RESEARCH ARTICLE OPEN ACCESS

A Novel Method of Moving Object Detection

JielongHuang*, Hongqi Li**,Lianglu Yin**, Yuanliang Huang**
*(College of Science and Engineering, Jinan University, Guangzhou, PR China)
**(Institute of Electrical Automation, Jinan University, Zhuhai, PR China)

--************************----------------------------------

Abstract:

Visual background extractor (Vibe), a universal background subtraction algorithm, which has the characteristics of fast operation, high accuracy and wide application range. But there are also difficulties in eliminating the "ghost" and adapting to the light mutation, in this paper, an improved Vibe algorithm is proposed. It draws on the W^4 theory and the mean value method in establishing the background model to suppress ghosts and some shadows, the new background model of Vibe was created by adding light mutation criterion to the motion detection algorithm, so that it could adapt to the light mutation rapidly, the experimental results show that the proposed method have good performance in removing interfaces as ghost shadow, when any sudden light change of circumstance occurs, the algorithm can also detect and update pixels statistical parameters adaptively.

Keywords **—W^4 model,Vibe, motion detection, ghost suppression, light mutation.**

--************************----------------------------------

I. INTRODUCTION

The purpose of motion detection is to distinguish moving objects (hereafter referred to as foreground)from scene (called background). According to the different background conditions, there are two kinds of situations, one is static, and other is changes dynamically over time, including illumination changes (changing time of day, clouds,etc), motion changes and others. Target tracking, processing recognition and behavior understanding directly affected by motion detection results, so the motion detection is the basis and key of intelligent video surveillance system, it also one of the hot and difficult issues in the research of machine vision and digital image processing.

Over the years, increasingly motion detection algorithms have been proposed, including optical flow method [1], interframe difference method [2], background subtraction method[3-5] and so on.The optical flow method usescharacteristics of the moving target over time to establish the optical flow constraint equation for motion detection, which is computationally complicated and could not meet the real-time requirements in practical application.The interframe difference method extracts the moving objects by using two or three adjacent differential thresholds. The algorithm is simple, the calculation is small and easy to implement. However, its accuracy is not high and the detected objects often have holes.Background subtraction can be formulated as a technique that establishing a model of background and then compares this model with the current frame to detect region where a significant difference occurs and finally updating the background model.Compared with the former two methods, the background subtraction method is the most commonly used method to solve the problem of motion detection at present, with high accuracy, good performance in real-time and wide application.

The key to the detection effect of background subtraction is the establishment of background model, the initialization and the updating of the model. Haritaoglu[6, 7] and others proposed a motion detection algorithm based on background model in W^4 human behavior detection system. This algorithm is efficient and less memory consumption. In addition, there are also single Gaussian models [8, 9], mixed Gaussian models [10-12], codebook models, etc.Background subtraction based on the background algorithm is proposed, it also have their respective advantages and drawbacks.

The visual background extraction algorithm is a general purpose motion detection algorithm proposed by Olivier Barnich and other scholars,the advantages of the algorithm include simple calculation, fast computing speed, high detection accuracy and good detection effect in different environments.However, the method lacks of consideration for the suppression of ghosts and the detection of moving targets in the case of light mutation, and several surveys are devoted to these issues. A method of rapid inhibition of ghosting is proposed in [17], according to [18], combined with the frame difference method to obtain the real background model so as to eliminate the ghost effect, literature [19] focus on the requirements that the vibe algorithm could adapt to gradual or fast illumination changes. In addition, a number of scholars have proposed some improved algorithms based on Vibe [20-22].In our method, we improve the Vibe algorithm to combined with the W4 algorithm and the mean model ideas, eliminate the shadow and add the analyzing conditions of light mutation to make the algorithm more suitable for illumination change and enhance the robustness of the algorithm. The paper is organized as follows: In section 1, the advantages and disadvantages of main methods in motion detection has been briefly described. Section 2 introduces W^4 theory, mean value model idea and Vibe algorithm, which are the foundation of our algorithm. Section 3 describes our skills and details of our major innovations. Section 4 discusses experimental results, show that our method outperform the original Vibe algorithm. Finally, the conclusion is presented in section 5.

II. FOUNDATION

Before introducing our proposed algorithm, we hope that readers will understand some of the following theoretical algorithms

A. *Mean Model*

The mean value model is a rather simple but effective method. It is essentially a statistical filtering idea, in a period of time, the collected multi frame images are added together to get the average value. The average value is used as a reference background model. The specific formula is as follows:

$$Background(x,y) = \frac{1}{n}\sum_{i=1}^{n} f_i(x,y) \quad (2)$$

Where background (x, y)denoted the background model, $n$refers to frames,$f_i(x,y)$ represents the pixel value of the i frame image at the location(x,y).

B. *W^4 Background Model*

Haritaogluetal presented a technique for the background model called W^4 model. It uses three values to represent each pixel in the background image: the minimum and maximum intensity values, and the maximum intensity difference between consecutive images of the training sequence. It is mainly divided into two steps:

- The first n frames of the video are buffered, and the median of gray values and standard deviations of each pixel are calculated, and some noises are filtered out by certain conditions to obtain a set of stable samples for each pixel.
- Then the maximum and minimum intensity values, and the maximum intensity difference between consecutive frames of the training sequence for each pixel are calculated, and the background model is built and initialized. Suppose that V represents a set of frames initialized by the model, $Vi(x)$ represents the gray value of the i^{th}frame location at x in the set V, and $\sigma(x)$, $\lambda(x)$ denote the gray standard deviation and gray mean value at pixel x, $m(x)$, $n(x)$ and $d(x)$ denote the minimum and maximum intensity values, and maximum interframe intensity difference value at pixel x, respectively. The initial background model for pixel x is:

$$\begin{bmatrix} m(x) \\ n(x) \\ d(x) \end{bmatrix} = \begin{bmatrix} \min_z\{V^z(x)\} \\ \max_z\{V^z(x)\} \\ \max_z\{|V^z(x) - V^{z-1}(x)|\} \end{bmatrix}$$

While $|V^Z(x) - \lambda(x)| < 2 * \sigma(x)$.

Calculate [$m(x)$, $n(x)$, $d(x)$] for each pixel x, and then we got a completed W^4 model.

C. Analysis of Vibe Algorithm

Vibe algorithm is a motion detection method based on sample random clustering. The underlying idea is to store a set of samples for each background pixel, then the current value of the pixel is compared to its closest samples within the collection of samples to be classified as foreground pixel or background pixel. The technique mainly consists of three processes:

1) ***Background model building and initialization:***There is an assumption as the author of [23], which is that neighboring pixels share a similar temporal distribution. So the Vibe algorithm populate the pixel models with values found in the spatial neighborhood of each pixel, and fill them with values randomly taken in their neighborhood in the first frame, it will build and initialize background model from a single frame.

2) ***Motion detection:***Motion detection is a classification problem, which is to determine whether pixels are background or foreground (also called moving objects), so as to avoid the effect of any outliers, we classify a new pixel value with respect to its immediate neighborhood in the chosen color space. The process is shown in Fig. 1. Let $v(x)$ denote the value in a given Euclidean color space taken by the pixel located at x in the frame, and by $v_i(x)$ a background sample value with an index i. Each background pixel x is modeled by a collection of N background sample values taken in previous frames: $M(x) = \{v_1, v_2, \dots, v_N\}$.To classify $v(x)$, we compare it to the closet values within the set of samples by defining a sphere SR($v(x)$) of radius R centered on $v(x)$, and count the number of samples of M(x) intersecting the $SR(v(x))$, the pixel value $v(x)$ is then classified as foreground if the number is less than a given threshold *#min*, otherwise is classified as background, the formula is as follows:

$$g(x) = \begin{cases} 0, \{S_R(v(x)) \cap M(x)\} \geq \#\min \\ 1, \{S_R(v(x)) \cap M(x)\} < \#\min \end{cases} \quad (2)$$

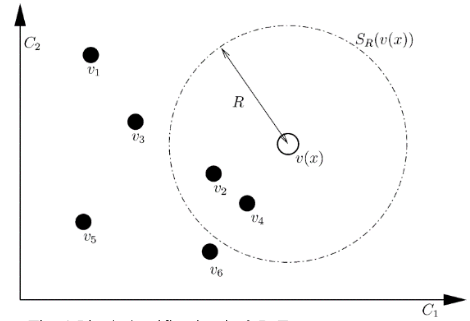

Fig. 1 Pixel classification in 2-D European space

3) ***Background model update:***Vibe algorithm adopted a conservative update policy. If a pixel value $v(x)$ is classified as the foreground pixel, the background samples being preserved. If not, then, a background pixel value has one chance in φ of being selected to update its pixel values model, the probability of a sample being preserved for the interval (t, $t+dt$) is equal to:

$$P(t, t+dt) = e^{-\ln(\frac{N}{N-1})dt} \quad (3)$$

This expression shows that the expected remaining lifespan of any sample value of the model decays exponentially. It has nothing to do with the time t, it is a property called memoryless property. The life cycle of the historical sample values of the background model is prolonged by time subsampling, which shortens the frequency of background updating. The random update of neighborhood points also enhances the adaptability to its variability over time.

III. OUR DESIGN

As stated earlier, Vibe is refer to as a universal method for motion detection with good speed and high accuracy. We combine the mean value method and the core principle of W^4 algorithm to improve it, in view of some problems existing in the practical application. The main process is as follows.

D. Construction of Background Model

As we known, a ghost is irrelevant for motion interpretation and has to be discarded, in order to eliminate the influence of ghost, noise, and others, a relatively real background is obtained, and a real and stable background model is established by the mean value method and the W^4 algorithm. Firstly, for each pixel x, a new value is calculated by linear weighting the maximum gray value, denoted $n(x)$, and the minimum gray value, denoted $m(x)$, to form a background model I, and the reason why is to minimize the impact of foreground motion noise as far as possible, the formula is as follows:

$$nm(x) = \mu_1 * n(x) + \mu_2 * m(x) \quad (4)$$

Then it is merged with the background model obtained from the mean method to obtain a more stable and reliable background II:

$$BG = \theta_1 * nm(x) + \theta_2 * Background(x) \quad (5)$$

where μ_1, μ_2, θ_1, and θ_2 are all weighted coefficients, μ_1 and μ_2 are generally assumed to be 0.8 and 0.2, respectively, and in order to make the background model II more stable, θ_2 is usually larger than θ_1, in fact, the effect will be better when θ_1=0.3, θ_2=0.7.

E. Moving Object Detection

The relative real background model denoted by *BG*, is regarded as the first frame of the video sequence to be initialized, then get a completed set of background sample values. From the beginning of the i^{th} frame, comparison of a pixel value with a set of samples in a 2-D Euclidean color space. To classify pixel *x*, we count the number of samples of model intersecting the sphere of radius R centered on pixel *x*. While number is larger than or equal to a given threshold *#min*, the pixel *x* is classified as background region, otherwise, it is the moving object, also called foreground.

F. Background updating and light mutation judgment

The background model updating strategy is a crucial step if we want to achieve accurate results over time, and it need to be able to adapt to lighting changes. The classical approach to updating the background model is to discard and replace old values by new ones after a number of frames (or after a given period of time). Compared to this method, the Vibe algorithm adopt a random subsampling policy to update the corresponding pixel values model, a time subsampling factor denoted by φ, if the pixel of current frame have been classified as a background pixel, which has one chance in φ of being selected to update its pixel model. At the same time, it randomly replace a sample of its neighborhood pixel values model with a $1/\varphi$ probability, in spite of this, it is still not stable to light mutation and ghosts situations. For the sake of improve it, a lot of work has been done for this. Firstly, if the times that a pixel is continuously detected as a foreground region pixel is larger than a given threshold *k*, the pixel should be updated to a background pixel value. The reason for this is to effectively prevent the occurrence of an object that is still detected as a moving target after it has entered a background area and then becomes a part of a new background (another kind of ghost situation). Secondly, when the light changes suddenly, the original background model is no longer applicable, then we are inspired by the background update strategy of W^4 method. If the ratio between pixels were classified as moving objects and all the pixels in entire image is greater than eighty percent, it is reasonable to believe that there is a mutation in the light, and we should discard the original background model, rebuild the background model and initialize it, then detect the moving objects. Our method flowchart is illustrated in Fig. 2.

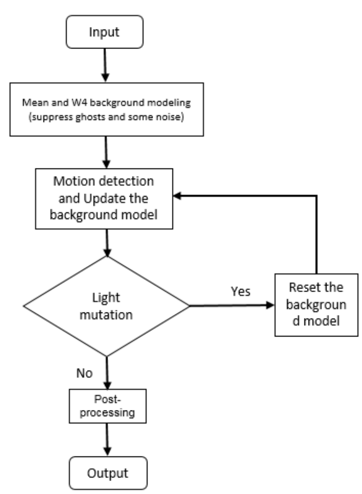

Fig. 2 Method flow chart

The detailed steps of the algorithm are as follows:
Input: video sequence frame.
Output: moving target image.
Step1: After caching the first n frames of the video sequence, combined with the mean method and the W^4 model to build a stable and reliable real background model.
Step2: Initialize the background model from a single frame obtained in Step1.
Step3:Motion detection.
Step4: According to the illumination situation, if the light breaks, the background model is reset by

the current frame, then Step3; if not, return to Step3 directly.

Step5:Output after image post-processing (mainly includes the removal of small target points and filtering for the detection results).

TABLE I
COMPARISON OF PROCESSING SPEED (S / FRAME)

Algorithm	Average time of process（s/fame）	
	Video1	Video2
Vibe	0.0177813	0.0175939
Our method	0.0182597	0.0188417

IV. EXPERIMENT AND ANALYSIS

Our platform: 2.7GHz Intel (R) Core i5-4590 CPU, 4GB of RAM, 32 bit operating system based on x64 processor, Configuring VS2010 for opencv2.4.8, C implementation. Parameters were set to:

μ_1=0.8, μ_2=0.2, θ_1=0.3, θ_2=0.7, N=20, R=20, *#min*=2, φ=16, *Background*＝0, *Foreground*＝255.

Two videos were tested in our experiment, one is highwayII_raw.avi from the general test video set, refered to as Video1 with a resolution of 320*240 and framerate of 15 FPS. Another video simulates the change of light, which is also called Video2 with a resolution of 320*240 and the framerate of 25 FPS. Fig. 3 (a) is the inaccurate background image generated during the initialization process of the original Vibe algorithm (At this time, the moving object is contained in the first frame of the video sequence), and Fig. 3 (b) is the stable real background image obtained through our algorithm, it takes about 1.23228 seconds.

Fig. 3 Background image

Comparing our technique with the original vibe algorithm, it is easy to know that our method has a better performance, and the results are as follows: From Fig. 4, we can see that in the tested video frames, there is a ghosting phenomenon in the original Vibe algorithm, and now we can get rid of the ghosts that are not needed and detect the real moving targets in ours method.

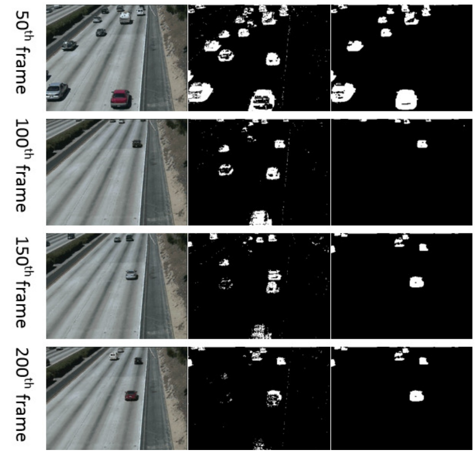

Fig. 2 Comparison of suppression ghosts

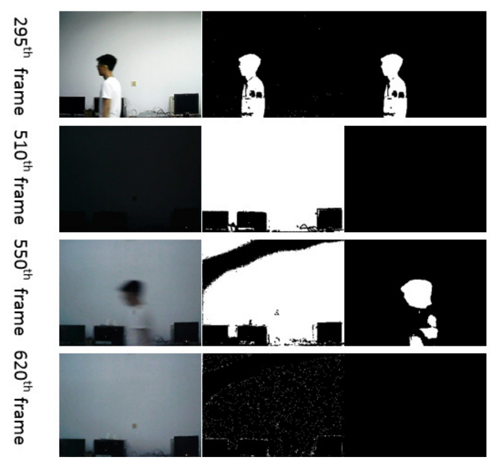

Fig. 3Comparison of the effect of adaptation to light change

The results of testes Video2 verify that the proposed algorithm is also efficient in detecting targets in the environment where there is a light mutation. As shown in Figure 5, we selected 295^{th} 510^{th}, 550^{th}, 620^{th} frame of a video sequence, at frame 295^{th} (before the light mutation), Vibe algorithm and ours algorithm can both accurately detect the moving target. However, when the light changes abrupt, the Vibe algorithm has been seriously misdiagnosed at the 510^{th}, 550^{th} frame. It failed until the 620^{th} frame is gradually restored to normal, and the process is slow. But our method can quickly reduce the influence of light mutation and detect the moving target accurately.

Comparing the processing speed of the two algorithms as shown in Table 1. In two different video sequences, the average processing time per frame of Vibe algorithm were 0.0177813s and 0.0175939s respectively, while the average processing time per frame of our algorithm are 0.0182597s and 0.0188417s, compared to the former, increased time is almost negligible, most importantly, the new algorithm can meet the requirements of real-time perfectly.

V. CONCLUSION

In this paper, a new algorithm to detect moving object detection based on Vibe algorithm is proposed, it combines the advantages of W^4 model and mean method, and considers the situation of light mutation and has good robustness. Experiments show that the new algorithm has good performance in suppressing ghost and has good adaptability to illumination change, even in case of sudden change of illumination, it can still detect moving objects accurately, in addition, the algorithm satisfies real-time requirement in speed. The disadvantage of the algorithm is that it does not take into account the impact of the shadow onthe results of the detection, in the future, maybe we should add module of shadow-remove to reduce interference by shadow for the sake of better results (higher accuracy), of course, the algorithm still needs to meet the real-time conditions.

ACKNOWLEDGMENT

The authors sincerely thank Telecommunications and Imaging Laboratory of Institute Montefiore-University of Liege, Belgium share the Vibe-sources and programs at http://www.telecom.ulg.ac.be/research/vibe/.

REFERENCES

[1] B. K. Horn and B. G.Schunck,"Determing Optical Flow," *Artificial Intelligence*., vol. 17, pp.185-203, Nov. 1981.

[2] R. T. Collins, A. Lipton, and T. Kanade, "A System for Video Surveillance and Monitoring," Pittsburg: Carnegie Mellon University, the Robotics Institute, 2000:3-39.

[3] M. Piccardi, Ed., "Background subtraction techniques: a review," Systems, man and cybernetics, 2004 IEEE international conference on. IEEE, 2004, 4: 3099-3104.

[4] S.Brutzer, B. H?ferlin, and G.Heidemann, "Evaluation of background subtraction techniques for video surveillance," Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011: 1937-1944.

[5] Y. Benezeth, P. M.Jodoin, and B.Emile, "Comparative study of background subtraction algorithms," Journal of Electronic Imaging, 2010, 19(3): 033003-033003-12.

[6] I. Haritaoglu,D. Harwood, and L. S.Davis,"W/sup 4: Who? When? Where? What? A real time system for detecting and tracking people," Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on. IEEE, 1998: 222-227.

[7] I. Haritaoglu, D. Harwood, andL. S.Davis, "W/sup 4: real-time surveillance of people and their activities," IEEE Transactions on pattern analysis and machine intelligence, 2000, 22(8): 809-830.

[8] C. Wren, A. Azarhayejani, and T. Darrell, "Pfinder: real-time tracking of the human body," IEEE. Transactions on Pattern Analysis and Machine Intelligence, 1997,19(7):780-785.

[9] X. P. Wang, L. J. Zhang, and J. Chang,"Improved method of moving objects detection based on single-gaussian background model," Computer Engineering and Applications, 2009,45(21):118-120. (in Chinese)

[10] C. Stauffer, and W. E. Grimson, "Adaptive background mixture models for real-time tracking," Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on. IEEE, 1999, 2: 246-252.

[11] Z. Zivkovic "Improved adaptive Gaussian mixture model for background subtraction," Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. IEEE, 2004, 2: 28-31.

[12] Z. Zivkovic, and Van Der HeijdenF, "Efficient, "adaptive density estimation per image pixel for the task of background subtraction," Pattern recognition letters, 2006, 27(7): 773-780.

[13] K. Kim, T. H. Chalidabhongse,and D. Harwood, "Real-time foreground–background segmentation using codebook model," Real-time imaging, 2005, 11(3): 172-185.

[14] C. Xu, Z. Tian, and R. F. Li, "A Fast Motion Detection Method Based on Improved Codebook Model," Journal of Computer Research and Development,2010,47(12):2149-2156. (in Chinese)

[15] O. Barnich, and D. M. Van, "ViBe: A Universal Background Subtraction Algorithm for Video Sequences," IEEE Transactions on Image Processing 2011,20(6):1709-1724.

[16] D. M. Van, O. Paquot, "Background subtraction: Experiments and improvements for ViBe," Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on. IEEE, 2012: 32-37.

[17] L. Chen, X. Z.Chen, and Z. T. FAN, "Ghost suppression algorithm based on Vibe," Journal of China University of Metrology,2013, 24(4):425-429.

[18] X. R. Hu, H. Sun, "Novel moving object detection method based on ViBe," Computer Science,2014,41(2):149-152.

[19] J. Ding, L. Q. Kuang, and X. Han, "Improved moving object detection method based on ViBe," Computer Engineering and Design,2017,38(2): 374-378.

[20] H. B. Li, "An Improved Moving Object Detection Algorithm Based on ViBe," Journal of Image and Signal Processing,2017,06(01): 52-61.

[21] Z. H. He,S. Huang, and G. Ran, "An improved visual background extractor model for moving objects detection algorithm," Journal of Chinese Computer Systems,2015, 36(11): 2559-2562.

[22] Y. Yu, M. W. Cao, and F. Yu, "EVibe: An improved Vibe algorithm for detecting moving objects," Chinese Journal of Scientific Instrument, 2014, 35(4): 924-931.

[23] P. M. Jodoin, M. Mignotte, and J. Konrad, "Statistical background subtraction using spatial cues," IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(12): 1758-1763.