

## A Background Modeling Approach Based on Visual Background Extractor

Taotao Liu<sup>1, a</sup>, Lin Qi<sup>2, b</sup> and Guichi Liu<sup>2, c</sup>

<sup>1</sup>School of Zhengzhou University, Henan 450001, China;

<sup>2</sup>School of Zhengzhou University, Henan 450001, China;

<sup>3</sup>School of Zhengzhou University, Henan 450001, China

<sup>a</sup>iettliu@163.com, <sup>b</sup>ielqi@zzu.edu.cn

**Keywords:** ViBe, Background modeling, Moving objects detection, Frame subtraction.

**Abstract.** Moving objects detection is an important research content in computer vision field, and it plays a very important role in many vision applications such as smart video surveillance, intelligent transportation, and human-computer interaction, traffic control, activity recognition, object tracking and behavior understanding. The visual background extractor (ViBe) is a classical approach, however, when a moving objects in the initialized frame, it will take a lot time to fuse the gost and when there braches or surface of the water in the frame, it can make more false detection. In this paper, an algorithm based on ViBe is proposed. Though frame subtraction we acquire a real background, moreover we initial the model. On the other hand, we use 3-5 frames initial the model. The improved approaches solve the problems of gost or blinking pixels. Experiment show that, the algorithm can improve the effective of detection in the specific scene with similar computer load.

### Introduction

Moving objects detection which extracts foreground objects from the background for a static camera has been extensively studied for many years [1,2,3,4,5,6]. Moving objects detection plays a very important role in many vision applications with the purpose of subtracting interesting target area and locating the moving objects from image sequences. It is widely used in vision systems such as smart video surveillance, intelligent transportation, and human-computer interaction, traffic control, activity recognition, object tracking and behavior understanding. Accurate moving object detection is essential for the robustness of intelligent video-surveillance systems. Background subtraction and temporal differencing are two popular approaches for moving objects detection in an image sequence under a stationary camera. The Gaussian mixture model (GMM) [7] which is a widely used because of its self learning capacity and its robustness to variations in lighting. However, some false positive pixels will be produced in highly dynamic scenes where dynamic texture does not repeat exactly. W4 [8] use three parameters to model each pixel of background, but it generally can only recover partial edge shapes of moving objects. Li [9] use spatial information at feature level, such as color and gradient. This can improve the accuracy of the background model and is most suitable for the stationary background. Olivier Barnich and Marc Van Droogenbroeck proposed Vibe algorithm [10,11]. The method adopt neighboring pixels to create the background model, but if there is a moving object in the background when the modeling initializing there will make gost in detecting. Moreover, because of ViBe use only one frame to initialize the model, when the background have many blinking pixels which can make more false detection.

In order to overcome the shortcoming of origin VIBE and provide accurate result with good spatial and temporal consistency, In this paper we present a novel moving object detection method with frame difference and 3-5 frames fuse based on VIBE to solve the problem in special scene. Firstly we obtain a real model not including moving objects present in the beginning frames which cause gost in the result of detection use the frame difference to initialize the models of background. Moreover, to decrease false detection in the scenes of surface of water or shaky branch which make a lot of blinking pixels, we make use of 3-5frames to initialize the models that includes possible pixel values of

blinking pixels. Experiment show that, the improved algorithm can improve the effective of detection in the specific scene with similar computer load.

### Visual background extractor

Visual Background Extractor (ViBe) adopts neighboring pixels to establish the background model and distinguish background and foreground by comparing the background model with the current pixel value. The implementation of the algorithm which is a universal background modeling method is subdivided into three steps [11].

**Background Model Initialization.** The first step of ViBe is to initialize each pixel in the background model from the single frame image which is an important innovation of ViBe. As we know there is no temporal information in a single frame and it is supposed that neighboring pixels share a similar temporal distribution. This means that the value of a pixel and its neighbor pixel values in spatial domain has a similar distribution. The right size range of the neighborhood can be 4 or 8, and a suitable range can make sure that the background model includes a sufficient number of different samples, while keeping in mind that as the neighborhood scope increases, the correlation between pixel values at different locations decreases. Let's  $t=0$  denotes the first frame, so we can get the model as follow:

$$BG^0(x, y) = \{P_m^0(x^n, y^n) | (x^n, y^n) \in N_G(x, y)\} \quad (1)$$

Where  $BG^0(x, y), N_G(x, y), P^0(x, y)$  denotes the pixel background model value, spatial neighborhood value, pixel value respectively, and  $(x^n, y^n)$  in  $N_G(x, y)$  is selected such as probability.  $m = 1, 2, \dots, N$  is the number of samples. The method of initializing of model was showed in Fig.1

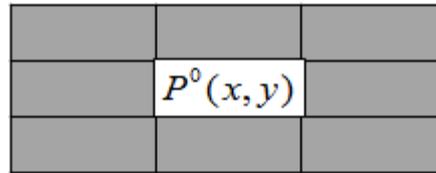


Fig. 1 Initialize of ViBe

**Pixel Classification Process.** Through the initialized model, we can use a simple decision process to determine whether a pixel of the input frame is belongs to background or not. As described previous, each background pixel was modeled with a set of samples which from the pixel's neighbor, though compare a new value to the background samples and then determine whether the new value belongs to background or not. Formally, let us denote by  $P(x, y)$  the value in a given Euclidean color space taken by the pixel located at  $(x, y)$  in the image, and by  $P_i(x, y)$  a background sample value with an index  $i$ . Each background pixel  $(x, y)$  is modeled by a collection of  $N$  background sample values

$$S(x, y) = \{P_1(x, y), P_2(x, y), \dots, P_N(x, y)\} \quad (2)$$

To classify a pixel value  $P(x, y)$  according to its corresponding model  $S(x, y)$ , we compare it to the closest values within the set of samples by defining a sphere  $S_R(P(x, y))$  of radius  $R$  centered on  $P(x, y)$ . The process can be showed as Fig.2

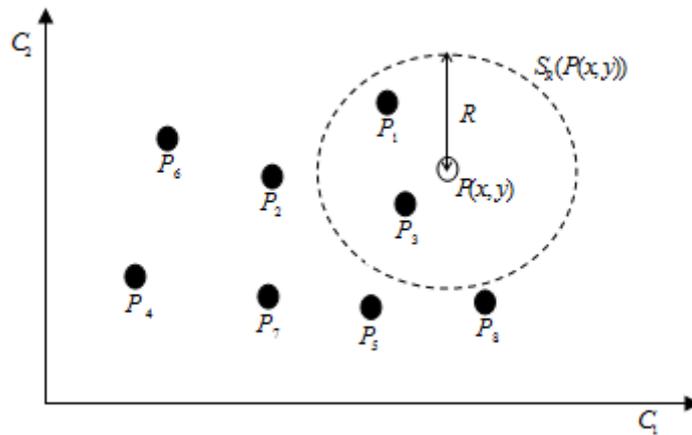


Fig. 2 Pixel Classification Process in a Euclidean color space ( $C_1, C_2$ )

As indicated in Fig.2, the pixel value  $P(x, y)$  is then classified as background if the number of the set intersection of this sphere and the collection of model samples  $S_R(P(x, y))$  is larger than or equal to a given threshold  $T$ . With many experiments, when a unique radius ( $R$ ) of 20, a cardinality ( $T$ ) of 2, and number of samples ( $N$ ) of 20 it can have a more appropriate.

**Updating the Background Model Over Time.** Another important step of a background model is how to updating the background model over time which ensure to achieve accurate results with each new frame. ViBe uses a method similar to conservative update policy which never includes a sample belonging to a foreground region in the background model. When a pixel classified as background, a random process determines whether this value is used to update the corresponding pixel model, if the time sampling is accord with the requirement predefined, it will replace a sample of the model randomly, at the same time its neighbor pixel is highly possible to be considered as the background pixel with high possibility, so the value of the pixel will replace a sample of its neighbor model ensure spatial coherence. In this update method incorporates there are three important components: a memoryless update policy because of replacing the sample randomly which ensures the samples stored in the background pixel models decay in the model smoothly. And the random replacing scheme of time subsampling extends the time windows covered by the background pixel models. A mechanism ensures spatial consistency though propagates background pixel samples spatially which merge the foreground objects which halt suddenly or stay long into the background model. The updating scheme may have different output because of the random renew method. It can be shown as Fig.3.

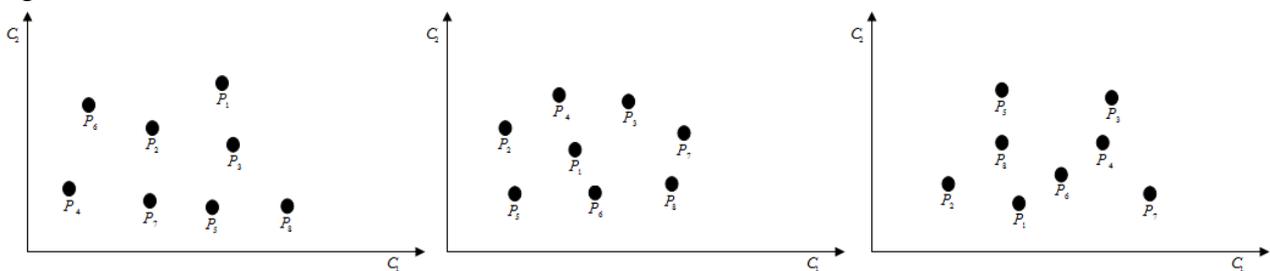


Fig. 3 Three possible outcomes of updating when a pixel model size  $N$  is equal to 8

### Improved Algorithm based on Visual Background Extractor

**Fuse with Frame Subtraction to Initialize the Model.** The algorithm of frame subtraction may be the earliest and simplest method of moving objects detection, but in the result of detections it have holes in the moving objects especially there is a big object in the scene. As represented above, it is a challenge that moving objects in the scene at the beginning of initializing model of the traditional ViBe. To solve that problem, we fuse with the frame subtraction to initialize the model, we draw a flow chart as Fig. 4 to express the method more clearly.

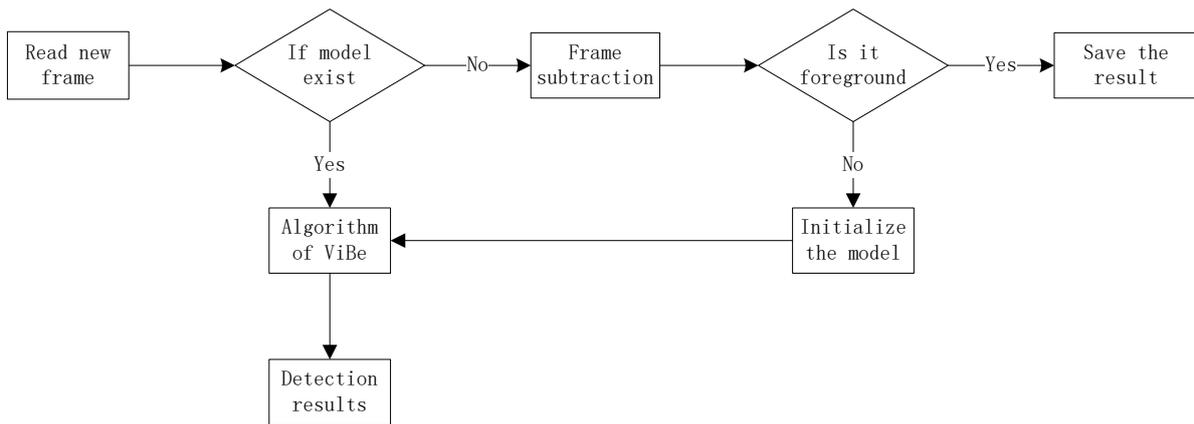


Fig. 4 The flow char of the method of fuse with frame subtraction to initialize the model

From the flow char above, when a new frame was read, firstly we judge whether the model of the pixel is exist. If the model existed, though the method of ViBe we can know if the pixel is a background or not; otherwise we though frame subtraction judge whether we initialize the model or not. By this way, we build a better model and we can have a more accurate detection result with the gost fuse more quickly.

**Multi Frame Fuse to Initialize Model.** When blinking pixels like surface of water or shaky branch exist in the first frame, it can lead much false detection. To get an effective detection, we fuse multi frame maybe 3or 5 frames to initialize the model which include more background information that is circularly. Once the model built, the algorithm can detect the moving object more effectively. The progress of the improved algorithm as the Fig. 5

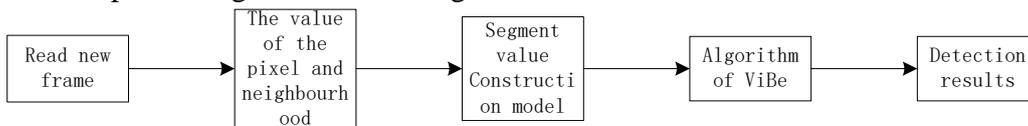


Fig.5 The flow char of the method of multi frame fuse to initialize model

### Result of the Improved Algorithm and Analysis

To validate the improved algorithm, we simulate the algorithm described above in many different scenes, including indoor and outdoor and so on. In a general scene, the improved algorithms have a same effect. But in the special scenes especially in the scenes there is a moving object in the first frame which is very important for the model initializing, and in the scenes that there blinking pixels like water face or shaky branches the improved algorithm has a better effective with only a little computer load. The effective of the improved algorithm can see clearly from the Fig.6-Fig.7

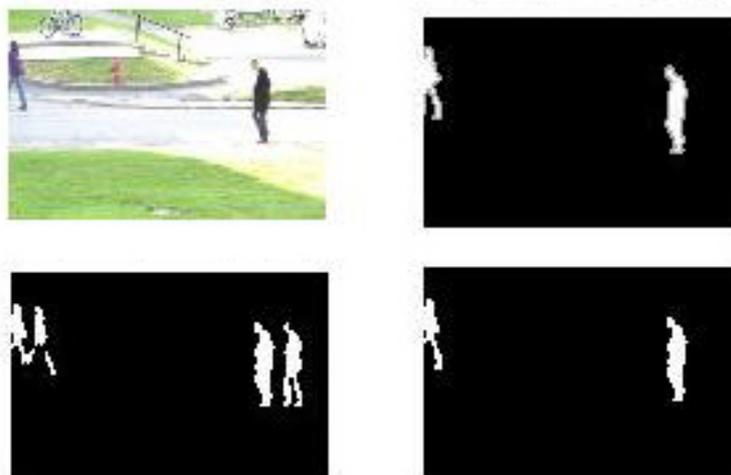


Fig. 6 The effect of improve with fuse with frame subtraction to initialize model



Fig. 7 The effect of improve with multi frame fuse to initialize model

In the Fig. 6 and Fig. 7, the first image is the new frame, and the second image is the ground truth of the original image, the third image is result of the original ViBe, and the last one is the effective of our improve algorithm. It is clearly that our improved algorithms have accurate effective.

### Summary

Though fuse with frame subtraction to initialize the model can solve the problem that a moving object in the first frame which is very important for initializing the model. And though multi frame fuse to initialize model we can have a better effective detection with water face or shaky branch in the scenes. As expressed above, the improved algorithm can have better effective without increase the computation complexity.

### Acknowledgment

This work was supported by the National Natural Science Foundation (61331021)

### References

- [1] T. Bouwmans, Recent advanced statistical background modeling for foreground detection: a systematic survey, *RPCS* 4 (3) (2011) 147–176.
- [2] T. Bouwmans, F. El Baf, B. Vachon, Statistical background modeling for foreground detection: a survey, in: *Handbook of Pattern Recognition and Computer Vision (volume 4)*, World Scientific Publishing, Singapore, 2010, pp. 181–199 (Ch. 3).
- [3] L. Tong et al., “Encoder combined video moving object detection,” *Neurocomputing* 139, 150–162 (2014).
- [4] O. Oreifej, X. Li, and M. Shah, “Simultaneous video stabilization and moving object detection in turbulence,” *IEEE Trans. Pattern Anal. Mach. Intell.* 35(2), 450–462 (2013).
- [5] J. Guo et al., “Fast background subtraction based on a multilayer codebook model for moving object detection,” *IEEE Trans. Circuits Syst. Video Technol.* 23(10), 1809–1821 (2013).

- [6] C. Cuevas and N. García, “Improved background modeling for realtime spatio-temporal non-parametric moving object detection strategies,” *Image Vis. Comput.* 31(9), 616–630 (2013).
- [7] C. Stauffer, E. Grimson, Learning patterns of activity using realtime tracking, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (8) (2000) 747–757.
- [8] I. Haritaoglu, D. Harwood, L. Davis, W4: real-time surveillance of people and their activities, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (8) (2000) 809–830.
- [9] L. Li, W. Huang, I. Gu, Q. Tian, Foreground object detection from videos containing complex background, in: *ACM International Conference on Multimedia*, ACM, Berkeley, USA, November, 2003, pp. 2–10.
- [10] O. Barnich, M. Van Droogenbroeck, ViBe: a powerful random technique to estimate the background in video sequences, in: *International Conference on Acoustics, Speech, and Signal Processing*, April, 2009, pp. 945–948.
- [11] O. Barnich, M. Van Droogenbroeck, ViBe: a universal background subtraction algorithm for video sequences, *IEEE Trans. Image Process.* 20 (6) (2011) 1709–1724.