Research on Improved Adaptive ViBe Algorithm For Vehicle Detection

Kun Jiang

School of Computer Science and Engineering

Xi'an Technological University

Xi'an, China

E-mail: 416578078@qq.com

Jianguo Wang

School of Computer Science and Engineering

Xi'an Technological University

Xi'an, China

E-mail: wjg_xit@126.com

Abstract—**Vehicle detection is an important step in vehicle tracking and recognition in video environment. Vibe algorithm is a moving target detection algorithm based on background difference method. Based on the traditional ViBe algorithm, this paper introduces three-frame difference method combined with ViBe algorithm to speed up the elimination of ghosts, and proposes an adaptive Vibe algorithm, which defines two kinds of vehicle detection errors and their corresponding error functions. Then, according to the range of these two errors, a set of reasonable judgment methods are determined to adjust the unreasonable threshold, which ensures the adaptive updating of the background model. It improves the environmental adaptability of vehicle detection and ensures higher accuracy of vehicle detection under complex illumination conditions.**

Keywords-Vehicle Detection; Background Difference Method; Vibe Algorithm; Three-Frame Difference Algorithm

I.   INTRODUCTION

Vehicle detection is the key step of video vehicle recognition, which aims to obtain the location of vehicle for further recognition. For each pixel, its background can usually be built using a model. At present, there are three methods for moving object detection: optical flow method[1], frame difference method[2], background subtraction method[3]. Based on the motion vector of pixels, the optical flow method can detect and track the target, but it has a large amount of computation and poor real-time performance. Moreover, the method lacks sensitivity to noise, illumination change and background interference. Frame difference method detects moving objects according to the difference between two or three consecutive frames. It has strong adaptability to the background change, but it does not perform well in detecting the contour of moving objects. In addition, it is very sensitive to the speed of moving objects, so it cannot effectively detect slow moving objects; background difference method is a commonly used moving object detection algorithm. The main idea is to make a distinction between each frame and background model to build the background model and get the moving foreground objects. Background difference method has the ability to adapt to the scene changes in the video background, but if the background model contains foreground objects, it may generate ghosting.

Background difference method is one of the most widely used vehicle detection methods because of its fast and accurate. Traditional background subtraction algorithms include Gaussian mixture model (GMM)[4] and codebook[5]. GMM method is simple and low cost. However, the initialization time is too long and the algorithm is complex to meet the real-time requirements. Cookbook has the advantages of dealing with time fluctuations well, but its memory consumption is quite high. In 2011, Olivier barnechand

Marc van droogenbroeck proposed the vibe algorithm [6]. Because the algorithm only needs the first frame to complete the model initialization, it can meet the real-time requirements compared with GMM model. In addition, in the process of execution, the algorithm only needs to record the corresponding sample set for each pixel, so it has smaller memory consumption. In addition, vibe algorithm has goodanti noise ability. Although the foreground object can be mixed in pixel initialization, it will produce ghost phenomenon.

In this paper, an improved adaptive algorithm of vibe is proposed, and a moving target detection method based on three frame difference method is introduced. Through experimental verification, the algorithm proposed in this paper effectively solves the problems of "ghost" existing in traditional VIB algorithm and insufficient adaptability to complex light environment. It has the advantages of simple algorithm, good real-time performance and high detection accuracy.

II. VIBE BACKGROUND MODELING

The vibe background modeling algorithm was proposed by Olivier barnech et al in 2011.[7] can be used for fast background extraction and moving object detection. Vibe algorithm uses two mechanisms of random selection and neighborhood propagation to build and update the background model, which includes three steps: background modeling and initialization, foreground detection and background model update. In this paper, based on the traditional vibe algorithm, an adaptive background model of vibe is added. According to the range of vehicle detection error, a set of judgment method is determined to evaluate the rationality of the current threshold. When the threshold is not reasonable, adjust according to a certain step to ensure that the background model is updated automatically, and finally get more accurate vehicle detection results.

Initialization of background model

v_x represents the pixel value at point x, and each pixel builds the number of background sample sets N ($N = 20$):

$$M(x) = \{v_1, v_2, \ldots, v_N\} \qquad (1)$$

As shown in Fig. 1, the gray space makes the region centered on X, the radius of the gray space $S_R(v(x))$ is R, and the threshold $\#\min$ is set($\#\min = 2$). Then find the intersection of $M(x)$ and $S_R(v(x))$, C is the total number of elements in the intersection:

$$C = \# M(x) \left[\{ S_R(v(x)) \cap \{v_1, v_2 \ldots, v_n\} \} \right] \qquad (2)$$
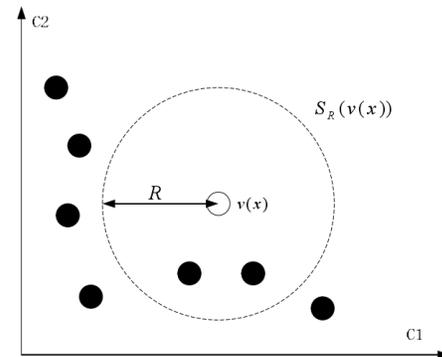
Figure 1. ViBe background model.

The initialization of vibe model only needs the first image, but only one image does not contain the spatial information of pixels. According to the similar spatial characteristics of the close pixels, the sample set is filled with the approximate pixels. When the first image is input, the background model of the pixels in the image is as follows (3):

$$M_0(x) = v_0\{y \mid y \in N_G(x)\} \qquad (3)$$

Where, $N_G(x)$ represents the neighborhood pixel; v_0 is the currently selected pixel. The probability that the pixels in $N_G(x)$ are selected in N initialization is $(1-N)/N$.

A. Foreground detection

After initialization, vehicle detection starts from the second image. Separating foreground target and background is the process of moving target detection. At time t, the pixel value of random pixel x is v_t, According to formula (2) C is judged according to formula (4):

$$v_t = \begin{cases} C < \#\min & (foreground) \\ C \geq \#\min & (background) \end{cases} \qquad (4)$$

T ($T = 20$) is a preset threshold. When the number of times for the background is greater than the threshold $\#\min$, the pixel x are considered to belong to the background area; if not, it belongs to the foreground target area. According to formula (4), the binary image obtained after vehicle detection is the initial vehicle detection binary image.

B. Dynamic update of background model

In the process of updating the vibe background model, not only the relationship between the current pixel and its historical samples, but also the relationship between the current pixel and other pixels in other spatial neighborhood should be considered. In other words, the updating of vibe background model is a random process in both time and space.

In the background model in the previous frame, if the current pixel v_t is marked as background, its background model $M_t(x)$ is updated in time. If the current pixel is marked as vehicle, the model is not updated. This update strategy is called conservative update strategy. The method of updating the background model is to randomly select the sample m in the sample set $M_t(x)$.

The method to update the background model in space is to calculate the gradient amplitude of the current pixel v_t, If the gradient amplitude is greater than 50, the space update is not implemented. Otherwise, in the 8 neighborhood of the current pixel v_t, randomly select the pixel marked as v_t, in the background model $M_t'(x)$ of pixel v_t' randomly select a sample m_j, and the characteristic value of the current pixel v_t is assigned to m_j. If the current pixel v_t is at the edge of the image, it is randomly selected in its incomplete 8 neighborhood. The spatial update strategy ensures the continuity of spatial information in the background image.

III. ADAPTIVE VIBE BACKGROUND MODEL

In the vibe background model, threshold R represents the range of background eigenvalues (as shown in Figure 1). Threshold R has a great influence on vehicle detection results. If the fixed threshold value is greater than the expected threshold value, it should be the vehicle's area, which will lead to inaccurate detection of the vehicle area.

A. Error functions of vehicle

There are several situations of vehicle detection error: the detection background area is mistakenly

regarded as the vehicle area, or the vehicle area is mistakenly regarded as the background area. The former belongs to the error vehicle area; the latter is the error background area. The size of these two types of error areas will change with the change of threshold R.

When the threshold R is very small, the fluctuation range of the sample eigenvalues in the vibe background model is also very small, which helps to improve the detection accuracy of the vehicle area. The noise area may be mistakenly regarded as the wrong vehicle area, and the small noise area can be removed by morphological method, but the large noise area will not be easily removed. Therefore, in order to prevent this situation, it is necessary to minimize the wrong vehicle area.

In theory, the wrong vehicle area is part of the background and it is static. Therefore, if the connected region does not overlap all the moving regions in the binary mapping of frame difference, the connected region is considered as the wrong vehicle region, and the error function of the wrong vehicle region can be defined as:

$$Err_1(R) = \frac{\sum_{i=1}^{n_1} a_1^i}{L \times W} \tag{5}$$

Where, a_1^i is the ith connected domain, which does not overlap all moving regions in the binary mapping of frame difference, and n_1 is the number of such connected domains. Is the total area of the wrong vehicle area, L and W represent the length and width of the image respectively, and their units are pixels. The higher the resolution of the image, the more accurate the value of $\sum_{i=1}^{n_1} a_1^i$. The wrong vehicle area function is defined as the ratio of the total area of the wrong vehicle area to the total area of the image.

When the threshold value R is large, the fluctuation range of the sample eigenvalues in the vibe background model is also large, which is conducive to improving the detection accuracy of the background region. The second type of error area is to detect the area originally belonging to the vehicle as the background area [9].

Firstly, error background area error a_2^i is defined, which represents the difference between the area of the smallest external rectangle containing the ith vehicle and the area of the same vehicle detected. s^i is the area of the smallest external rectangle containing the first vehicle, and v^i is the area of the ith vehicle detected.

$$a_2^i = s^i - v^i \tag{6}$$

Therefore, error background area error function can be further defined as:

$$Err_2(R) = \frac{\sum_{i=1}^{n_2} a_2^i}{n_2} \tag{7}$$

Where n_2 is the number of vehicles. $\sum_{i=1}^{n_1} a_2^i$ is the total area of the wrong vehicle area, take the average value to the error $Err_2(R)$ of the wrong background area.

From the above, the error $Err_1(R)$ of the error background area and the error $Err_2(R)$ of the error vehicle area are obtained. The total error of the error area can be defined as:

$$Err(R) = Err_1(R) + Err_2(R) \tag{8}$$

B. Adaptive adjustment of threshold

If the current threshold R is too small, the area of the area originally belonging to the background and mistakenly detected as the vehicle is too large, which means that the area $Err_1(R)$ of the wrong vehicle area is relatively large. If the current threshold is too large, the area $Err_2(R)$ of the area originally belonging to the vehicle and mistakenly detected as the background is too small, which means that the area of the wrong background is relatively large. According to this situation, we use the following adaptive scheme[10]:

$$\begin{cases} R = R - N, \ if \ Err(R) > T_2 \\ R = R + N, \ else \ if \ Err(R) > T_1 \ and \ Err(R) \le T_2 \\ R = R, \qquad else \end{cases} \quad (9)$$

T_1 and T_2 is the parameter to judge the rationality of threshold, and N is the adjustment step of threshold R. After a lot of experiments, the range of T_1 is 0.02 to 0.04, the range of T_2 is 0.13 to 0.26, the range of N is 1 to 3. A large number of experiments show that these values can ensure that the total error of the error background area and the error vehicle area proposed in this paper can be minimized.

IV.   IMPROVED ADAPTIVE VIBE ALGORITHM OF THREE FRAME DIFFERENCE METHOD

This paper presents an improved adaptive vibe background modeling algorithm, which uses the vibe algorithm to model the background, adjusts the threshold adaptively, and then introduces the three frame difference method to improve. Vibe background model algorithm is based on the first image to establish the background model[10], but the traditional vibe algorithm will appear the phenomenon of "ghost". At present, many domestic and foreign literatures have carried out relevant research on the problem of "ghost".

At present, the more commonly used method is to combine the traditional vibe algorithm with other algorithms, or to change the initialization of the first image of the original algorithm to a multi frame image Initialization.

The traditional vibe algorithm needs hundreds of frames to completely eliminate the "ghost" in the first frame. Using the improved adaptive vibe background modeling algorithm, the speed of eliminating the "ghost" is accelerated, and the "ghost" can be eliminated within dozens of frames. At the same time, the proposed adaptive vibe algorithm greatly improves the traditional vibe background modeling algorithm for complex light environment detection. After the introduction of three frame difference method, the speed of eliminating "ghost" is obviously speeded up, and the problem of "hole" existing in the three frame difference method itself is solved. Finally, the accuracy of moving object detection is improved by morphological processing of detection results. The flow chart of the improved vibe algorithm based on the three frame difference method is shown in Figure 2, and the specific implementation steps are as follows:

1) Input video image, and carry out image pre-processing such as graying and binarization.

2) Background modeling of three frame difference method and background modeling of vibe algorithm are carried out respectively for the image preprocessed. The final image is the "and" of the image calculated by the two methods.

3) Background modeling of three frame difference method and background modeling of vibe algorithm are carried out respectively for the image preprocessed. The final image is the "and" of the image calculated by the two methods.

4) Through the adaptive threshold adjustment algorithm proposed in this paper, the appropriate threshold is calculated to update the current background.

5) The updated image is processed by morphology to get the final detection results.
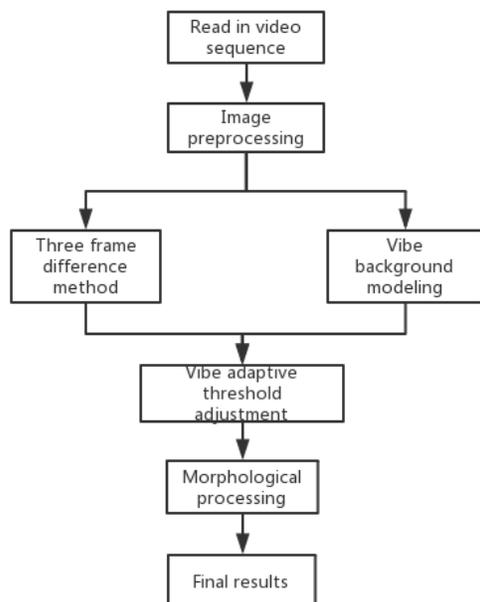
Figure 2.   Improved ViBe algorithm of three frame difference method.

V. EXPERIMENTAL RESULT

Based on the above theory and processing flow, the algorithm is tested in the following environment: operating system: Microsoft Windows 10, experimental platform: Visual Studio 2019, CPU: intel5, RAM: 8g, third-party open source library: openCV 4.1.0. In order to verify that the method proposed in this paper can accurately detect moving objects in complex environment, the video selected in this experiment is the road monitoring video with more vehicles. This video is the traffic situation of a certain intersection at a certain time, with 650 frames in total, and the frame size is 640*480. An improved three frame difference algorithm based on vibe background modeling is used to detect moving vehicles. During the experiment, the 20th frame, 122 frame and 380 frame of video image sequence are randomly selected to analyze the detection effect.

20th frames122th frames380th frames

(a) Original video

(b) Original ViBe algorithm

(c) Gaussian mixture model

(d) Three frame difference method

(e) Improved ViBe algorithm
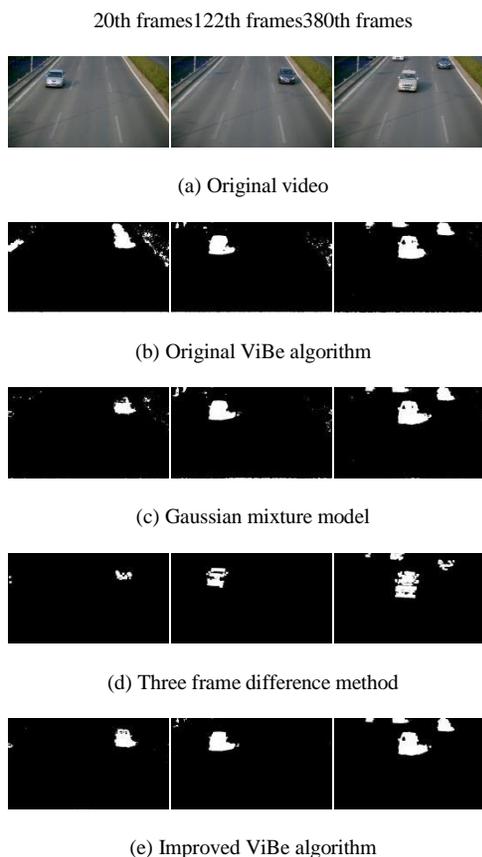
Figure 3.   Comparison between the algorithm in this paperand the traditional target detection algorithm.

It can be seen from Figure 3 that the original vibe algorithm of frame 20 has obvious ghost phenomenon and interference of complex lighting environment factors, and the effect of Gaussian mixture model is good, but the calculation of Gaussian mixture model is too complex to meet the real-time requirements of vehicle detection, the real-time performance of three frame difference method is good but the accuracy is not high enough, and there is an obvious "empty" phenomenon. In contrast, this algorithm effectively solves the ghost phenomenon and reduces the impact of complex environmental factors.

TABLE I.   EVALUATION RESULTS OF VEHICLE INSPECTION

	GMM	Three frame difference	Original vibe	Improved ViBe
Recall	72.1%	70.0%	80.5%	87.1%
Precision	91.2%	83.1%	90.2%	92.3%
F1	80.1%	80.0%	85.1%	90.0%

VI.   CONCLUSION

In order to solve the ghost phenomenon in vibe algorithm and the problem of low detection accuracy in complex illumination environment, this paper first proposes an adaptive threshold vibe algorithm in order to update the background accurately under the condition of complex illumination change. By defining two kinds of vehicle detection error functions, according to the error range calculated by these two functions, an algorithm is used to determine the rationality of threshold. Vehicle detection and background update are performed by using the adaptive algorithm.

In order to solve the problem of ghost, three frame difference method and adaptive vibe algorithm are introduced. Finally, the experimental results show that the improved adaptive vibe algorithm can effectively remove the "ghost" phenomenon and improve the accuracy of vehicle detection in complex lighting environment. Figure 4 shows the change of the number of ghost pixels with the number of frames. The X axis represents the number of frames, and the Y axis represents the number of ghost pixels. It can be seen from the figure that the improved algorithm in this paper greatly accelerates the ghost elimination speed, which is significantly faster than the Gaussian mixture algorithm and the vibe algorithm.
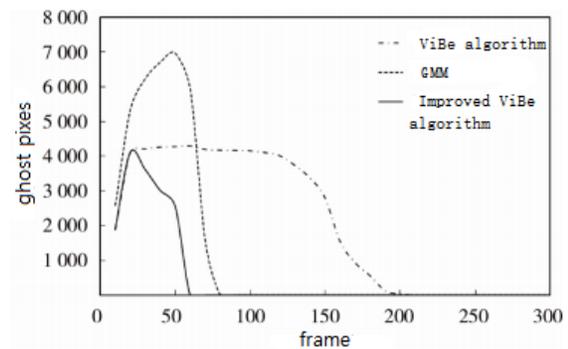
Figure 4.   Ghost elimination speed.

REFERENCES

[1] Delpiano J, Jara J, Scheer J, et al. Performance of optical flow techniques for motion analysis of fluorescent point signals in confocal microscopy. Machine Vision & Applications, 2012, 23(4):675-689.

[2] J.-G. Yan, W.-H Xu. Moving object real-time detection algorithm based on new frame difference. Computer Engineering & Design, 2013, 34(12):4331-4335.

[3] Yang W, Zhang T. A new method for the detection of moving targets in complex scenes. Journal of Computer Research & Development, 1998.

[4] Kaewtrakulpong P, Bowden R. An improved adaptive background mixture model for realtime tracking with shadow detection. Springer US, 2002.

[5] Kim K, Chalidabhongse T H, Harwood D, et al. Real-time foreground– background segmentation using codebook model. Real-Time Imaging, 2005, 11(3):172-185.

[6] Barnich O, Van D M. ViBe: a universal background subtraction algorithm for video sequences. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2011, 20(6):1709-1724.

[7] Barnich O, Van Droogenbroeck M. ViBe: A unrsal background subtraction algorithm for video sequences[J]. IEEE Transactions on Image Processing, 2011, 20(6):1709-1724.

[8] Hu Changhui, Lu Xiaobo, Ye Mengjun, Zeng Weili. Singular value decomposition and local near neighbors for face recognition under varying illumination [J]. Pattern Recognition, 2017,64: 60-83.

[9] Z. Qiming and M. ChengQian, A vehicle detection method in tunnel video based on ViBe algorithm,2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, 2017, pp. 1545-1548.

[10] C. Pan, Z. Zhu, L. Jiang, M. Wang and X. Lu, "Adaptive ViBe background model for vehicle detection," 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, 2017, pp. 1301-1305.

[11] Ekpar F. A framework for intelligent video surveillance. Proceedings of the IEEE 8th International Conference on Computer and Information Technology Workshops. Sydney, QLD, Australia. 2008. 421–426.