Moving Vehicle Detection with Shadow Elimination Based on Improved ViBe Algorithm

Xue Xia¹², Xiaobo Lu¹²*, Yichao Cao¹², Shudong Xia³ and Changqing Fu³

¹ School of Automation, Southeast University, Nanjing 210096, China
² Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education, Southeast University, Nanjing 210096, China
³ Beijing PKU TransInfo Technology Co., Ltd, Beijing 100193, China
E-mail: xblu2013@126.com
*Corresponding Author

Abstract: Moving vehicle detection based on video processing has been widely used in intelligent transportation system recently. However there are also many problems, such as dynamic background, ghost region, and shadow of moving objects. This paper proposes an improved ViBe object detection algorithm. First, an accurate background image is obtained by using the multi-frame averaging method, and then the background model is initialized by this accurate background image, thus effectively reducing the generation of ghost region. Whenever there is no moving object for a fixed number of consecutive frames in the video, this frame is updated to the background image. Conservative update strategy and foreground point counting method are adopted to update the background and reduce the impact of dynamic background on the foreground detection. Next, the foreground image detected by improved ViBe algorithm is input into the shadow elimination method proposed in this paper. Shadows in foreground pixels are detected in RGB color space, and then the pixels determined as shadows are eliminated. Finally, accurate moving vehicles are obtained. Our algorithm can effectively eliminate the shadows of moving vehicles, quickly adapt to background changes and illumination mutation, and get accurate moving objects, which is helpful for vehicle contour extraction and subsequent image processing.

1. Introduction

Moving vehicle detection based on video processing has many advantages, so it has been widely used in intelligent transportation system in recent years. However, the common moving object detection algorithms are generally affected by the dynamic background such as illumination changes, and do not suppress the shadow, which is not conducive to subsequent image processing.

Background modelling is the most commonly used method in motion detection. Among them, Van proposed a PBAS algorithm [2] with feedback information. Charles et al. constructed each pixel background model using spatio-temporal feature descriptors, and proposed a robust SuBSENSE algorithm [14], but this algorithm is time-consuming. LOBSTER algorithm [3] and PAWCS algorithm [4] use local binary similarity pattern (LBSP) [13] to construct background model, but the computation of LBSP is large, which reduces the real-time performance of the algorithm.

At present, the common moving object detection algorithms generally do not suppress the shadow [15]. In [6] detects shadows in HSV colour space, because shadows do not change the hue of the covered background, but reduce its intensity and saturation. In [7], the Gabor function is used to extract texture features of images. In [10], support vector machines are used to distinguish moving
shadows from moving objects. All statistical-based shadow detection requires a learning process to initialize model parameters [12].

This paper is focus on how to accurately detect and extract moving objects from videos taken on different time periods and different road sections. However, illumination and road conditions will change in varying degrees, and the commonly used object detection methods cannot detect the shadow of moving objects, so how to effectively remove the shadow of moving objects is also a key problem to be solved urgently. In this paper, an improved background initialization method of ViBe algorithm is proposed. The idea of multi-frame averaging is used to get accurate background image first, and then ViBe algorithm initializes background model with accurate background image, which can effectively eliminate ghost region. In the aspect of background updating, the specific updating strategy is to combine conservative updating strategy with foreground points counting method; and whenever there is no moving object for a fixed number of consecutive frames in the video, this frame is updated to the background image. In order to suppress the shadow of moving objects, the shadow elimination algorithm based on colour space is studied. The foreground image detected by improved ViBe algorithm is input into the proposed shadow elimination method. According to the shadow detection method in RGB colour space, the shadow of moving objects can be effectively removed. Accurate moving objects can be obtained.

2. Vibe Foreground Detection Algorithm

ViBe is a sample-based background model, which establishes a background sample set for each pixel, preserves the gray information of the background pixel, and updates the sample set with random update strategy. The ViBe background model mainly includes three aspects: the initialization method of the model, the update strategy of the model and the working principle of the model.

2.1. The Initialization Method of the Vibe Model

Initialization of ViBe background model is the process of model building. Generally, the first frame image of video is used to initialize background model. For each pixel, N pixels are randomly selected from its neighbourhood to establish a sample set, and 8 neighbourhoods are generally selected. As shown in the following formula (1):

\[ B_i(x, y) = \{P(x, y)|P(x, y) \in N_8(x, y)\}, BM(x, y) = \{B_1(x, y), B_2(x, y), ..., B_N(x, y)\} \]  

Among them, \( BM(x, y) \) is the sample set, \( B_i(x, y) \) is the sample point, \( i = 1, 2, ..., N \), and \( P(x, y) \) is the randomly selected point in the neighbourhood \( N_8(x, y) \).

2.2. The Update Strategy of the Vibe Model

The updating of ViBe background model is the updating of background sample set. When a pixel is classified as background point, it has a probability of \( 1/\beta \) to update its own background sample set with its current pixel value, where \( \beta \) is a time sub-sampling element. And each background point has a probability of \( 1/\beta \) to update the background sample set of any point in its spatial neighbourhood \( N_8(x, y) \) with its current pixel value.

2.3. Working Principle of ViBe

In ViBe model, the background model store a sample set for each pixel, and then compares each new pixel value with the sample set to determine whether it belongs to the background point. Specifically, \( V(x, y) \) is the pixel value at \( (x, y) \) position. \( R \) is the judgment threshold of the distance difference between the pixel and the sample set. If the distance is less than threshold \( R \), the number of approximate samples \( Num \) increases. When the number of approximate samples is greater than threshold \( \#min \), the pixel is determined as background point, otherwise it is foreground point. The rules for judging pixel \( (x, y) \) are as formula (2):

\[ Num = num\{\text{dist}(V(x, y), B_i(x, y)) < R\}, F(x, y) = \begin{cases} 
fg, & Num < \#min \\
bg, & Num > \#min
\end{cases} \]  

(2)
where $F(x, y) = fg$ indicates that the pixel $(x, y)$ belongs to the foreground, $F(x, y) = bg$ means that the pixel $(x, y)$ belongs to the background. The distance determination formula is:

$$\text{dist}(V(x, y), B_i(x, y)) = |V_{\text{gray}}(x, y) - B_{i\text{gray}}(x, y)|$$

(3)

where $V_{\text{gray}}(x, y)$ represents the gray value at the pixel point $(x, y)$, and $B_{i\text{gray}}(x, y)$ represents the gray value of the sample point in the background model $BM(x, y)$.

3. Improved Vibe Background Algorithm

ViBe algorithm uses the first frame initialization method. The disadvantage is that the first frame image may contain foreground objects, it will lead to inaccurate objects detected at the beginning time, and which is ghost region will appear in the early detection image. This paper proposes an improved ViBe detection method from the aspects of background initialization and background updating.

3.1. Improved Background Initialization Method

Background subtraction is the most commonly used method in the field of object detection. The background modeling method used in background subtraction method is multi-frame average method. Assuming the average of $N$ frames, when $N$ is large enough, because the position of the moving object in the background changes constantly and the background remains unchanged, the influence of the accumulated multi-frame moving object on the background will weaken continuously, so that a relatively stable background can be obtained. When the background is initialized, the initial period of video is selected, and the pixel values of each frame are accumulated according to the video sequence, and then the average value is calculated as the background model. As shown in formula (4):

$$BM_0 = \frac{1}{N}(\text{frame}_0 + \text{frame}_1 + \cdots + \text{frame}_{N-1})$$

(4)

where $BM_0$ is the background model calculated by multi-frame averaging method for the first $N$ frame images. The computed background image is used for the initial sample set of ViBe algorithm. This background image does not contain foreground objects, and can be used as the reference background image for the next shadow detection algorithm.

3.2. Improved Background Updating Method

After the background initialization is completed, the frame is updated to the background picture whenever there is no moving object in the video for a fixed number of consecutive frames. At this time, the background model can adapt to the changing background. This background image is used as a reference background image for the next shadow detection algorithm.

In the aspect of background updating, the specific updating strategy is to combine conservative updating strategy with foreground points counting method. The foreground point counting is introduced to count the pixels while judging whether it is foreground point. If a pixel is detected for foreground point for $n$ times, it can be considered that the point is mistaken detected, and then this pixel is updated as background point.

4. Shadow Elimination

In this paper, a shadow elimination algorithm based on color space is used to eliminate some shadows in the foreground, so as to obtain more accurate moving objects. Through the research and analysis of Phong illumination model [11], it is not difficult to get the criteria for distinguishing shadows in RGB color space. Firstly, the R, G, B components of the shadowed area are lower than the background pixels, The R channel decreases the most, the G channel takes the second place, and the B channel decreases the least.

In this paper, in the foreground region obtained by the improved vibe algorithm, the following formula (6) is used to determine whether a pixel is a shadow point. And $th_{i1}$, $th_{i2}$, $th_{i3}$ are the threshold values for judging the decline of color components in R, G and B channels respectively. When the difference between pixel and reference pixel exceeds the threshold, it is considered as a shadow point. When shooting video on highway, the moving object is usually the vehicle driving on
the highway, while some vehicles are dark vehicles. In order to avoid the false detection of dark vehicles as shadows, three thresholds $th_{h1}, th_{h2}, th_{h3}$ are set to distinguish the dark vehicle’s body from the shadow so as to protect the integrity of the moving object while removing more shadows. The specific shadow determination formula is as follows:

$$S_{RGB}(x,y) = \begin{cases} 1, & \text{if } th_{h1} > R(x_{ref},y_{ref}) - R(x,y) > th_{h2} \\ th_{h2} > G(x_{ref},y_{ref}) - G(x,y) > th_{h3} \\ th_{h3} > B(x_{ref},y_{ref}) - B(x,y) > th_{h3} \\ 0, & \text{otherwise} \end{cases}$$  \hspace{1cm} (5)

Among them, $R(x,y), G(x,y), B(x,y)$, $R(x_{ref},y_{ref}), G(x_{ref},y_{ref}), B(x_{ref},y_{ref})$ denotes the R, G and B components of the pixels located at $(x,y)$ position in the current frame image and in the reference background image, and where $S_{RGB}(x,y) = 1$ indicates that the pixel $(x,y)$ is a shadow pixel, $S_{RGB}(x,y) = 0$ means that the pixel $(x,y)$ is a non-shadow pixel.

Then the pixels detected as shadows are reset as background points, and the non-shadowed pixels are still foreground points, so as to achieve shadow elimination to obtain more accurate moving objects. Formula (6) is the foreground pixel determination rule with shadow elimination:

$$F(x,y) = \begin{cases} f, & \text{if Num} < \# \text{min and } S_{RGB}(x,y) = 0 \\ bg, & \text{otherwise} \end{cases}$$  \hspace{1cm} (6)

where $F(x,y) = f$ indicates that the pixel $(x,y)$ belongs to the foreground, $F(x,y) = bg$ means that the pixel $(x,y)$ belongs to the background.

5. Experiment Results and Analysing

5.1. Result of Improved Vibe Algorithm

In order to verify the effectiveness of the improved Vibe algorithm proposed in this paper, we compare the improved algorithm with the original Vibe algorithm.

We choose to add up to 1125 frames of video images to calculate the initial background image. The first frame of video with moving vehicles is shown in Fig.1 (a) which contains vehicles, and the accurate initial background image we calculated is shown in Fig.1 (b). There is no obvious foreground trace in the background model, which is close to the real background, and it takes a short time to start the subsequent detection algorithm.

![Figure 1](image1.png)

Figure 1. (a) The first frame of video. (b) The accurate initial background image. (c) Foreground binary image obtained by Vibe algorithm. (d) Foreground binary image obtained by improved Vibe algorithm.

![Figure 2](image2.png)

Figure 2. (a) The 50th frame of video. (b) Foreground binary image obtained by Vibe algorithm of the 50th frame. (c) Foreground binary image obtained by improved Vibe algorithm of the 50th frame. (d) The 165th frame of video. (e) Foreground binary image obtained by Vibe algorithm of the 165th frame. (f) Foreground binary image obtained by improved Vibe algorithm of the 165th frame.
As shown in Fig.2, within 165 frames of the start of the video, the traditional Vibe algorithm still has obvious ghost region. When the foreground object in the first frame leaves, a large ghost region is left in the original position and exists for a long time. However, the improved Vibe algorithm proposed in this paper has no obvious ghost region.

In the aspect of background updating, when there is no moving object in the video for 50 consecutive frames, this frame is updated to the background image. And we use conservative updating strategy with foreground points counting method, where the improved ViBe sample set size is 20, time sub-sampling element $\beta$ is 16. If a pixel is detected for foreground point for 50 times, and then this pixel is updated as background point.

Figure 3. (a) The 18650th frame of video. (b) Background image of the 18650th frame. (c) Foreground binary image obtained by improved ViBe algorithm of the 18650th frame. (d) The 52350th frame of video. (e) Background image of the 52350th frame. (f) Foreground binary image obtained by improved ViBe algorithm of the 52350th frame.

As shown in Fig.3, in the 18650th frame and 52350th frame of the video, the background images shown in Fig.3 (b) and Fig.3 (e) are nearly close to the current real background.

5.2. Result of Shadow Elimination Algorithm
In order to verify the effectiveness of the proposed shadow detection method, we conduct experiments on different videos to test the actual effect of this method. We improves Vibe algorithm to detect the foreground image of each frame, and then input the shadow detection method in this paper to eliminate the shadow. The reference background image in the experiment is obtained from the improved ViBe algorithm in the previous section. Through experiments and analysis, the selection of parameter thresholds in RGB spatial shadow determination is as follows: $th_{v1} = 0.06$, $th_{v2} = 0.05$, $th_{v3} = 0.02$, $th_{h1} = th_{h2} = th_{h3} = 110$

Figure 4. (a), (b) and (c) show video images containing moving objects with shadows; (d), (e) and (f) show the foreground binary image obtained by improved ViBe algorithm, and (g), (h) and (i) shows foreground binary image obtained by improved ViBe algorithm with shadow elimination.
As shown in Fig.4, shadow elimination algorithms in this paper can remove shadow effectively and get more accurate moving vehicles.

6. Conclusion
The work of this paper is to detect moving vehicles accurately from different videos. This paper mainly studies the object detection algorithm based on ViBe. The main work is as follows: To solve the problem of ghost region, we improve the initialization method of ViBe background model. First, we use the idea of multi-frame averaging to get a accurate background image, and then ViBe algorithm uses accurate background image to initialize background model, so as to effectively reduce the generation of ghost region; Aiming at the problem of dynamic background, When there is no moving object in the video for 50 consecutive frames, the frame is updated to the background image, and the foreground point counting method is used to update the background and reduce the impact of the dynamic background on the algorithm. Compared with the traditional ViBe algorithm, the experimental results show that the improved algorithm can detect moving vehicles more accurately. Next, a shadow elimination method is proposed. The foreground image detected by improved ViBe algorithm is input into the proposed shadow elimination method. Shadows in foreground pixels are detected in RGB color space, and then the pixels determined as shadows are eliminated. The experimental results show that the shadow of moving vehicles can be effectively eliminated and the accurate moving vehicles can be finally obtained.

7. Acknowledgments
This work was supported by the National Natural Science Foundation of China (No.61871123), Key Research and Development Program in Jiangsu Province (No.BE2016739) and a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

8. References
[1] O Barnich and M Van Droogenbroeck, "ViBe: A Universal Background Subtraction Algorithm for Video Sequences," in IEEE Transactions on Image Processing, vol. 20, no. 6, pp. 1709-1724, June 2011.
[2] M Hofmann, P Tiefenbacher and G Rigoll "Background segmentation with feedback: The Pixel-Based Adaptive Segmenter," 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Providence, RI, 2012, pp. 38-43.
[3] P St-Charles and G Bilodeau, "Improving background subtraction using Local Binary Similarity Patterns," IEEE Winter Conference on Applications of Computer Vision, Steamboat Springs, CO, 2014, pp. 509-515.
[4] P St-Charles, G Bilodeau and R Bergevin, "A Self-Adjusting Approach to Change Detection Based on Background Word Consensus," 2015 IEEE Winter Conference on Applications of Computer Vision, Waikoloa, HI, 2015, pp. 990-997.
[5] M Van Droogenbroeck and O Paquot, "Background subtraction: Experiments and improvements for ViBe," 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Providence, RI, 2012, pp. 32-37.
[6] R Cucchiara, C Grana, M Piccardi, A Prati and S Sirotto, "Improving shadow suppression in moving object detection with HSV color information," ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No.01TH8585), Oakland, CA, 2001, pp. 334-339.
[7] Salvador E , Cavallaro A. and Ebrahimi T. (2004). Cast shadow segmentation using invariant color features. Computer Vision and Image Understanding, 95(2), 238-259.
[8] Zhang W, Fang X Z, Yang X K and Wu Q M J (2007). Moving cast shadows detection using ratio edge. IEEE Transactions on Multimedia, 9(6), 1202-1214.
[9] Tanaka T, Shimada A, Arita D and Taniguchi R I. (2007). Non-parametric background and shadow modeling for object detection. Asian Conference on Computer Vision 159-168.
[10] Joshi A J and Papanikolopoulos N P. (2008). Learning to detect moving shadows in dynamic environments. IEEE Transactions on Pattern Analysis & Machine Intelligence,30(11), 2055-2063.
[11] Phong B T. (1975). Illumination for computer generated pictures. *Communications of the Acm, 18*(6), 311-317.

[12] Russell M, Zou J J and Fang G. (2016). An evaluation of moving shadow detection techniques. *Computational Visual Media, 2*(3), 195-217.

[13] G Bilodeau, J Jodoin and N Saunier, "Change Detection in Feature Space Using Local Binary Similarity Patterns," *2013 International Conference on Computer and Robot Vision, Regina, SK, 2013, pp. 106-112.*

[14] St-Charles P L, Bilodeau G A and Bergevin R. (2014). Subsense : a universal change detection method with local adaptive sensitivity. *IEEE Transactions on Image Processing, 24*(1), 359-373.

[15] Bouwmans and Thierry. (2014). Traditional and recent approaches in background modeling for foreground detection: an overview. *Computer Science Review, 11-12, 31-66.*

[16] Kim K, Chalidabhongse T H, Harwood D and Davis L. (2005). Real-time foreground–background segmentation using codebook model. *Real-Time Imaging, 11*(3), 172-185.