Research Article

An improved ViBe-based approach for moving object detection

Guangyi Tang[1], Jianjun Ni[1,2], Pengfei Shi[1,2], Yingqi Li[1], Jinxiu Zhu[1]

[1]College of Internet of Things Engineering, Hohai University, Changzhou 213022, Jiangsu, China.
[2]Jiangsu Key Laboratory of Power Transmission & Distribution Equipment Technology, Hohai University, Changzhou 213022, Jiangsu, China.

Correspondence to: Prof. Jianjun Ni, College of Internet of Things Engineering, Hohai University, No.200, North Jinling Road, Xinbei District, Changzhou 213022, Jiangsu, China. E-mail: njjhhuc@gmail.com

Abstract

Moving object detection is a challenging task in the automatic monitoring field, which plays a crucial role in most video-based applications. The visual background extractor (ViBe) algorithm has been widely used to deal with this problem due to its high detection rate and low computational complexity. However, there are some shortcomings in the general ViBe algorithm, such as the ghost area problem and the dynamic background problem. To deal with these problems, an improved ViBe approach is presented in this paper. In the proposed approach, a mode background modeling method is used to accelerate the process of the ghost elimination. For the detection of moving object in dynamic background, a local adaptive threshold and update rate is proposed for the ViBe approach to detect foreground and update background. Furthermore, an improved shadow removal method is presented, which is based on the HSV color space combined with the edge detection method. Finally, some experiments were conducted, and the results show the efficiency and effectiveness of the proposed approach.

Keywords: Moving object detection; ViBe-based approach; dynamic background; shadow detection

1. INTRODUCTION

The real-time detection of moving objects is an essential task in the computer vision field, which has wide applications, including target tracking, video surveillance, abnormal behavior analysis, intelligent robot, etc[1–5]. There are still many challenges of the moving object detection under natural scenes, such as illumination changes, swaying leaves, and shadow changes[6,7]. Therefore, it has attracted more and more attention from researchers recently.

There are many research achievements in moving object detection. For example, Sengar and Mukhopadhyay[8] proposed a motion detection method using block based bi-directional optical flow method. Chen *et al.*[9] proposed an end-to-end deep sequence learning architecture for moving object detection. Li *et al.*[10] presented a novel technique for background subtraction based on the dynamic autoregressive moving average (ARMA) model. These methods used for moving objects detection can be divided into three main types: the optical flow method, the deep learning method, and the difference method. In addition, the difference methods are further divided into three categories[11,12], namely the unsupervised method[13], the supervised method[14,15], and the semi-supervised method[16,17]. There are some drawbacks in the optical flow method, such as complex computation and sensitivity to illumination mutation, which is not suitable for real-time moving objects detection[18]. Compared with traditional algorithms, deep learning methods have the advantages of high detection accuracy and strong fitting ability, but the size of the dataset determines the effect of detection, and it is difficult to meet the needs of deploying in some special scenarios at any time without sufficient samples. At the same time, they have higher requirements on the hardware environment, so the computational cost of deep learning-based algorithms is higher than that of traditional algorithms[19–21]. The background difference method has become the most widely used method for its outstanding superiorities in computation complexity and efficiency, which is the hot spot in moving object detection field[22]. However, the detection results of the background difference method depend on the accuracy of the background model. The way of establishing a robust background model is the key to this method.

There are many methods for moving object detection based on background difference methods, including Gaussian single model (GSM), Gaussian mixture model (GMM), and visual background extractor (ViBe) method[23,24]. ViBe algorithm is a sample-based moving object detection method, which has the advantages of less calculation, small footprint, and fast processing speed. It is suitable for the real-time detection of moving objects. Many researchers are focusing on the ViBe-based method of moving object detection. For example, Talab *et al.*[25] proposed an approach for moving crack detection in video based on ViBe and multiple filtering. Gao and Cheng[26] presented the use of the ViBe algorithm to extract smoke contours and shapes, which finally makes the detection of smoke root more accurate. However, there are some deficiencies of the general ViBe algorithm. For example, when the first frame of the video contains a moving object, there will be a ghost area left in the current location, which will need a long time to be removed. In addition, there is often a shadow problem in moving object detection based on the general ViBe algorithm.

To deal with the problems above for moving object detection based on ViBe method, various improvements have been proposed. For example, Huang *et al.*[27] proposed a moving target detection algorithm based on the improved ViBe algorithm by joining TOM (time of map) mechanism in the process of detection, where both the spatial domain and the time domain information of the pixels were used to eliminate the ghost area. Qiu *et al.*[28] presented a moving object detection method based on the strategy of ViBe algorithm and fused the infrared imaging features, which can establish the pure background in a variety of complex conditions. Yue *et al.*[29] introduced ant colony clustering algorithm and integrated it into the traditional ViBe framework and extended the ViBe based on local modeling to a global modeling algorithm, which can deal with the target adhesion problems but cannot effectively process shadows. The works above improve the performance of the ViBe-based method to some extent. However, few of them considered the problems comprehensively. For example, some methods considered the shadow problem, but they need a long computation time[30,31].

In this paper, an improved ViBe-based approach is proposed, where the problems of moving object detection under natural scenes are fully considered including the ghost area problem, the dynamic background problem, and the shadows problem, and some solutions are presented. Finally, various experiments were conducted under different scenes for moving object detection task. The results show the efficiency and effectiveness of the proposed approach.

The main contributions of this paper are summarized as follows: (1) A new background model based on mode background modeling method is proposed to eliminate the ghost areas quickly; (2) An improved ViBe approach is proposed based on an adaptive foreground detection and background updating method, where the value of the eight neighboring pixels difference between the background and the current frame is used. (3) A novel shadow elimination approach is presented, which is based on the HSV color space combined with the edge detection method. Furthermore, the computation time and background updating mechanism of the proposed approach are discussed.

This paper is organized as follows. Section 2 provides the related works about the ViBe-based method. Section 3 presents the improved ViBe-based method for moving object detection. The moving object detection experiments under various natural scenes are given in Section 4. Section 5 discusses the performance of the proposed approach. Finally, the conclusions are given in Section 6.

2. RELATED WORKS

In the past few years, various foreground target detection methods have been proposed to build powerful and flexible background models that can be used in surveillance scenarios with different challenges. One of the most widely used probabilistic models is the GMM[32], which models each pixel using a mixture of Gaussian models rather than modeling all pixel values as a distribution. For example, Kaewtrakulpong and Pakorn[33] modified the update equation of GMM for improving the accuracy and proposed a shadow detection scheme based on the existing GMM. Hofmann *et al.*[34] used a constantly adapted number of Gaussian distributions of the GMM for each pixel.

As for nonparametric approaches, Barnich and Droogenbroeck[35,36] proposed the ViBe-based method, where the current pixel value is compared to its closest sample within the collection of samples. First, the pixel values of the detected frames are matched with the corresponding models. The threshold value determines whether it belongs to the background or the foreground; for the matching pixel, the background model of the pixel and its neighborhood is updated by a random update mechanism. The method is simple to operate and detects well in static backgrounds but has fixed parameters. This limits the algorithm's ability to adapt to dynamic backgrounds (surface ripples, leaf shaking, etc.), and its neighborhood diffusion update strategy causes slower-moving foreground targets to blend into the background too quickly, increasing false detections. Its single-frame input image initialization strategy creates a "ghost" area when the input image contains foreground targets. In addition, there is often a shadow problem in moving object detection based on the general ViBe algorithm, which affects the accuracy of the background model.

To deal with the problems above for moving object detection based on ViBe method, various improvements have been proposed. For example, Zhu *et al.*[37] proposed a fast and efficient improvement of ViBe algorithm based on the edge characteristic info and neighborhood mean filter, but there are a lot of holes inside the detection area. Chen *et al.*[38] combined physical shadow theory and $C_1C_2C_3$ color space for the shadow removal. Yang *et al.*[39] used two thresholds to describe the uncertainty in the ViBe-based color video detection, and they used evidence theory to model and handle the uncertainty. Liu *et al.*[40] used the temporal and spatial information of the pixels to initialize the background model, and then combined the background sample set with the neighborhood pixels to determine the complexity of the background and obtain an adaptive segmen-
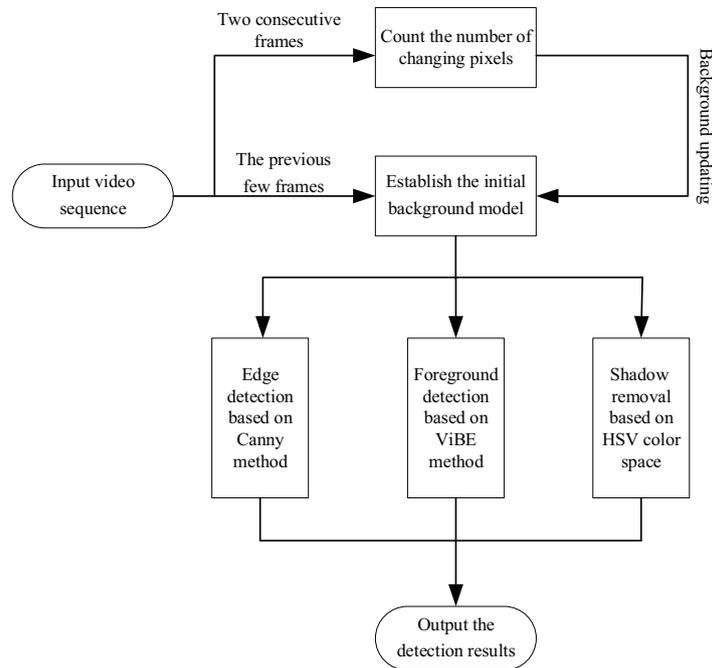
Figure 1. Flow diagram of the improved ViBe-based approach for moving object detection.

tation threshold, which can also obtain a better performance in complex dynamic backgrounds but cannot effectively remove shadows.

3. METHODS

In this paper, the problem of moving object detection based on ViBe method is studied. The basic idea of the ViBe method uses neighboring pixels to establish the background model and then compares the background model with the current pixel value to detect the foreground. There are three main steps in the ViBe method, namely the background initialization, the foreground object detection, and the background model updating. Aiming at the problems in the three main steps, some improvements are proposed in this study. The flow chart of the proposed approach is shown in Figure 1 and the main steps of the proposed approach are introduced in detail as follows.

3.1. Mode method based background modeling

The initial background selection is the first step in the ViBe-based method, which will directly influence the detection results. If it can be extracted correctly, the accuracy of the object detection will increase. In general, the ViBe method uses the first frame as the initial background[41], namely

$$B(x, y) = V_F(x, y) \tag{1}$$

where $B(x, y)$ is the pixel value of the background and $V_F(x, y)$ is the pixel value of the first frame in the video. Although the method using the first frame is simple and efficient, it will fail when there is a moving object in the first frame. To deal with this problem, some improvements are proposed, such as the mean method, which needs to store more video frames and has the problem of shadows[42]. In this paper, the mode method is introduced to extract the initial background frame[43]. The basic idea of the mode method for background modeling is that few previous frames are used to obtain an optimized background model. The pixel value of

//// **The pseudo-code of the Mode Background Method** ////

Initialize the parameters; Input the prepared image to matrix I$_m$;

For $n = 1{:}N\_Frame$ % *N_Frame is first few frames of video*;

   I_y = Convert_gray (I_m)   % *Convert the image I$_m$ to a gray image I$_y$*;

   Θ = Save_gray (I_y)   % *Save the grayscale value of I$_y$ to the array*;

End for

For $j = 1{:}N\_Pixel$ % *N_Pixel is the number of pixels in the image*;

   C = Ceil (Θ/E)

   % Ceil ( ) *is a function to rounds the elements of the data to the nearest integers towards*

   infinity; E is an integer number.

   Mf = Mode (C)

   % Mode ( ) *is a function to returns the sample mode of C, which is the most frequently*

   occurring value in C;

   Num = Count ($C{=}{=}Mf$)

   %Count ( ) *is a function to calculate the number of the pixel where C==Mf*

   $Mode\_save$ = Mf

   % *Save the mode value of Mf to the array*;

   $R\,(x,\,y)$ = Mode_save $(x,\,y){*}E$

   % *Calculate the value of the array*;

End for

Return $B\,(x,\,y) = R\,(x,\,y)$

   % *Output the initial background image I$_b$, which is constructed by the pixel B(x, y)*;

Figure 2. The pseudo-code of the mode background method.

the background is calculated by

$$B(x, y) = \frac{\sum_{k=1}^{Num} V_k\,(Mode(C))}{Num} \tag{2}$$

where $Mode(C)$ is a function to return the mode number of the sample, which is the most frequently occurring value in this sample. Num is the number of the mode numbers in the sample. Here, C is defined as follows:

$$C = Ceil(\Theta/E) \tag{3}$$

where $Ceil(\cdot)$ is a function to round the elements of the data to the nearest integers towards infinity. Θ is the grayscale value of the gray image obtained from the original image. To extract most of the numbers appearing in the array Θ by the function $Ceil(\cdot)$, the range of Θ is reduced by dividing an integer E. In this study, E is set as 5, namely the grayscale images are separated with five grayscale levels, which can improve the contrast of different elements in the image and reduce the influence of small speckles on target extraction. The pseudo-code of the mode method for background modeling is shown in Figure 2.

After the background of the video is established, the ViBe method is used to initialize the background model, which is based on the domain model. For each pixel $p(x, y)$ of the background image, the sample set $M(x, y)$ of it is:

$$M(x, y) = \{V_1, V_2, \cdots V_n\}, i = 1, 2, \cdots n \tag{4}$$

where n is the number of the neighboring sample. V_i is the value of a sample that is randomly chosen from the 8-connected neighborhood of each pixel (see Figure 3a). When the sample sets of all the pixels in the background are obtained, a background model is set up.
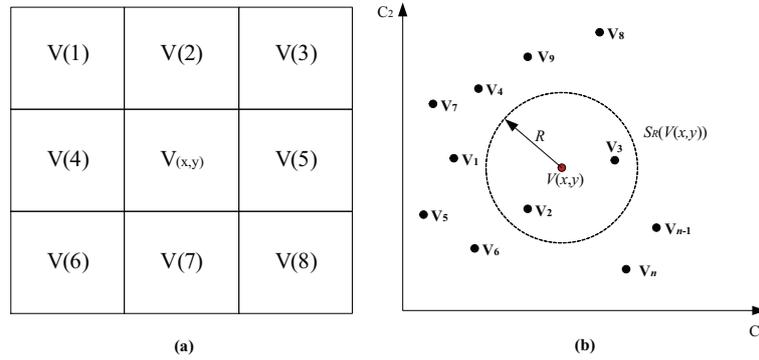
Figure 3. The ViBe-based method: (a) the eight neighbor domain; and (b) the background model of ViBe.

3.2. Adaptive updating mechanism for ViBe method

When the background of the video is established, the next step is to detect the moving objects. The basic discrimination mechanism for the general ViBe method is as follows: for each pixel in the new frame of the video, a sphere $S_R(V(x, y))$ of radius R centered on the value $V(x, y)$ of the pixel is defined (see Figure 3b). Then, the pixel of the new frame can be determined as the background or foreground by[44]:

$$Flag1(x, y) = \begin{cases} 1, & \Psi\{S_R(V(x, y)) \cap M(x, y)\} \leq K \\ 0, & \Psi\{S_R(V(x, y)) \cap M(x, y)\} > K \end{cases} \tag{5}$$

where function $\Psi\{S_R(V(x, y)) \cap M(x, y)\}$ means the cardinality of the set intersection of the sphere $S_R(V(x, y))$ and the collection of $M(x, y)$. K is a threshold. If $Flag1(x, y) = 1$, it means the pixel point $p(x, y)$ belongs to foreground. Otherwise, it means the pixel point $p(x, y)$ belongs to background.

The last step is to randomly update the background model with each new frame. Because of the strong statistical correlation between a pixel and its neighboring pixel, when a pixel is detected as the background pixel, it has a probability of $1/\alpha$ to update model sample set (where α is called update rate). Meanwhile, it also has the probability of $1/\alpha$ to update the background model of neighboring pixels.

From the discrimination mechanism of the original ViBe algorithm in Figure 3a,b, we can see that the detection radius R and the update rate α are two very important parameters. In general, the detection radius R should be larger and the update rate α should be smaller in the dynamic background, to make more pixels be classified as background, and vice versa. However, in the general ViBe algorithm, the values of the parameters R and α are predefined by the designers, which reduce the adaptivity of the ViBe algorithm. Because the value of the eight neighboring pixels difference between the background and the current frame is the factor that can reflect the complex degree of background, it is used to determine the values of the detection radius R and the update rate α adaptively. Namely,

$$R = \begin{cases} R_0 \cdot (1 + a), & a > \tau_0 \\ R_0 \cdot (1 - a), & a \leq \tau_0 \end{cases} \tag{6}$$

$$\alpha = \begin{cases} \alpha_0 \cdot (1 - a), & a > \tau_0 \\ \alpha_0 \cdot (1 + a), & a \leq \tau_0 \end{cases} \tag{7}$$

where R_0 and α_0 are the initial values of the detection radius R and the update rate α; τ_0 is a threshold; and a is a parameter to judge the change of the current scenario, which is calculated by

$$a = \frac{\sum D_{k+1}(x, y)}{N} \tag{8}$$

Here, N is the number of pixels. $D(x, y)$ is the difference of the pixels between two images $I_{k+1}(x, y)$ and $I_k(x, y)$, namely

$$D_{k+1}(x, y) = \begin{cases} 0, & |I_{k+1}(x, y) - I_k(x, y)| < \tau_1 \\ 1, & \text{Otherwise} \end{cases} \tag{9}$$

where τ_1 is a threshold to reduce the effects of the moving objects.

Remark: The mode background method can eliminate the foreground target that appears in the previous frames. The subsequent frames of the video sequence continuously update the "ghost" area to set it as the background, which can effectively speed up the "ghost" area removal.

3.3. Shadow removal strategy

Shadow is a common problem in moving object detection, and how to remove the shadow is a hot topic in the field of computer vision [45,46]. In this paper, an improved method based on the HSV color space is used to complete the shadow removal task. The main reason for using the HSV color space is that it is very close to the characteristics of human vision considering the existing methods, which is more accurate than RGB color space for shadow removal. However, there are many parameters of the traditional HSV that need to be set in different video environments, such as the thresholds used for the shadow judgment [47]. In addition, when there is no significant difference on the color attribute between the moving object and the shaded area, the accuracy of shadow removal based on the traditional HSV color space will be decreased. To deal with these problems, an improved shadow removal strategy is proposed in this paper. The basic idea of the proposed method is that the shadow area can be effectively distinguished by using the characteristics of shadow intensity reduction and color invariance theory, because the HSV color space can directly reflect the color characteristics of the image. The main procedures of the proposed method are as follows:

(1) The HSV space transformation is done. Then, the values of the H, S, and V components are obtained. Since the value V is a direct measure of the brightness of the color, the brightness of these pixels is significantly reduced in the shadow part. The difference of the brightness is denoted as $D_V(x, y)$, which is defined as follows:

$$D_V(x, y) = t_V(x, y)/B_V(x, y) \tag{10}$$

where $t_V(x, y)$ is the V value of current image frame. $B_V(x, y)$ is the V value of background frame. For any pixel points $p(x, y)$, the brightness value between the current frame and background frame is used to determine whether the current pixel is a shadow point. The decision strategy is as follows:

$$Flag2(x, y) = \begin{cases} 1, & \tau_2 \leq D_V(x, y) \leq \tau_3 \\ 0, & \text{Otherwise} \end{cases} \tag{11}$$

where $Flag2(x, y)$ is a flag. τ_2 and τ_3 are two thresholds for shadow detection.

(2) When the chromaticity of the object is similar to the shadow, the shadow area will be enlarged based on the brightness detection above. To deal with this problem, an improved method is proposed based on the forming mechanism of shadow. Namely, for each shadow pixel $p(x, y)$, its darkness level is limited, because it is darkened for the blocking out of the illumination source, but there is the presence of ambient illumination. In addition, the shadow pixels are mostly in gray areas. The decision strategy is as follows:

$$Flag3(x, y) = \begin{cases} 1, & t_S(x, y) \leq \lambda_1 \text{ and } t_V(x, y) \geq \lambda_2 \\ 0, & \text{Otherwise} \end{cases} \tag{12}$$

where $Flag3(x, y)$ is a flag; $t_S(x, y)$ is the saturation of the pixel point of current image frame; λ_1 is the maximum value of the saturation in the gray range; and λ_2 is the minimum value within the gray range. Then, the shadow area can be detected by:

$$I(x, y) = \begin{cases} 1, & Flag2(x, y) = 1 \text{ and } Flag3(x, y) = 1 \\ 0, & \text{Otherwise} \end{cases} \tag{13}$$

At the same time, to ensure the integrity of the foreground targets, the Canny edge detection is performed after finding the different image between the current frame and the background frame[48].

The whole work flow of the proposed approach for the moving objects detection is as follows:

Step1: Initialize the background model based on the mode method.

Step2: Convert the current frame and the background model to gray space, and then detect the foreground objects which include shadows, based on the proposed ViBe algorithm with the adaptive detection radius R and update rate α.

Step3: Convert HSV color space transformation for the current frame and detect the shadows by the color invariance theory at the shadow and the background.

Step4: Carry out an "AND" operation on the results obtained from Steps 2 and 3 to remove the shadow of the foreground targets.

Step5: Find the difference image between the current image frame and the background frame and perform Canny edge detection.

Step6: Carry out an "OR" operation on the results obtained from Steps 4 and 5 to ensure the integrity of the foreground objects.

4. RESULTS

To test the performance of the proposed approach, some experiments were carried out on several benchmark datasets including Highway, Bungalow, Cars, and People[49,50]. These experiments were coded by Python on a computer with 8G RAM and i7-4720HQ 2.60GHz CPU. Seven indices were used to evaluate the performance of detection: recognition rate of foreground (RE), recognition rate of background (SP), false positive rate (FPR), false negative rate (FNR), percentage of wrong classification (PWC), precision (PRE), and F-score (F) (see[51] for the details of these indices). For these indices, the larger are the RE, SP, PRE, F, the more accurate is the detected target area, and the smaller are the FPR, FNR, PWC, the more accurate is the detected background. The values of the parameters used in these experiments are the same and listed in Table 1. To show the efficiency of the proposed improved approach (I-ViBe), it was compared with the Gaussian mixture model based method (GMM) and the general ViBe-based method (G-ViBe). In the general ViBe-based method, the detection radius R and the update rate α are equal to R_0 and α_0 in the proposed approach.

4.1. The experiment for single object detection

To test the basic performance of the proposed approach, two experiments were conducted where only one object was detected. The datasets used for this experiment were Walk (Clip1) and Bungalows (Clip2). Two clips of the two videos were used to test the three detection methods, where the frame with the moving object was used as the detection frame (see Figure 4b). The results of the two experiments are shown in Figure 4. The evaluations for the three methods are listed in Table 2.

The results in Figure 4 show that all the three methods can detect the moving object effectively in this simple experiment, and the results in Table 2 show that the proposed approach has better detection results in most of the indices than the other two methods. In addition, the detection results on Walk (Clip1) show that the general ViBe cannot deal with the ghost problem, while the proposed ViBe can remove the ghost area very well. The detection results on Bungalows (Clip2) show that the proposed ViBe can remove the shadow more

Table 1. Parameters of the proposed method

Parameters	Values	Remarks
E	5	A given threshold in Equation (3)
K	1	A given threshold in Equation (5)
R_0	20	The initial detection radius
α_0	16	The initial update rate
τ_0	0.2	A given threshold in Equations (6) and (7)
τ_1	1	A given threshold in Equation (9)
τ_2	0.2	A given threshold in Equation (11)
τ_3	0.7	A given threshold in Equation (11)
λ_1	43	A given threshold in Equation (12)
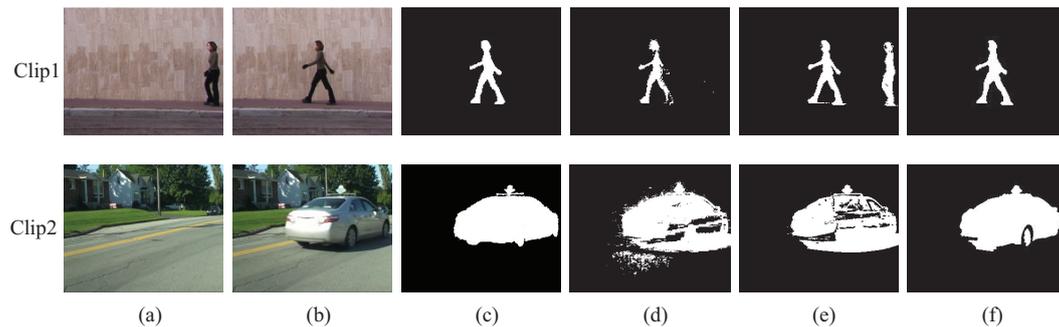
Figure 4. The moving object detection experiments on the video Walk (Clip1) and Bungalows (Clip2): (a) the first frame; (b) the frame for detection; (c) the ground-truth; (d) the result of GMM; (e) the result of G-ViBe; and (f) the result of I-ViBe.

Table 2. The valuation of the three methods for moving object detection in Walk and Bungalows

The valuation indices	The video clip of Walk			The video clip of Bungalows		
	GMM [32]	G-ViBe [35]	I-ViBe	GMM [32]	G-ViBe [35]	I-ViBe
SP	**0.9995**	0.9804	0.9985	0.9310	0.9401	**0.9854**
RE	0.7758	0.9483	**0.9612**	0.8570	0.7999	**0.9636**
FPR	**0.0004**	0.0195	0.0014	0.0689	0.0598	**0.0145**
FNR	0.2241	0.0516	**0.0387**	0.1429	0.2000	**0.0363**
PWC	0.0060	0.0203	**0.0021**	0.0826	0.0879	**0.0186**
PRE	**0.9779**	0.5647	0.9272	0.7395	0.7702	**0.9390**
F	0.8652	0.7079	**0.9493**	0.7939	0.7847	**0.9511**

effectively than the other two methods (see Figure 4e,f).

4.2. The experiment for multiple objects detection

To test the performance of the proposed approach in multiple moving objects detection, two experiments were conducted on the dataset Highway (Clip1) and People (Clip2). The results are shown in Figure 5, and the evaluations for the three methods in this experiment are shown in Table 3.

The results of the experiment on Highway (Clip1) show that there are lots of errors based on the GMM method and the general ViBe method, because there are some leaves shaking in the background having similar color attribute with the vehicles. However, the proposed approach can deal with this problem efficiently, which is
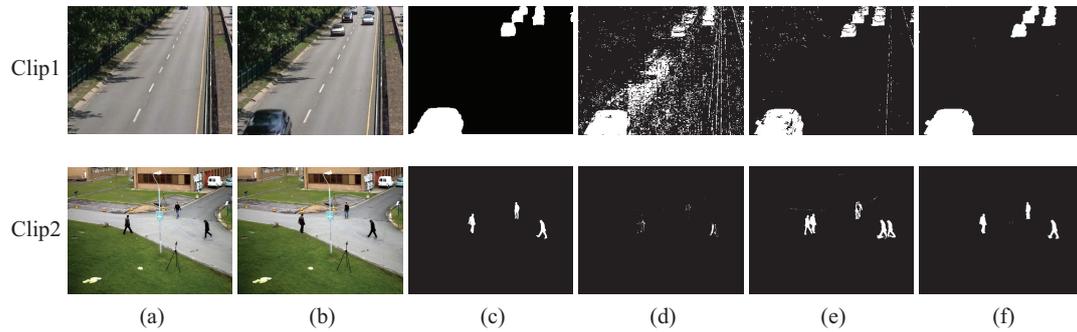
Figure 5. The moving object detection experiments on the video Highway (Clip1) and People (Clip2): (a) the first frame; (b) the frame for detection; (c) the ground-truth; (d) the result of GMM; (e) the result of G-ViBe; and (f) the result of I-ViBe.

Table 3. The valuation of the three methods for moving objects detection on Highway and People

The valuation indices	The video clip1			The video clip2		
	GMM [32]	G-ViBe [35]	I-ViBe	GMM [32]	G-ViBe [35]	I-ViBe
SP	0.8765	0.9910	**0.9979**	**0.9998**	0.9915	0.9992
RE	0.7843	0.8554	**0.9674**	0.1545	0.8616	**0.9849**
FPR	0.1234	0.0089	**0.0020**	**0.0001**	0.0084	0.0007
FNR	0.2156	0.1445	**0.0325**	0.8454	0.1383	**0.0150**
PWC	0.1307	0.0196	**0.0041**	0.0059	0.0097	**0.0008**
PRE	0.3532	0.8918	**0.9713**	0.8834	0.5110	**0.8943**
F	0.4871	0.8732	**0.9694**	0.2630	0.6415	**0.9374**
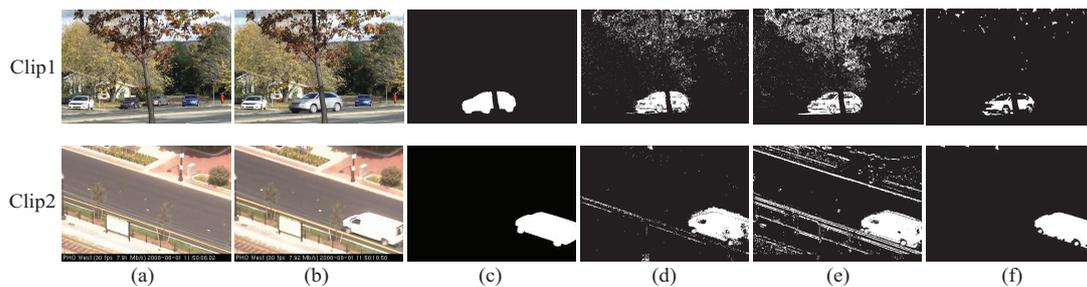
Figure 6. The moving object detection experiments under challenging conditions: (a) the first frame; (b) the frame for detection; (c) the ground-truth; (d) the result of GMM; (e) the result of G-ViBe; and (f) the result of I-ViBe.

combined with the edge information (see Figure 5 and Table 3). Furthermore, there are also ghost problems in the detection results of the experiment on People (Clip2) based on the G-ViBe, because the first frame includes the moving objects (see Figure 5e).

4.3. The experiment under challenging conditions

To further test the performance of the proposed method for moving object detection under some challenging conditions, two extensive experiments were conducted in the dataset of Fall (Clip1) and Boulevard (Clip2), respectively. In the Fall dataset, the background is changing obviously because of the leaves shaking violently. In the Boulevard dataset, the video is blurry due to the shake of the camera. The results of these experiments are shown in Figure 6 and Table 4.

The results in the two experiments show that the performances of all the three methods decrease under dy-

Table 4. The valuation of the three methods for moving object detection under challenging conditions

The valuation indices	The video Clip1			The video Clip2		
	GMM [32]	G-ViBe [35]	I-ViBe	GMM [32]	G-ViBe [35]	I-ViBe
SP	0.9547	0.8765	**0.9960**	0.9850	0.9025	**0.9974**
RE	**0.8755**	0.7496	0.7184	0.8643	0.9258	**0.9574**
FPR	0.0452	0.1234	**0.0039**	0.0149	0.0974	**0.0025**
FNR	**0.1244**	0.2503	0.2815	0.1356	0.0741	**0.0425**
PWC	0.0481	0.1285	**0.0106**	0.0219	0.0957	**0.0051**
PRE	0.4245	0.2016	**0.8202**	0.7819	0.4173	**0.9617**
F	0.5718	0.3178	**0.7659**	0.8210	0.5753	**0.9595**

(a) (b) (c) (d) (e) (f)
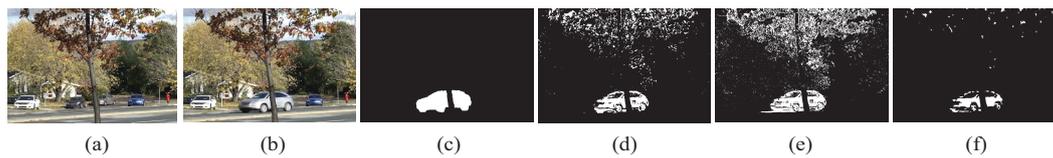
Figure 7. The moving object detection experiments on the video Fall: (a) the first frame; (b) the frame for detection; (c) the ground-truth; d) the result of G-ViBe; (e) the result of F-ViBe; and (f) the result of I-ViBe.

namic environments. The main reason is that all the three methods are based on the mechanism of background subtraction. However, the performance of the proposed approach does not decrease dramatically compared with other two methods (see the values of PRE and F in Table 4). This performance of the proposed approach is very important for the real application of moving object detection.

5. DISCUSSION

The results presented in Section 3 show that the proposed approach can deal with the ghost area problem and remove the shadow very well. In addition, the evaluation indices of the proposed approach are better than the GMM method and the general ViBe method. In this section, some performances of the proposed approach are discussed, including the computation complexity and the background updating mechanism.

One key part of the ViBe-based approach is the background updating mechanism, so the performance of the improvement in this part for the proposed method is discussed first. An experiment was conducted in the dataset of Fall, where the proposed approach was compared with two methods. The first one is the general ViBe. The second one is a method which has the same parameters and work flow as the proposed approach, except that the background updating mechanism is based on the fixed detection radius and updating rate, and this method is called F-ViBe. The experimental results of Section 3.3 are used as reference, as shown in Figure 7 and Table 5. The experimental results show that the proposed approach can deal with the dynamic environment better than the other two methods. Thus, the background updating mechanism is very efficient for moving object detection under complex environment. In addition, the detection radius and updating rate of the F-ViBe method are given by the designer, which need more experience and time.

Another important index of the moving object detection method is the real time problem, because the speed of the moving object is very high sometimes. The proposed approach has two main differences with the general ViBe method, the background modeling and updating mechanism and the shadow removal strategy. Thus, the time needed in all the three experiments of Clip1 in Section 3 is divided into two parts, the time for the background modeling and the time for moving object detection (including background updating).

Table 5. The moving object detection experiments based on different background discrimination mechanism

The valuation indices	G-ViBe [35]	F-ViBe	I-ViBe
SP	0.8765	0.9738	**0.9960**
RE	**0.7496**	0.7394	0.7184
FPR	0.1234	0.0261	**0.0039**
FNR	**0.2503**	0.2605	0.2815
PWC	0.1285	0.0328	**0.0106**
PRE	0.2016	0.4551	**0.8202**
F	0.3178	0.5634	**0.7659**

Table 6. The moving object detection experiments based on different background discrimination mechanism

The video	Computation time (s)	GMM [32]	G-ViBe [35]	I-ViBe
Clip1 of Section 4.1	background modeling		**0.4042**	1.9662
(180 * 144)	object detection	**0.0337**	0.1028	0.1508
Clip1 of Section 4.2	background modeling		**0.4077**	2.0265
(320 * 240)	object detection	**0.0353**	0.1049	0.2503
Clip1 of Section 4.3	background modeling		**0.4966**	2.2822
(720 * 480)	object detection	**0.0798**	0.1229	0.4138

The results in Table 6 show that more time for the object detection is needed using G-ViBe and I-ViBe than the GMM method, because the GMM method selects the initial background frame randomly. For high resolution videos, the proposed ViBe method takes more time to compare the values of pixels in each channel of the HSV space, so the time for object detection increases. In addition, the results show that more time is needed in the ViBe based approach during the background modeling process, which can be off-line proceeded and will not affect the real-time moving object detection. For off-line processing, multiple images of the detection area can be collected in advance, and the mode background method can be used for modeling. In the subsequent detection tasks, there is no need to repeat the modeling. Thus, the proposed approach has a better comprehensive performance than both the GMM method and the G-ViBe method, although the computation time of the proposed approach is relatively higher than the other two methods, which is a problem for further study.

6. CONCLUSIONS

In this paper, we present an improved moving object detection approach based on ViBe algorithm. During the process of foreground region extraction, the initial background is obtained by the previous few frames and then updated by the value of the eight neighboring pixel difference between the background and the current frame. In addition, a shadow removal strategy is adopted by combining the HSV color space and the edge information. Most of the parameters in the proposed method are calculated adaptively, which is very important for the adaptivity of moving object detection method. The experiments showed that the proposed approach can deal with moving object detection efficiently in various situations, such as the severe shadow problems in the foreground and the presence of moving objects in the first frame. In addition, the proposed approach can be used for real-time moving object detection. In future work, some more efficient methods based on artificial intelligence algorithms should be studied to improve the accuracy and real-time ability for moving object detection.

DECLARATIONS

Authors' contributions
Funding acquisition: Ni J
Project administration: Ni J, Shi P
Writing-original draft: Tang G
Writing-review and editing: Li Y, Zhu J

Availability of data and materials
Not applicable.

Financial support and sponsorship
This work has been supported by the National Natural Science Foundation of China (61873086) and the Science and Technology Support Program of Changzhou (CE20215022).

Conflicts of interest
All authors declared that they have no conflicts of interest to this work.

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Copyright
© The Author(s) 2022.

REFERENCES

1. Ni J, Zhang X, Shi P, Zhu J. An Improved kernelized correlation filter based visual tracking method. *Mathematical Problems in Engineering* 2018;2018:1-12. [DOI: 10.1155/2018/6931020]
2. Zhang Z, Chen B, Yang M. Moving target detection based on time reversal in a multipath environment. *IEEE Trans Aerosp Electron Syst* 2021;57:3221-36. [DOI: 10.1109/TAES.2021.3074131]
3. Ni J, Yang L, Wu L, Fan X. An improved spinal neural system-based approach for heterogeneous AUVs cooperative hunting. *Int J Fuzzy Syst* 2018;20:672-86. [DOI: 10.1007/s40815-017-0395-x]
4. Tivive FHC, Bouzerdoum A. Toward moving target detection in through-the-wall radar imaging. *IEEE Trans Geosci Remote Sensing* 2021;59:2028-40. [DOI: 10.1109/TGRS.2020.3005199]
5. Ni J, Gong T, Gu Y, Zhu J, Fan X. An improved deep residual network-based semantic simultaneous localization and mapping method for monocular vision robot. *Comput Intell Neurosci* 2020;2020:7490840. [DOI: 10.1155/2020/7490840]
6. Lu X, Mao X, Liu H, Meng X, Rai L. Event camera point cloud feature analysis and shadow removal for road traffic sensing. *IEEE Sensors J* 2022;22:3358-69. [DOI: 10.1109/JSEN.2021.3138736]
7. Negri P. Estimating the queue length at street intersections by using a movement feature space approach. *IET Image Processing* 2014;8:406-16. [DOI: 10.1049/iet-ipr.2013.0496]
8. Sengar SS, Mukhopadhyay S. Motion detection using block based bi-directional optical flow method. *Journal of Visual Communication and Image Representation* 2017;49:89-103. [DOI: 10.1016/j.jvcir.2017.08.007]
9. Chen Y, Wang J, Zhu B, Tang M, Lu H. Pixelwise deep sequence learning for moving object detection. *IEEE Trans Circuits Syst Video Technol* 2019;29:2567-79. [DOI: 10.1109/TCSVT.2017.2770319]
10. Li J, Pan ZM, Zhang ZH, Zhang H. Dynamic ARMA-based background subtraction for moving objects detection. *IEEE Access* 2019;7:128659-68. [DOI: 10.1109/ACCESS.2019.2939672]
11. Garcia-Garcia B, Bouwmans T, Silva AJR. Background subtraction in real applications: challenges, current models and future directions. *Computer Science Review* 2020;35:100204. [DOI: 10.1016/j.cosrev.2019.100204]
12. Cristani M, Farenzena M, Bloisi D, Murino V. Background subtraction for automated multisensor surveillance: a comprehensive review. *EURASIP J Adv Signal Process* 2010;2010. [DOI: 10.1155/2010/343057]
13. Javed S, Narayanamurthy P, Bouwmans T, Vaswani N. Robust PCA and robust subspace tracking: a comparative evaluation. In: 2018 IEEE Statistical Signal Processing Workshop (SSP). Freiburg im Breisgau, Germany; 2018. pp. 836–40. [DOI: 10.1109/SSP.2018.8450718]
14. Mandal M, Vipparthi SK. An Empirical Review of Deep Learning Frameworks for Change Detection: Model Design, Experimental Frameworks, Challenges and Research Needs. IEEE Transactions on Intelligent Transportation Systems 2021:Article in Press. [DOI:

10.1109/TITS.2021.3077883]

15. Minematsu T, Shimada A, Uchiyama H, Taniguchi R. Analytics of deep neural network-based background subtraction. *J Imaging* 2018;4:78. [DOI: 10.3390/jimaging4060078]

16. Giraldo JH, Javed S, Sultana M, Jung SK, Bouwmans T. The emerging field of graph signal processing for moving object segmentation. In: International Workshop on Frontiers of Computer Vision. Virtual, Online; 2021. pp. 31–45. [DOI: 10.1007/978-3-030-81638-4-3]

17. Giraldo JH, Javed S, Bouwmans T. Graph Moving Object Segmentation. *IEEE Trans Pattern Anal Mach Intell* 2022;44:2485-503. [DOI: 10.1109/TPAMI.2020.3042093]

18. Sengar SS, Mukhopadhyay S. Moving object area detection using normalized self adaptive optical flow. *Optik* 2016;127:6258-67. [DOI: 10.1016/j.ijleo.2016.03.061]

19. Ni J, Chen Y, Chen Y, et al. A survey on theories and applications for self-driving cars based on deep learning methods. *Applied Sciences* 2020;10:2749. [DOI: 10.3390/app10082749]

20. Wang Y, Zhu L, Yu Z. Foreground detection for infrared videos with multiscale 3-D fully convolutional network. *IEEE Geosci Remote Sensing Lett* 2019;16:712-6. [DOI: 10.1109/LGRS.2018.2881053]

21. Ni J, Shen K, Chen Y, Cao W, Yang SX. An improved deep network-based scene classification method for self-driving cars. *IEEE Trans Instrum Meas* 2022;71:1-14. [DOI: 10.1109/TIM.2022.3146923]

22. Mahmoudabadi H, Olsen MJ, Todorovic S. Detecting sudden moving objects in a series of digital images with different exposure times. *Computer Vision and Image Understanding* 2017;158:17-30. [DOI: 10.1016/j.cviu.2017.01.004]

23. Wang C, Cheng J, Chi W, Yan T, Meng MQH. Semantic-aware informative path planning for efficient object search using mobile robot. *IEEE Trans Syst Man Cybern, Syst* 2021;51:5230-43. [DOI: 10.1109/TSMC.2019.2946646]

24. Liu Z, An D, Huang X. Moving target shadow detection and global background reconstruction for VideoSAR based on single-frame imagery. *IEEE Access* 2019;7:42418-25. [DOI: 10.1109/ACCESS.2019.2907146]

25. Talab AMA, Huang Z, Xi F, Haiming L. Moving crack detection based on improved VIBE and multiple filtering in image processing techniques. *IJSIP* 2015;8:275-86. [DOI: 10.14257/ijsip.2015.8.2.27]

26. Gao Y, Cheng P. Full-scale video-based detection of smoke from forest fires combining ViBe and MSER algorithms. *Fire Technol* 2021;57:1637-66. [DOI: 10.1007/s10694-020-01052-3]

27. Huang W, Liu L, Yue C, Li H. The moving target detection algorithm based on the improved visual background extraction. *Infrared Physics & Technology* 2015;71:518-25. [DOI: 10.1016/j.infrared.2015.06.011]

28. Qiu S, Tang Y, Du Y, Yang S. The infrared moving target extraction and fast video reconstruction algorithm. *Infrared Physics & Technology* 2019;97:85-92. [DOI: 10.1016/j.infrared.2018.11.025]

29. Yue Y, Xu D, Qian Z, Shi H, Zhang H. AntViBe: improved vibe algorithm based on ant colony clustering under dynamic background. *Mathematical Problems in Engineering* 2020;2020:1-13. [DOI: 10.1155/2020/7478626]

30. Nagarathinam K, Kathavarayan RS. Moving shadow detection based on stationary wavelet transform and zernike moments. *IET Computer Vision* 2018;12:787-95. [DOI: 10.1049/iet-cvi.2017.0273]

31. Khare M, Srivastava RK, Khare A. Moving shadow detection and removal-a wavelet transform based approach. *IET Computer Vision* 2014;8:701-17. [DOI: 10.1049/iet-cvi.2014.0028]

32. Stauffer C, Grimson WEL. Adaptive background mixture models for real-time tracking. In: Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'99). vol. 2. Fort Collins, CO, USA; 1999. pp. 246–52. [DOI: 10.1109/CVPR.1999.784637]

33. KaewTraKulPong P, Bowden R. In: An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection. Boston, MA: Springer US; 2002. pp. 135–44. [DOI: 10.1007/978-1-4615-0913-4-11]

34. Hofmann M, Tiefenbacher P, Rigoll G. Background segmentation with feedback: The Pixel-Based Adaptive Segmenter. In: 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Providence, RI, United states; 2012. pp. 38–43. [DOI: 10.1109/CVPRW.2012.6238925]

35. Barnich O, Van Droogenbroeck M. ViBE: a powerful random technique to estimate the background in video sequences. In: 2009 IEEE International Conference on Acoustics, Speech and Signal Processing. Taipei, Taiwan; 2009. pp. 945–48. [DOI: 10.1109/ICASSP.2009.4959741]

36. Barnich O, Van Droogenbroeck M. ViBe: a universal background subtraction algorithm for video sequences. *IEEE Trans Image Process* 2011;20:1709-24. [DOI: 10.1109/TIP.2010.2101613]

37. Zhu F, Jiang P, Wang Z. ViBeExt: The extension of the universal background subtraction algorithm for distributed smart camera. In: 2012 International Symposium on Instrumentation Measurement, Sensor Network and Automation (IMSNA). vol. 1. Sanya, Hainan, China; 2012. pp. 164–68. [DOI: 10.1109/MSNA.2012.6324539]

38. Chen F, Zhu B, Jing W, Yuan L. Removal shadow with background subtraction model ViBe algorithm. In: 2013 2nd International Symposium on Instrumentation and Measurement, Sensor Network and Automation (IMSNA). Toronto, ON, Canada; 2013. pp. 264–69. [DOI: 10.1109/IMSNA.2013.6743265]

39. Yang Y, Han D, Ding J, Yang Y. An improved ViBe for video moving object detection based on evidential reasoning. In: 2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI). Baden-Baden, Germany: IEEE; 2016. pp. 26–31. [DOI: 10.1109/MFI.2016.7849462]

40. Liu L, Chai Gh, Qu Z. Moving target detection based on improved ghost suppression and adaptive visual background extraction. *J Cent South Univ* 2021;28:747-59. [DOI: 10.1007/s11771-021-4642-9]

41. Bo G, Kefeng S, Daoyin Q, Hongtao Z. Moving object detection based on improved ViBe algorithm. *IJSEIA* 2015;9:225-32. [DOI:

10.14257/ijsh.2015.9.12.23]

42. Zhang X, Liu K, Wang X, Yu C, Zhang T. Moving Shadow Removal Algorithm Based on HSV Color Space. *TELKOMNIKA* 2014;12. [DOI: 10.11591/telkomnika.v12i4.4991]

43. Zhang B, Jiao D, Lv X. A target detection algorithm for SAR images based on regional probability statistics and saliency analysis. *International Journal of Remote Sensing* 2019;40:1394-410. [DOI: 10.1080/01431161.2018.1524593]

44. Tian Y, Wang D, Jia P, Liu J. Moving Object Detection with ViBe and Texture Feature. In: Pacific Rim Conference on Multimedia. Xi'an, China; 2016. pp. 150–59. [DOI: 10.1007/978-3-319-48890-5_15]

45. Liu Z, Yin H, Mi Y, Pu M, Wang S. Shadow Removal by a Lightness-Guided Network With Training on Unpaired Data. *IEEE Trans Image Process* 2021;30:1853-65. [DOI: 10.1109/TIP.2020.3048677]

46. Hu X, Fu CW, Zhu L, Qin J, Heng PA. Direction-aware spatial context features for shadow detection and removal. *IEEE Trans Pattern Anal Mach Intell* 2020;42:2795-808. [DOI: 10.1109/TPAMI.2019.2919616]

47. Huang W, Kim K, Yang Y, Kim YS. Automatic Shadow Removal by Illuminance in HSV Color Space. *csit* 2015;3:70-5. [DOI: 10.13189/csit.2015.030303]

48. Long Z, Zhou X, Zhang X, Wang R, Wu X. Recognition and classification of wire bonding joint via image feature and SVM model. *IEEE Trans Compon, Packag Manufact Technol* 2019;9:998-1006. [DOI: 10.1109/TCPMT.2019.2904282]

49. Goyette N, Jodoin PM, Porikli F, Konrad J, Ishwar P. Changedetection.net: A new change detection benchmark dataset. In: 2012 IEEE computer society conference on computer vision and pattern recognition workshops. Providence, RI, USA; 2012. pp. 1–8. [DOI: 10.1109/CVPRW.2012.6238919]

50. Zhang H, Qian Y, Wang Y, Chen R, Tian C. A ViBe Based Moving Targets Edge Detection Algorithm and Its Parallel Implementation. *Int J Parallel Prog* 2020;48:890-908. [DOI: 10.1007/s10766-019-00628-z]

51. Zhang E, Li Y, Duan J. Moving object detection based on confidence factor and CSLBP features. *The Imaging Science Journal* 2016;64:253-61. [DOI: 10.1080/13682199.2016.1168977]