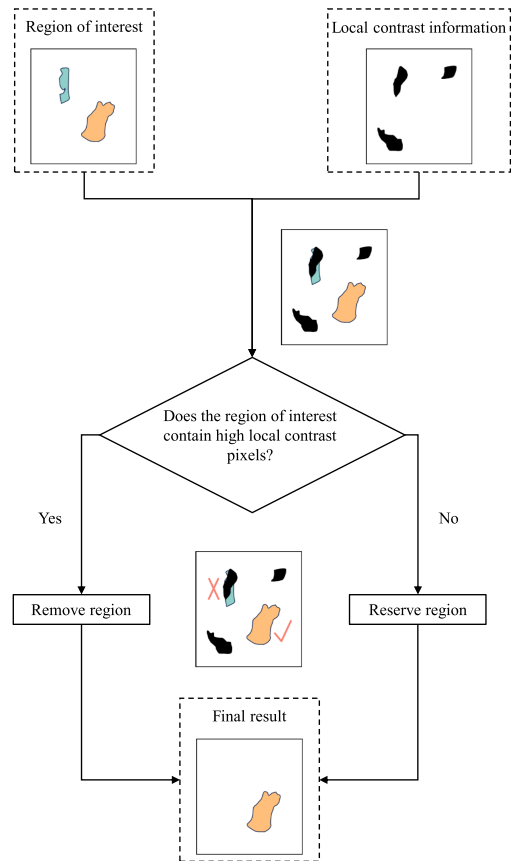


Fig. 9. Flowchart for filtering ROIs using the obtained local contrast information.

Markov random field theory, the background of shadow regions can be reconstructed using neighboring pixels of the shadow. The specific shadow filling method involves the random sampling of the shadow neighborhood, followed by filling the sampled neighborhood pixel values into the shadow region to complete the background reconstruction. The schematic of the shadow filling is shown in Fig. 10.

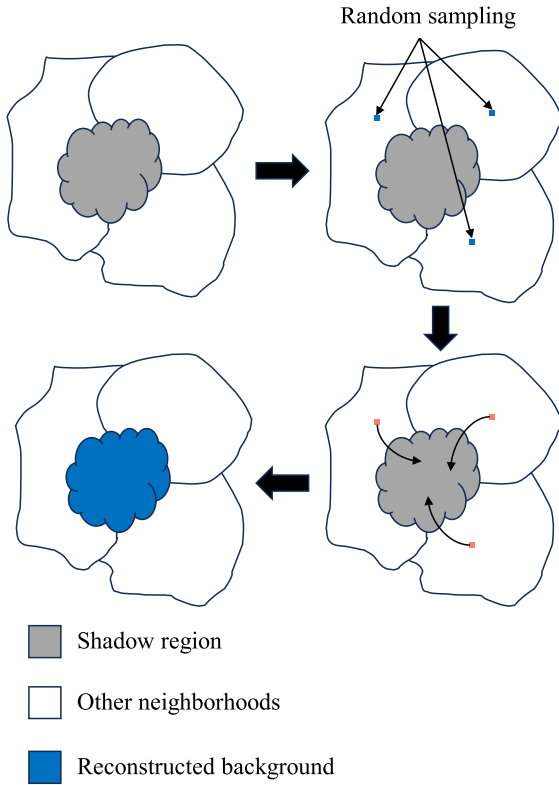


Fig. 10. Schematic of shadow filling.

### III. IMPROVED ViBe ALGORITHM

The ViBe algorithm is a nonparametric background modeling method distinguished by its rapid response, robust performance, and high detection accuracy. It fits the probability distribution of the background at each point directly based on the frequency of appearance of the background sampling points, without being limited to any specific distribution or parameters. After the background reconstruction, there are no longer any moving target shadows in the scene. Thus, a background model can be constructed based on the reconstructed background image, and the ViBe algorithm can be used to continue the shadow detection of moving targets in VideoSAR images. The original ViBe algorithm was designed for the optical videos for the moving target detection and did not utilize the prior information that the detected targets in VideoSAR images are shadows. Therefore, this article improves the ViBe algorithm to make it more suitable for detecting moving target shadows in VideoSAR images. The improved ViBe algorithm can be summarized into six steps, i.e., the background model initialization, image pixel classification, pixel filtering, morphological filtering, background update, and false alarm suppression, as shown in Fig. 11.

#### A. Background Model Initialization

The original ViBe algorithm repeats a random strategy 20 times within a  $3 \times 3$  neighborhood, randomly selecting one pixel each time. A total of 20 pixels are chosen as the background sample set for the central pixel, thus completing the initialization of the background model. This characteristic grants the ViBe

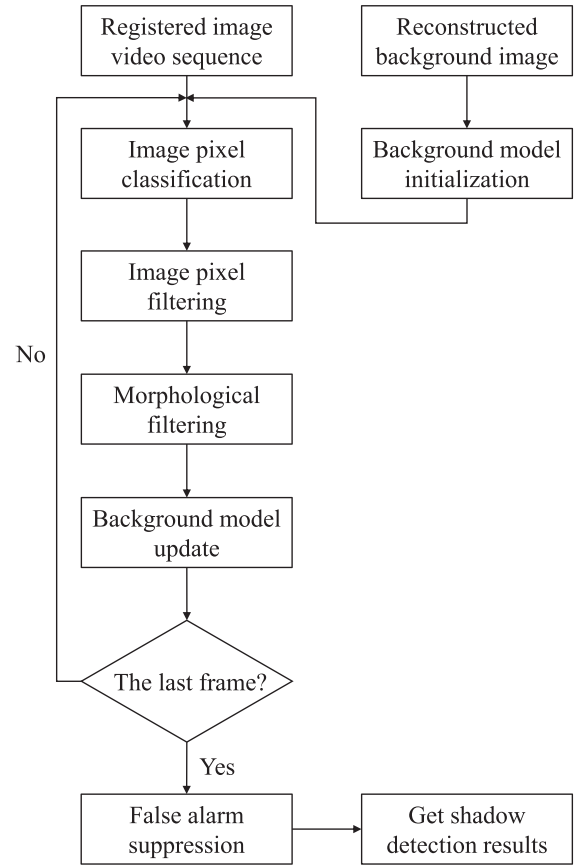


Fig. 11. Flowchart of the improved ViBe algorithm.

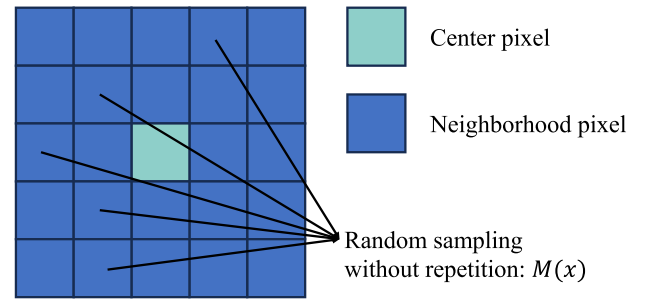


Fig. 12. Diagram of the background modeling in the improved ViBe algorithm.

algorithm strong robustness, enabling it to effectively address the noise interference. However, selecting 20 pixels randomly within each eight-neighborhood inevitably results in the selection of duplicate pixels [28]. To address this issue, this article opts to expand the neighborhood and perform 20 nonrepetitive samplings within a  $5 \times 5$  neighborhood, which can complete the initialization of the background model, as illustrated in Fig. 12. The background model sample set  $M(x)$  for pixel  $x$  can be represented as

$$M(x) = \{v_1, v_2, v_3, \dots, v_{20}\} \quad (4)$$

where  $v_i$  represents the pixel value sampled at the  $i$ th time in the neighborhood of pixel  $x$ .

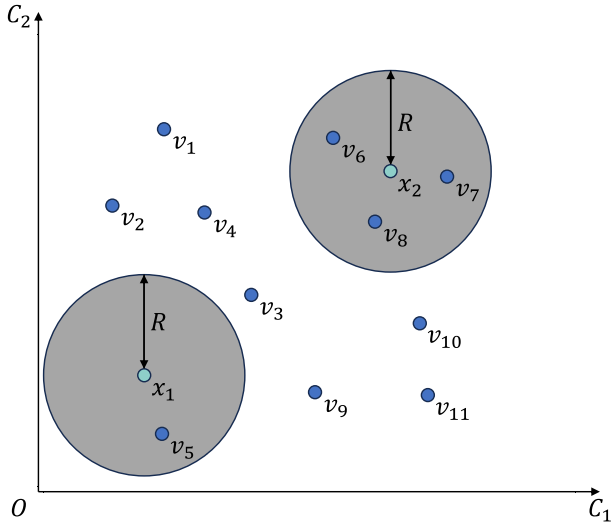


Fig. 13. Schematic of the pixel classification process.

### B. Image Pixel Classification

After completing the initialization of the background model, the next step is to classify the image pixels. That is to determine whether the image pixels belong to the foreground or the background. The classification process is shown in Fig. 13. The original ViBe algorithm first calculates the Euclidean distance in color space between each pixel  $x$  to be classified and the pixels in the background model sample set  $M(x)$ . The formula for calculating the distance,  $d_i(x)$ , which is given by

$$\begin{aligned} d_i(x) &= \sqrt{(v(x) - v_i)^2} \\ &= |v(x) - v_i|, \quad i = 1, 2, 3, \dots, 20 \end{aligned} \quad (5)$$

where  $v(x)$  denotes the gray value of the pixel point  $x$ . Define the distance threshold  $R$ . If the number of samples in the background model sample set  $M(x)$  whose distance from the current pixel point  $x$  is less than  $R$  is less than the threshold  $\text{Min}$ , the pixel point  $x$  is adjudicated as the foreground. Conversely, the pixel point is adjudicated as the background. The formula is expressed as

$$x = \begin{cases} \text{foreground, } [v(x) - R, v(x) + R] \cap M(x) < \text{Min} \\ \text{background, } [v(x) - R, v(x) + R] \cap M(x) \geq \text{Min.} \end{cases} \quad (6)$$

It is known that the target to be detected is the shadow of a moving target, and the grayscale value of the shadow region is lower than that of the background. When the shadow region moves to pixel  $x$ , the trend of the grayscale value of the pixel  $x$  should change from high to low. If the distance defined by (5) is used, when a nonshadow region with a high grayscale value moves to pixel  $x$ , the trend of the grayscale value of the pixel  $x$  changes from low to high, resulting in  $|v_i - v(x)| > R$ , which tends to classify the current pixel  $x$  as the foreground during the image pixel classification. However, when the trend of the grayscale value of the pixel  $x$  changes from low to high, it should be more inclined to classify the pixel  $x$  as the background. To

address this issue, the distance  $d_i(x)$  is redefined as

$$d_i(x) = v_i - v(x), \quad i = 1, 2, 3, \dots, 20. \quad (7)$$

When the trend of the gray value of the pixel  $x$  is from low to high,  $v_i - v(x) < -R < R$ , the image pixel classification is performed with a tendency to adjudicate the current pixel  $x$  as the background, as shown in Fig. 14. The judgment formula for the pixel point  $x$  is

$$x = \begin{cases} \text{foreground, } [-\infty, v(x) + R] \cap M(x) < \text{Min} \\ \text{background, } [-\infty, v(x) + R] \cap M(x) \geq \text{Min.} \end{cases} \quad (8)$$

In addition, the distance threshold  $R$  is determined by the degree of dispersion of the background model gray values. The distance threshold  $R$  is calculated by the formula [29]

$$R = \frac{m}{0.68\sqrt{2}} \quad (9)$$

where  $m$  denotes the median of  $|v_i - v_{i+1}|$ .

### C. Pixel Filtering and Morphological Filtering

In the previous section, although modifying the calculation of the distance  $d_i(x)$  can suppress pixel classification errors caused by grayscale value changes from low to high, there are still many high grayscale value pixels classified as the foreground. This is because the grayscale value trend of some pixels changes from high to low, but after the change, the grayscale value remains high. To address this issue, considering the characteristic that the average pixel value of moving target shadow areas in VideoSAR images is lower than the background pixel value, a background threshold  $T$  can be preset to filter out pixels classified as the foreground that clearly do not belong to shadows, and reclassify high grayscale value pixels as the background. If the pixel value is higher than the background threshold  $T$ , then the pixel is reclassified as the background. If the pixel value is lower than the background threshold  $T$ , then the pixel is classified correctly. The formula is represented as follows:

$$x = \begin{cases} \text{background,} & v(x) > T \\ \text{foreground,} & v(x) \leq T. \end{cases} \quad (10)$$

The results obtained after the reclassification using the background threshold  $T$  still contain speckle noise, road edges, and other areas that are not shadows of moving targets. Morphological filtering and connected component analysis can be used to eliminate the influence of these factors.

### D. Background Model Update

During the motion detection process using video image, the scene depicted in the video can undergo changes due to a variety of factors, including camera jitter, falling leaves, or other dynamic environmental elements. In SAR videos, changes in the scattering coefficient of targets at different angles can cause variations in the final image scene intensity, as shown in Fig. 15. Since the ViBe algorithm compares the current pixel value with the sample values in the background model set, it is necessary to update the background model in real time when using the ViBe

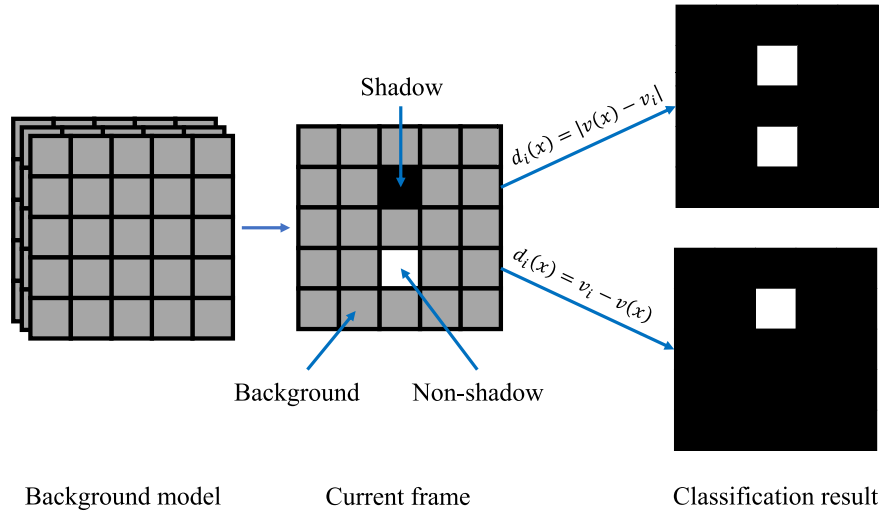


Fig. 14. Schematic of the pixel classification using two distances.

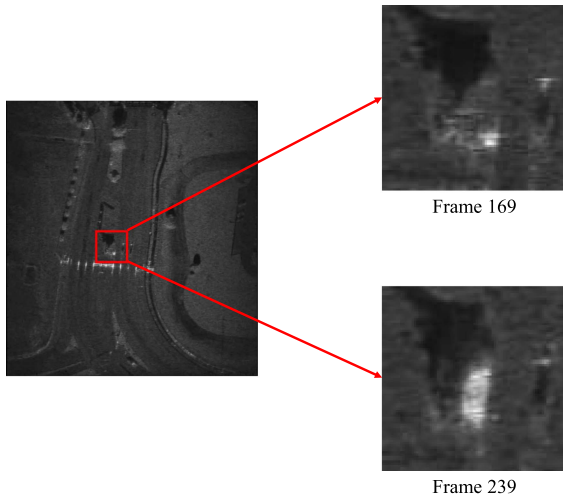


Fig. 15. Changes in the intensity of the imaging scene.

algorithm for the motion detection to ensure the more accurate detection results. The background model updating method used by the ViBe algorithm is conservative updating, meaning that a pixel is only included in the background model set when it is classified as the background [30].

### E. False Alarm Suppression

After the aforementioned steps, preliminary detection results of moving target shadows in each frame of the SAR image have been obtained. However, the previous analysis was based on single-frame images, without utilizing the interframe information, which still leads to many false alarms. To address this issue, the next step involves false alarm suppression using the joint detection results of multiple frames.

Real moving targets typically exhibit regular motion trajectories, while false targets appear the more chaotic in multiple frame detections. Exploiting this characteristic, the detection results of

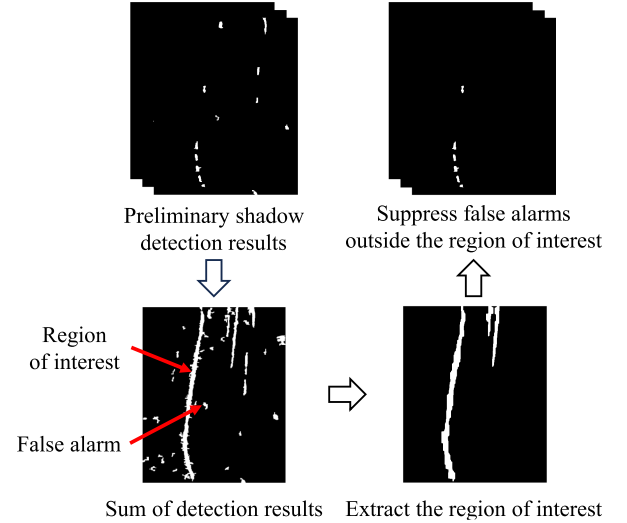


Fig. 16. Suppressing false alarms outside the ROIs.

all frames are summed up. If it is a real moving target, a long rectangular connected region will appear in the summation result. Utilizing the morphological filtering and connected component analysis based on this, the ROIs can be extracted. This approach helps in suppressing false alarms outside the ROIs, as depicted in Fig. 16.

In addition, another approach involves utilizing shadow detection methods based on the super-pixel segmentation for each frame to extract the shadow information. Then, this shadow information can be used to suppress false alarms within the ROIs. The specific procedure is similar to the one used in the background reconstruction and will not be reiterated here.

## IV. EXPERIMENT RESULTS AND ANALYSIS

In this section, the moving target detection was performed on a sequence of VideoSAR images using the proposed improved ViBe algorithm. A comparative analysis was conducted



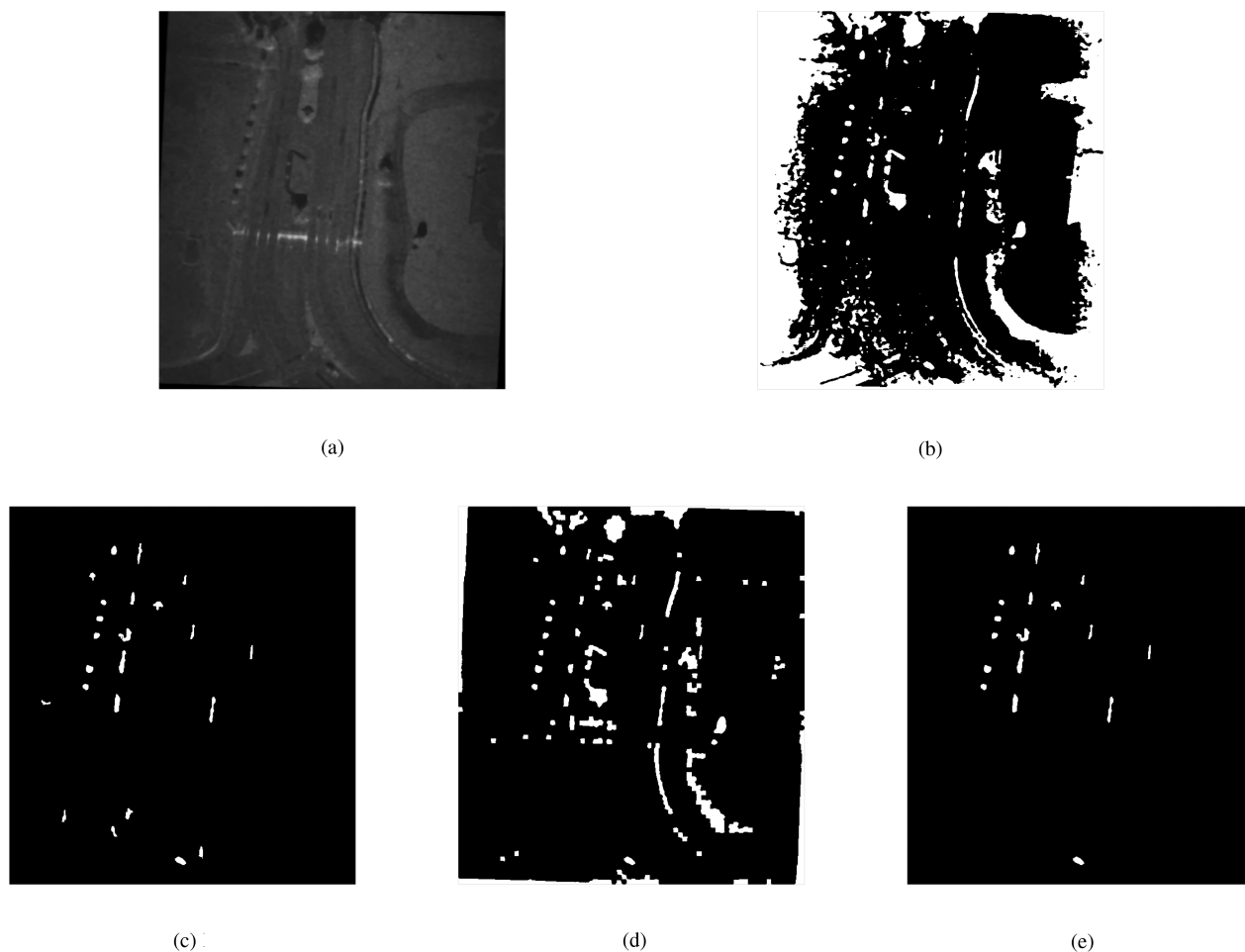


Fig. 17. Result of the shadow detection. (a) Original image. (b) Binary image. (c) Morphological filtering results. (d) Shadow information. (e) Result of the shadow information filtering.

with the results obtained from the original ViBe algorithm [17] and background subtraction algorithm [31]. The high-resolution VideoSAR data used in the experiment was obtained from the Sandia National Laboratories and captured a scene at the entrance of Kirtland Air Force Base, which features moving vehicles. In addition, the scene included various static targets, such as the trees, roads, buildings, and artificial islands. A subset of 100 frames of SAR images (frames from 153 to 252) was selected for the experiment, and preprocessing steps such as the registration, removal of invalid pixels, median filtering to suppress speckle noise, and intensity normalization were applied to the selected SAR images.

#### A. Background Reconstruction

The background reconstruction algorithm was applied to a single frame SAR image. The algorithm consists of three main steps, the shadow detection, removal of stationary target shadows, and pixel filling. The shadow detection results are shown in Fig. 17. From Fig. 17, it can be observed that the OTSU algorithm effectively extracts the dark regions in the image. After the morphological filtering, regions that are clearly not

shadows of moving targets are removed based on features, such as the area, aspect ratio, and rectangularity. Finally, using the shadow information extracted by the shadow detection method based on the super-pixel segmentation, nonshadow regions are effectively eliminated.

After removing the shadows of stationary targets, the result is shown in Fig. 18. Local contrast information is extracted from the image after performing the super-pixel segmentation. Then, utilizing this extracted local contrast information effectively eliminates the shadows of stationary targets, leaving only the shadows of moving targets. Finally, the shadows of moving targets are filled using neighboring pixels to complete the background reconstruction, as shown in Fig. 19. After the reconstruction, there are no longer any shadows of moving targets in the background. The reconstructed background image is applied to both the improved ViBe algorithm and the original ViBe algorithm. Since the background model in the background subtraction algorithm cannot be updated in real time, there is a significant difference between the backgrounds of SAR images taken at different times. If the reconstructed background image is used as the background model, it would lead to many false alarms caused by background differences in the detection results.

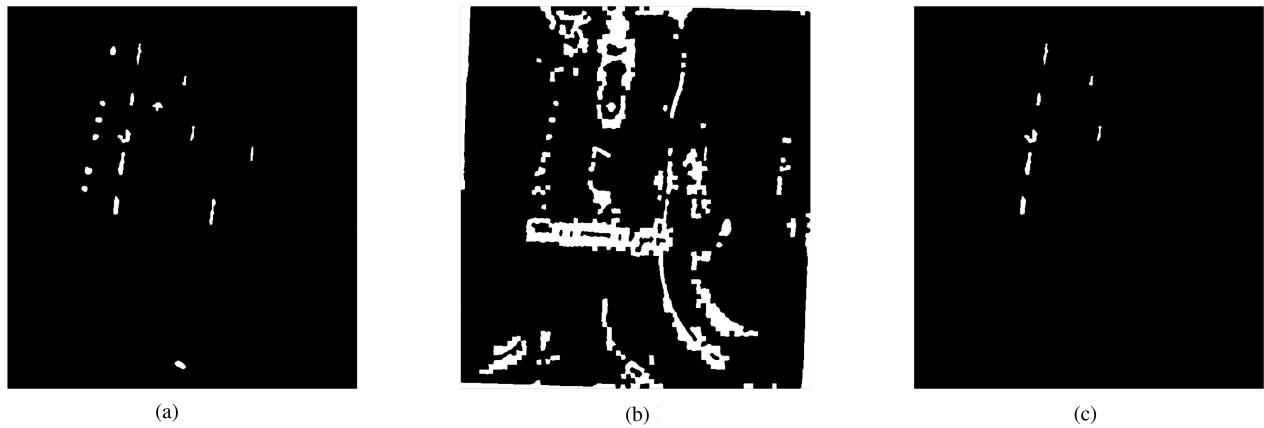


Fig. 18. Result after removing the shadows of stationary targets. (a) ROI. (b) Local contrast information. (c) Result after filtering with the local contrast information.

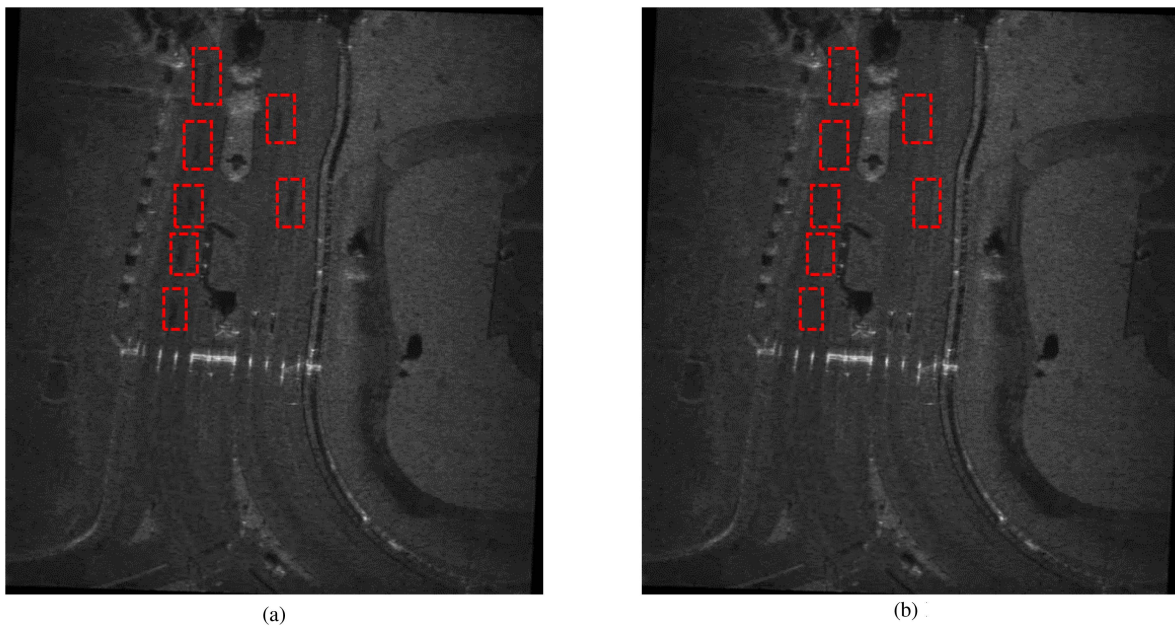


Fig. 19. Background reconstruction. (a) Original image. (b) Background reconstructed image.

Therefore, the mean image of 100 frames of SAR images is used as the background model for the background subtraction algorithm in the subsequent comparative verification experiments.

### B. Improved ViBe Algorithm

Fig. 20 shows the moving target shadow detection results of six different methods. Comparing the improved ViBe algorithm, the original ViBe algorithm with the false alarm suppression, and the background subtraction algorithm with the false alarm suppression. The detection results of the improved ViBe algorithm have only one missed detection, with no false alarms. The detection results of the original ViBe algorithm with the false alarm suppression have two missed detections and one false alarm. The detection results of the background subtraction algorithm with the false alarm suppression have one missed detection

and two false alarms. Among these three moving target shadow detection methods, the detection performance of the improved ViBe algorithm is the best. Comparing with the improved ViBe algorithm without the false alarm suppression, the original ViBe algorithm, and the background subtraction algorithm, it is also evident that the improved ViBe algorithm has the fewest missed detections and false alarms, and its detection performance is the best. This validates that the proposed improved ViBe algorithm can effectively detect moving target shadows. Comparing the improved ViBe algorithm with and without the false alarm suppression, the original ViBe algorithm with and without the false alarm suppression, and the background subtraction algorithm with and without the false alarm suppression, it can be seen that all three methods show a significant reduction in the false alarms after the false alarm suppression, removing false alarms outside the ROI as well as false alarms in nonshadow regions within the

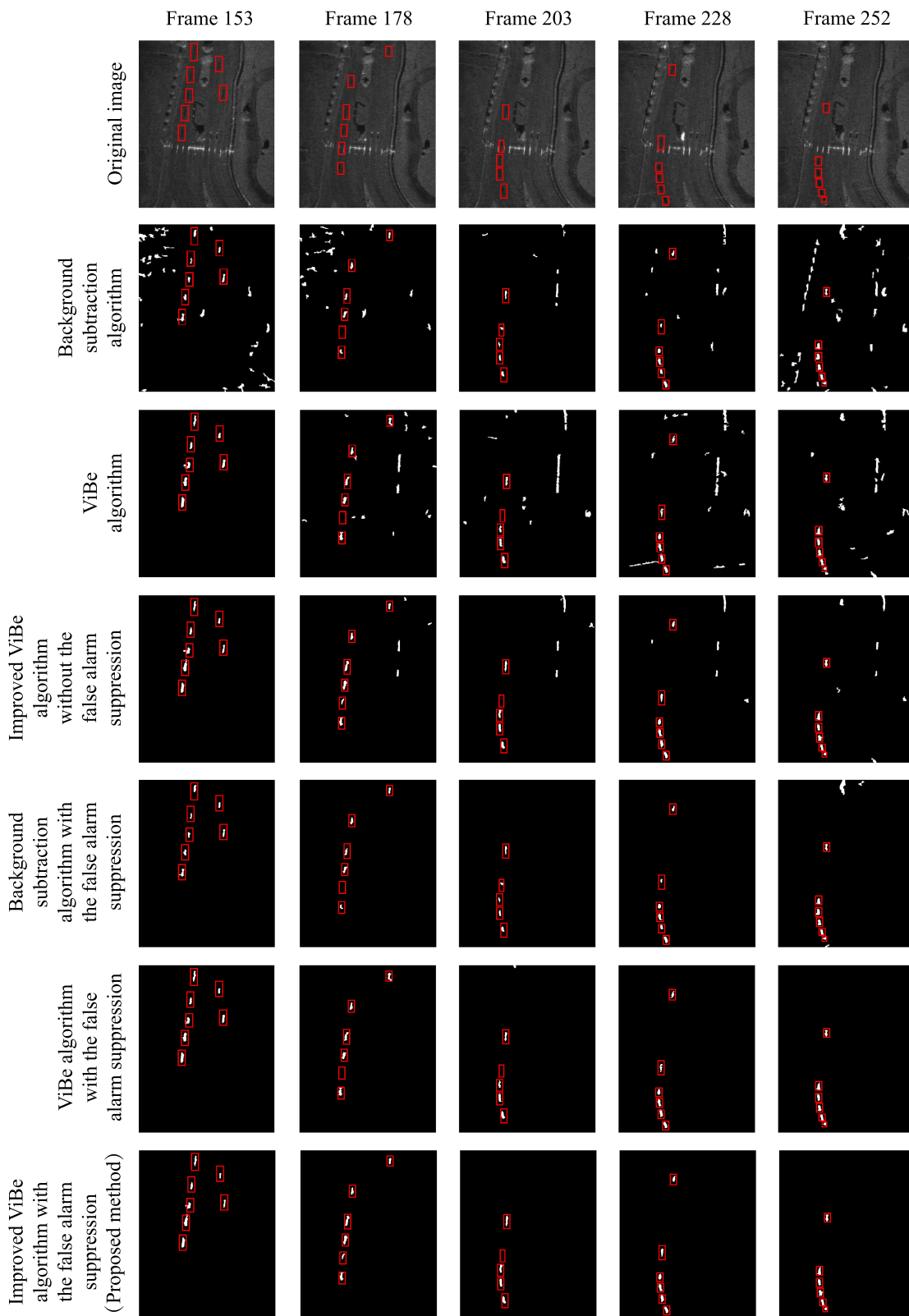


Fig. 20. Detection results of moving target shadows. The ground truth positions are marked with red boxes in both the original images and the detection results of all methods.

TABLE I  
TABLE OF STATISTICS FOR MOVING TARGET SHADOW DETECTION RESULTS

Methods	Correct detection number	Missed detection number	False alarm number	Detection rate(%)
Background subtraction algorithm	567	31	1780	94.82
ViBe algorithm	575	23	1357	96.15
Improved ViBe algorithm without the false alarm suppression	584	14	427	97.66
Background subtraction algorithm with the false alarm suppression	567	31	47	94.82
ViBe algorithm with the false alarm suppression	575	23	19	96.15
Improved ViBe algorithm with the false alarm suppression (Proposed method)	<b>584</b>	<b>14</b>	<b>8</b>	<b>97.66</b>

The blue values represent the best performance of all methods.

ROI, verifying the correctness and effectiveness of the proposed false alarm suppression method. Table I summarizes the moving target shadow detection results of 100 frames of SAR images. From Table I, it can be seen that the improved ViBe algorithm has the fewest missed detections and false alarms, with a detection rate of 97.66% and only 8 false alarms. The number of false alarms is significantly reduced after the false alarm suppression for all three methods. Once again, this validates the excellent moving target shadow detection performance of the proposed improved ViBe algorithm.

## V. CONCLUSION

In this article, a moving target shadow detection method based on the improved ViBe algorithm in VideoSAR images is investigated. First, the shadow information and local contrast information of the VideoSAR image are extracted by using the super-pixel segmentation technique to remove the stationary target shadow and filter out the moving target shadow. The moving target shadows are filled using neighboring pixels to achieve the background reconstruction. Besides, the reconstructed background image is used to initialize the background model of the ViBe algorithm, and the ViBe algorithm is improved by modifying the definition of the distance in the image pixel classification, adding the image pixel screening, and performing the morphological filtering. The improved ViBe algorithm is more suitable for the moving target shadow detection in VideoSAR. Finally, a false alarm suppression method utilizing joint multiframe detection results suppresses the vast majority of false alarms. Experiments were conducted on 100 frames of VideoSAR images from Sandia National Laboratories, and the experimental results showed that the moving target shadow detection rate reached 97.66%, and only 8 false alarms appeared, which verified the correctness and effectiveness of the proposed method.

## REFERENCES

- [1] J. Zhang, H. Xie, L. Zhang, and Z. Lu, "Information extraction and three-dimensional contour reconstruction of vehicle target based on multiple different pitch-angle observation circular synthetic aperture radar data," *Remote Sens.*, vol. 16, no. 2, pp. 401–416, Jan. 2024.
- [2] H. Xie, J. He, Z. Lu, and J. Hu, "Two-level feature-fusion ship recognition strategy combining HOG features with dual-polarized data in SAR images," *Remote Sens.*, vol. 15, no. 18, pp. 4393–4405, Sep. 2023.
- [3] X. Hu, H. Xie, S. Yi, L. Zhang, and Z. Lu, "An improved NLCS algorithm based on series reversion and elliptical model using geosynchronous spaceborne-airborne UHF UWB bistatic SAR for oceanic scene imaging," *Remote Sens.*, vol. 16, no. 7, pp. 1131–1154, Mar. 2024.
- [4] J. Zhu, Y. Song, N. Jiang, Z. Xie, C. Fan, and X. Huang, "Enhanced doppler resolution and sidelobe suppression performance for golog complementary waveforms," *Remote Sens.*, vol. 15, pp. 2452–2461, May 2023.
- [5] Z. Xie, Z. Xu, C. Fan, S. Han, and X. Huang, "Robust radar waveform optimization under target interpulse fluctuation and practical constraints via sequential lagrange dual approximation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 6, pp. 9711–9721, Dec. 2023.
- [6] Z. Xie, L. Wu, J. Zhu, M. Lops, X. Huang, and B. Shankar, "RIS-aided radar for target detection: Clutter region analysis and joint active-passive design," *IEEE Trans. Signal Process.*, vol. 72, pp. 1706–1723, Feb. 2024, doi: 10.1109/TSP.2024.3371292.
- [7] W. L. Melvin, "A STAP overview," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 19, no. 1, pp. 19–35, Jan. 2004.
- [8] D. Cerutti-Maori and I. Sikaneta, "A generalization of DPCA processing for multichannel SAR/GMTI radars," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 560–572, Jan. 2013.
- [9] B. Ge, L. Chen, D. An, J. Li, and Z. Zhou, "Parameter estimation of moving target in dual-channel circular synthetic aperture radar," in *Proc. IEEE Asia-Pac. Microwave Conf.*, Singapore, Dec. 2019, pp. 1381–1383.
- [10] A. M. Raynal, D. L. Bickel, and A. W. Doerry, "Stationary and moving target shadow characteristics in synthetic aperture radar," *Proc. SPIE*, vol. 9077, pp. 413–427, May 2014.
- [11] H. Fang, G. Liao, Y. Liu, C. Zeng, X. He, and Q. Meng, "A dual-mode framework for robust long-term tracking in video SAR," *IEEE Sens. J.*, vol. 24, no. 8, pp. 13028–13042, Apr. 2024.
- [12] Y. Zhang, X. Mao, H. Yan, D. Zhu, and X. Hu, "A novel approach to moving targets shadow detection in VideoSAR imagery sequence," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Fort Worth, TX, USA, Jul. 2017, pp. 606–609.
- [13] M. Jahangir, "Moving target detection for synthetic aperture radar via shadow detection," in *Proc. IET Int. Conf. Radar Syst.*, Edinburgh, U.K., 2007, pp. 1–5.
- [14] Z. Xie, Z. Xu, S. Han, J. Zhu, and X. Huang, "Modulus constrained minimax radar code design against target interpulse fluctuation," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 13671–13676, Oct. 2023.
- [15] X. Tian, J. Liu, M. Mallick, and K. Huang, "Simultaneous detection and tracking of moving-target shadows in ViSAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1182–1199, Feb. 2021.
- [16] C. Zhong, J. Ding, and Y. Zhang, "Joint tracking of moving target in single-channel video SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Oct. 2021, Art. no. 5212718, doi: 10.1109/TGRS.2021.3115491.
- [17] Z. Liu, D. An, and X. Huang, "Moving target shadow detection and global background reconstruction for VideoSAR based on single-frame imagery," *IEEE Access*, vol. 7, pp. 42418–42425, 2019.
- [18] J. Ding, L. Wen, C. Zhong, and O. Loffeld, "Video SAR moving target indication using deep neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 10, pp. 7194–7204, Oct. 2020.
- [19] H. Zhang and Z. Liu, "Moving target shadow detection based on deep learning in video SAR," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Brussels, Belgium, Jul. 2021, pp. 4155–4158.
- [20] J. Luan, L. Wen, and J. Ding, "Multifeature joint detection of moving target in video SAR," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Nov. 2022, Art. no. 4515805, doi: 10.1109/LGRS.2022.3222227.
- [21] L. Wen, J. Ding, and O. Loffeld, "Video SAR moving target detection using dual faster r-CNN," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2984–2994, Feb. 2021, doi: 10.1109/JSTARS.2021.3062176.
- [22] W. Zhang et al., "GNN-JFL: Graph neural network for video SAR shadow tracking with joint motion-appearance feature learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, Apr. 2024, Art. no. 5209117, doi: 10.1109/TGRS.2024.3383870.

- [23] J. Zhu et al., "Delay-doppler map shaping through oversampled complementary sets for high speed target detection," *Remote Sens.*, vol. 16, pp. 2898–2909, Aug. 2024.
- [24] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
- [25] R. Achanta and S. Süsstrunk, "Superpixels and polygons using simple non-iterative clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, HI, USA, Jul. 2017, pp. 4895–4904.
- [26] Y. Liu, M. Yu, B. Li, and Y. He, "Intrinsic manifold SLIC: A simple and efficient method for computing content-sensitive superpixels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 3, pp. 653–666, Mar. 2018.
- [27] Y. Liu, C. Yu, M. Yu, and Y. He, "Manifold SLIC: A fast method to compute content-sensitive superpixels," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 651–659.
- [28] J. Liu and H. Lin, "Adaptive threshold moving target detection algorithm based on ViBe method," *Laser Optoelectron, Prog.*, vol. 59, no. 2, pp. 1–10, Jan. 2022.
- [29] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," in *Proc. 6th Eur. Conf. Comput. Vis.*, Dublin, Ireland, Jan. 2000, pp. 751–767.
- [30] O. Barnich and M. Van Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1709–1724, Jun. 2011.
- [31] Z. He, Z. Li, X. Chen, A. Yu, T. Yi, and Z. Dong, "Detecting moving target on ground based on its shadow by using VideoSAR," *Remote Sens.*, vol. 13, no. 16, pp. 3291–3303, Aug. 2021.



**Zhitao Wu** received the B.S. degree in communications engineering in 2022 from the Sun Yat-Sen University, Shenzhen, China, where he is currently working toward the M.S. degree in information and communication engineering.

His research interests include CSAR-GMTI and moving target refocusing.



**Hongtu Xie** (Member, IEEE) received the B.S. degree in electronic information engineering from Hunan University, Changsha, China, in 2008, and the M.S. degree in electronic science and technology and the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2010 and 2015, respectively.

He is currently an Associate Professor and the M.S. Supervisor with the School of Electronics and Communication Engineering, Sun Yat-Sen University (SYSU), Guangzhou, China, and the judge's experts of the National Natural Science Foundation of China and the Guangdong Basic and Applied Basic Research Foundation. His research interests include high-resolution SAR imaging, and target detection and recognition.

Dr. Xie was the recipient of the Excellent Doctoral dissertation of the Chinese People's Liberation Army Award for his Ph.D. dissertation in 2018, and the Third Prize of Teaching Competition for the Young Teacher of the School of Electronics and Communication Engineering in 2020.



**Ting Gao** received the B.S. degree in computer science and technology from the Naval University of Engineering, Wuhan, China, in 2005, and the M.S. degree in computer application and the Ph.D. degree in image processing from the Air Force Early Warning Academy, Wuhan, China, in 2008 and 2014, respectively.

She is currently a Lecturer and Doctor with the Air Force Early Warning Academy and she has authored or coauthored more than 20 professional papers, published 3 monographs, and is responsible for more than 10 scientific research projects. Her research field is the intelligent target recognition.

Dr. Gao is a member of the First Batch of Innovative Teams and a recipient of the TuoQing Engineering Fund in the Air Force Early Warning Academy. She was the recipient of more than 10 military awards.



**Yuanjie Zhang** received the B.S. degree in electronic information engineering from the Nanjing University of Posts and Telecommunications University, Nanjing, China, in 2023. He is currently working toward the M.S. degree in information and communication engineering with Sun Yat-Sen University, Guangzhou, China.

His research interests include radar signal processing and target identification.



**Haozong Liu** received the B.S. degree in electronic science and technology in 2023 from the Sun Yat-Sen University, Shenzhen, China, where he is currently working toward the M.S. degree in electronic information.

His research interests include CSAR-MTI and autofocus.