

# An Improved Moving Object Detection Algorithm Based on Gaussian Mixture Models

Xuegang Hu, Jiamin Zheng\*

Chongqing University of Posts and Telecommunications, Chongqing, China  
Email: \*zhengjm3211@126.com

Received 13 June 2016; accepted 24 July 2016; published 27 July 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Aiming at the problems that the classical Gaussian mixture model is unable to detect the complete moving object, and is sensitive to the light mutation scenes and so on, an improved algorithm is proposed for moving object detection based on Gaussian mixture model and three-frame difference method. In the process of extracting the moving region, the improved three-frame difference method uses the dynamic segmentation threshold and edge detection technology, and it is first used to solve the problems such as the illumination mutation and the discontinuity of the target edge. Then, a new adaptive selection strategy of the number of Gaussian distributions is introduced to reduce the processing time and improve accuracy of detection. Finally, HSV color space is used to remove shadow regions, and the whole moving object is detected. Experimental results show that the proposed algorithm can detect moving objects in various situations effectively.

## Keywords

Moving Object Detection, Gaussian Mixture Model, Three-Frame Difference Method, Edge Detection, HSV Color Space

---

## 1. Introduction

The performance of moving object detection algorithm is very crucial in a sound surveillance system for reliable tracking and behavior recognition. In recent years, many scholars have successively proposed lots of classic moving object detection algorithms [1]. They mainly include the optical flow method, the frame difference method and the background subtraction method. The optical flow method uses the vector characteristics of the

---

\*Corresponding author.

moving object which changed over with time to detect moving objects in image sequences [2]. The method can be applied to dynamic scenes, but its calculation is complex and anti-disturbance ability is poor, which can't meet the requirements of real-time video processing. The frame difference method [3] uses the two adjacent frame subtraction to detect moving objects, the method is not sensitive to the change of light, and has good real-time performance. However, it's difficult to obtain complete contour of moving object, and prone to "ghost" and "holes" phenomenon when the speed of moving target is too fast or too slow. The background subtraction method builds first a background image, and then uses the subtraction of the current frame and background image to extract the moving objects [4]. The key of this method is that the background modeling and accuracy of the model. But the extracted background image is easy to be disturbed by light, weather and the other environmental conditions. Recently, the background modeling method based on statistical model has been developed for its good adaptability to complex scenes.

Stauffer *et al.* proposed Gaussian mixture model (GMM) [5] used for background modeling which had been found to cope reliably with slow illumination changes, repetitive motions from clutter, and long-term scene changes. However, this method suffers from many problems in the process of object detection. For example, it will judge the entire frame as foreground when suffering a sudden illumination change; the number of Gaussian distributions is constant in the background model; the result of moving object detection contains shadow region and contour of moving object is discontinuous. Many improved methods have been proposed to solve these problems. Ellis *et al.* used a method to calculate the ratio of the foreground region to deal with the interference of illumination change [6]. The method only has better performance for global illumination change rather than local illumination change. Zivkovic proposed an adaptive GMM algorithm with the maximum likelihood estimation [7], which adaptively chose the appropriate number of Gaussian distributions to model each pixel on-line. This method reduces the processing time and improves the accuracy of segmentation slightly. However, a priori negative coefficient introduced by artificiality may result in appearing unreasonable update to the weight of the Gaussian distribution during the update processing and affecting effect of the moving target detection. Zhiyan Bin *et al.* proposed a new effective algorithm which combines HSV color information with RGB normalized space and first-order gradient information for shadow removing in all kinds of complex scene [8]. The method can remove a portion of the shadow region, but the edge information of moving objects has been influenced, which brings lots of difficulties for subsequent processing.

Based on Gaussian mixture model and three-frame differencing method, we propose a moving object detection algorithm in this paper. The experiments show that our algorithm has improved in the aspects of accuracy, adaptability and real-time performance.

## 2. Gaussian Mixture Model

GMM algorithm considers the value of a particular pixel  $(x_0, y_0)$  over time as a "pixel process", which is a time series of pixel value. It presents the recent history of each pixel  $\{X_1, X_2, \dots, X_t\}$  by mixture of  $K$  Gaussian distributions, where  $X_t$  is the pixel value of  $(x_0, y_0)$  at time  $t$ . The probability of observing the current pixel value is

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \times \eta(X_t, \mu_{i,t}, M_{i,t}) \quad (1)$$

where  $\mu_{i,t}$  is the mean value of the  $i^{\text{th}}$  Gaussian in the mixture at time  $t$ ,  $\omega_{i,t}$  is an estimate of the weight of the  $i^{\text{th}}$  Gaussian in the mixture at time  $t$ ,  $M_{i,t}$  is the covariance matrix of the  $i^{\text{th}}$  Gaussian in the mixture at time  $t$ , and where  $\eta$  is a Gaussian probability density function

$$\eta(X_t, \mu_{i,t}, M_{i,t}) = \frac{1}{(2\pi)^{n/2} |M_{i,t}|^{1/2}} e^{-\frac{1}{2}(X_t - \mu_{i,t})^T M_{i,t}^{-1} (X_t - \mu_{i,t})} \quad (2)$$

where  $n$  is dimension of the  $X_t$ . For computational reasons, the covariance matrix is assumed to be of the form  $M_{i,t} = \sigma_{i,t}^2 I$ . This assumes that the red, green, and blue pixel values are independent and have the same variances, where  $I$  is the unit matrix,  $\sigma_{i,t}^2$  is the variance of the  $i^{\text{th}}$  Gaussian in the mixture at time  $t$ . Every new pixel value,  $X_t$ , is checked against the existing  $K$  Gaussian distributions, until a match is found. A match is defined as a pixel value within 2.5 standard deviations of a distribution. If none of the  $K$  distributions match the current pixel value, the least probable distribution is replaced with the current value as its mean value, an initially high

variance, and low prior weight. The prior weight of the  $K$  distributions at time  $t$ ,  $\omega_{i,t}$  are adjusted as follows

$$\omega_{i,t} = (1 - \alpha)\omega_{i,t-1} + \alpha R_{i,t} \quad (3)$$

where  $\alpha$  is the learning rate,  $R_{i,t}$  is 1 for the model which matched and 0 for the remaining models. The  $\mu$ ,  $\sigma$  for unmatched distributions remain the same. Parameters of the distribution which matches the new observation are updated as follows

$$\mu_{i,t} = (1 - \beta)\mu_{i,t-1} + \beta X_t \quad (4)$$

$$\sigma_{i,t}^2 = (1 - \beta)\sigma_{i,t-1}^2 + \beta(X_t - \mu_{i,t})^T(X_t - \mu_{i,t}) \quad (5)$$

where  $\beta$  is parameter learning rate, and  $\beta = \alpha/\omega_{i,t}$ . After this approximation, the weights are renormalized.

The Gaussian distributions are ordered by the value of  $P = \omega_{i,t}/\sigma_{i,t}$ . This ordering of the model is effectively an ordered, open-ended list, where the most likely background distributions remain on top and the less probable transient background distributions gravitate towards the bottom and are eventually replaced by new distributions. Then the first  $B$  distributions are chosen as the background model

$$B = \arg \min_b \left( \sum_{i=1}^b \omega_{i,t} \geq T \right) \quad (6)$$

where  $T$  is a measure of the minimum portion of the pixel that should be accounted for by the background. Each pixel which is matched with any of the former  $B$  Gaussian distributions will be marked as background pixel. Otherwise, it will be marked as foreground pixel.

### 3. Proposed Algorithm

This paper proposes a moving object detection algorithm which based on Gaussian mixture model and three-frame difference method. Firstly, we achieve the coarse segmentation of motion region which utilize the improved three-frame difference method. We make up the target boundary fracture portions by combining with edge detection technology, and overcome the impact of the sudden illumination change by adjusting the binary threshold value based on changes of the scenes. Secondly, in the process of mixture Gaussian background model, according to the change of model's weight decide to add model or remove unmatched model, the number of components of the mixture is constantly adapted for each pixel, which improves the ability of the algorithm to describe the scene. Finally, we utilize characteristic of HSV color space to eliminate the interference of the shadow region effectively.

#### 3.1. Improved Three-Frame Difference Method

The traditional three-frame difference method can quickly detect moving objects. But the algorithm has poor ability to deal with illumination change scenes because of the fixed segmentation threshold it utilized. In addition, holes phenomena will occur in object interior and part of the object contour is not continuous. To solve these problems, this paper proposes the improved of three-frame difference method. It combines three-frame difference method with canny edge detection [9] operator, which solves the problem of target edge discontinuity. What's more, with dynamic binary threshold value, it effectively adapt to the scene of illumination change. Steps of the improved three-frame difference method are shown as follows

- 1) Three frames, defined as  $f_{i-1}(x, y), f_i(x, y), f_{i+1}(x, y)$ , then we calculate the difference value of image between adjacent two frames respectively, binary image  $D_{i,i-1}(x, y)$  and  $D_{i,i+1}(x, y)$  are obtained by the segmentation threshold. They can be written as follows

$$D_{i,i-1}(x, y) = \begin{cases} 1, & |f_i(x, y) - f_{i-1}(x, y)| > T + \Delta Z_1 \\ 0, & \text{others} \end{cases} \quad (7)$$

$$D_{i,i+1}(x, y) = \begin{cases} 1, & |f_i(x, y) - f_{i+1}(x, y)| > T + \Delta Z_2 \\ 0, & \text{others} \end{cases} \quad (8)$$

where  $T$  is the fixed binary threshold,  $\Delta Z_1, \Delta Z_2$  are dynamic threshold. They are described as follows

$$\Delta Z_1 = \xi \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |f_i(x, y) - f_{i-1}(x, y)| \quad (9)$$

$$\Delta Z_2 = \xi \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |f_i(x, y) - f_{i+1}(x, y)| \quad (10)$$

where  $M \times N$  is the number of pixels of the image,  $\xi$  is inhibition coefficient. We can get binary image  $D_i(x, y)$  by doing and operation between  $D_{i,i+1}(x, y)$  and  $D_{i,i-1}(x, y)$ .

$$D_i(x, y) = \begin{cases} 1, D_{i,i+1}(x, y) \& \& D_{i,i-1}(x, y) = 1 \\ 0, \text{others} \end{cases} \quad (11)$$

2) We get edge image  $B_i(x, y)$  by using canny operator extract edge information of  $f_k$  frame image. Then we obtain binary image  $F_{i,i-1}(x, y), F_{i,i+1}(x, y)$  by doing and operation between edge image and  $D_{i,i-1}(x, y), D_{i,i+1}(x, y)$ . Finally, we get binary image  $G_i(x, y)$  by doing or operation between  $F_{i,i-1}(x, y)$  and  $F_{i,i+1}(x, y)$ .

$$G_i(x, y) = \begin{cases} 1, F_{i,i-1}(x, y) = 1 \parallel F_{i,i+1}(x, y) = 1 \\ 0, \text{others} \end{cases} \quad (12)$$

3) We obtain complete moving object region  $M_i(x, y)$  by doing or operation between  $D_i(x, y)$  and  $G_i(x, y)$ .

$$M_i(x, y) = \begin{cases} 1, D_i(x, y) = 1 \parallel G_i(x, y) = 1 \\ 0, \text{others} \end{cases} \quad (13)$$

When  $M_i(x, y) = 1$ , which is foreground pixel, and others are background pixels.

### 3.2. Improved Gaussian Mixture Model

Moving target detection algorithm based on Gaussian mixture model basically set the fixed number of Gaussian distributions for each pixel. In fact, when recently observed pixel values are roughly constant, all of the distributions approximate the same values. In such a case, only one distribution should exist and the other distributions are not necessary at all. On the contrary, if recently observed pixel values change frequently, a constant number of Gaussian is not always enough to estimate the background model, and which is very difficult to determine the appropriate number of Gaussians. Therefore, this paper proposes a new background estimation method, which can increase and decrease the number of distributions to handle the variations of each pixel. The update process of improved Gaussian mixture model is listed as follows:

1) Increment of distribution

Every new pixel value  $X_t$  is checked against the existing  $K$  Gaussian distributions, when none of the  $K$  distribution matches the current pixel value, if  $K < K_{\max}$ , we increase a new Gaussian distribution. Setting the Gaussian distribution of mean, variance, and weights are  $X_t, \sigma_{init}, \omega_{mit}$ ; if  $K = K_{\max}$ , the least probable distribution is replaced with the current value as its mean value, an initially high variance and low prior weight.

2) Decrement of distribution

In the process of Gaussian mixture model update, if a Gaussian distribution of the current pixel can't describe the background accurately, its weight will continue to decay according to  $\omega_{i,t} = (1 - \alpha)\omega_{i,t-1}$ . Therefore, we can set an initial weight  $\omega_{mit}$ , when a Gaussian distribution satisfies the formula (13). Which indicate that the Gaussian distribution can't describe background well. And this distribution will affect the convergence rate of the model, therefore, we should delete it.

$$\omega_{i,t} < \omega_{mit} \& \& \frac{\omega_{i,t}}{\sigma_{i,t}} < \frac{\omega_{mit}}{\sigma_{mit}} \quad (14)$$

3) The integrated Gaussian distribution

When the difference between means of two Gaussians (the one is  $\varphi_a$  and the other is  $\varphi_b$ ) is smaller than a threshold, these distributions are integrated into one Gaussian. The integrated Gaussian  $\varphi_c$  is calculated as follow

$$\omega_{c,t} = \omega_{a,t} + \omega_{b,t} \quad (15)$$

$$\mu_{c,t} = \frac{\omega_{a,t}\mu_{a,t} + \omega_{b,t}\mu_{b,t}}{\omega_{a,t} + \omega_{b,t}} \quad (16)$$

$$\sigma_{c,t} = \frac{\sigma_{a,t}\omega_{a,t} + \sigma_{b,t}\omega_{b,t}}{\omega_{a,t} + \omega_{b,t}}. \quad (17)$$

### 3.3. HSV Color Space Shadow Removal

Shadow is the dark area in which the light source can't directly irradiate to the surface of the object, and it can be divided into two types: shadows of the object and the cast shadows [10]. Shadows cause serious problems while segmenting and extracting moving objects due to the misclassification of shadow points as foreground. Usually, we are interested only in the objects and the pixels corresponding to the shadow should be detected. This algorithm works in Hue-Saturation-Value (HSV) [11] color space. We analyze pixels in HSV color space, the main reasons that HSV color space corresponds closely to the human perception of color and it has revealed more accuracy in distinguishing shadows. In fact, a shadow cast on a background does not change its hue and saturation significantly. However, whether in color or grayscale image, Value component always reflect the useful information of the image. According to this characteristic, this paper proposes a new method of shadow detection. The resulting decision process is reported in the following equation:

$$SP_i(x, y) = \begin{cases} 1, & 0 \leq \frac{I_i^V(x, y)}{B_i^V(x, y)} \leq 1 \ \& \ \& \ |I_i^V(x, y) - I_{i-1}^V(x, y)| \leq 1 \\ 0, & \text{others} \end{cases} \quad (18)$$

where  $I_i^V(x, y)$  is the Value component at coordinate  $(x, y)$  in the input image (frame  $i$ ) and  $B_i^V(x, y)$  is the Values component at coordinate  $(x, y)$  in the background model,  $I_{i-1}^V(x, y)$  is the value of the Value component of the previous frame. For each pixel belonging to the objects resulting from the segmentation step, we check if it is a shadow according to the formula (17). If  $SP_i(x, y) = 1$ , the point is shadow point, others are moving object points. Compared with previous algorithm which judges shadow region just by using change of the value of current frame and background frame, this algorithm increases the change of value between current frame and previous frame, which improves the accuracy of the shadow judging. Moreover, we set the fixed ratio threshold of Value component between the input image and the background model, which improve the processing speed of the algorithm.

## 4. Experimental Results and Analysis

In this paper, all experiments are performed on windows 7 system with 3.00 GHz Core 4 processor and 4.00 GB of memory. In order to analyze the robustness and effectiveness of the proposed method, four experiments under different conditions are demonstrated in our paper. The results are compared with GMM algorithm and the algorithm [12] in subjective visual and objective parameters statistics.

**Figure 1** shows results of the algorithm in situation of single object and multiple objects. GMM algorithm can't extract complete object region, and it isn't sensitivity to low-speed object. Literature [12] and our algorithm can detect the complete moving targets, but literature [12] still exist holes phenomenon. However, our algorithm is preferably made up for this shortcoming, and the detection sensitivity to low-speed moving object has improved a lot in the meanwhile.

**Figure 2** shows the detection results in shadow scenes. Due to the interference of noise and shadows, the detection result of GMM algorithm isn't accurate. The algorithm [12] is only able to remove a portion of shadows. But our method can eliminate the influences of noise and shadow, and retain complete edge and internal information of moving targets.

**Figure 3** shows detection results under the illumination change scenes. GMM algorithm can hardly recognize the moving objects, and it exist false detection objects in the result. Literature [12] is able to detect moving targets. However, as it is affected by illumination mutation, the information of edges and internal details has been seriously damaged. What's worse, some of small objects can't be detected. Compared with it, our algorithm can

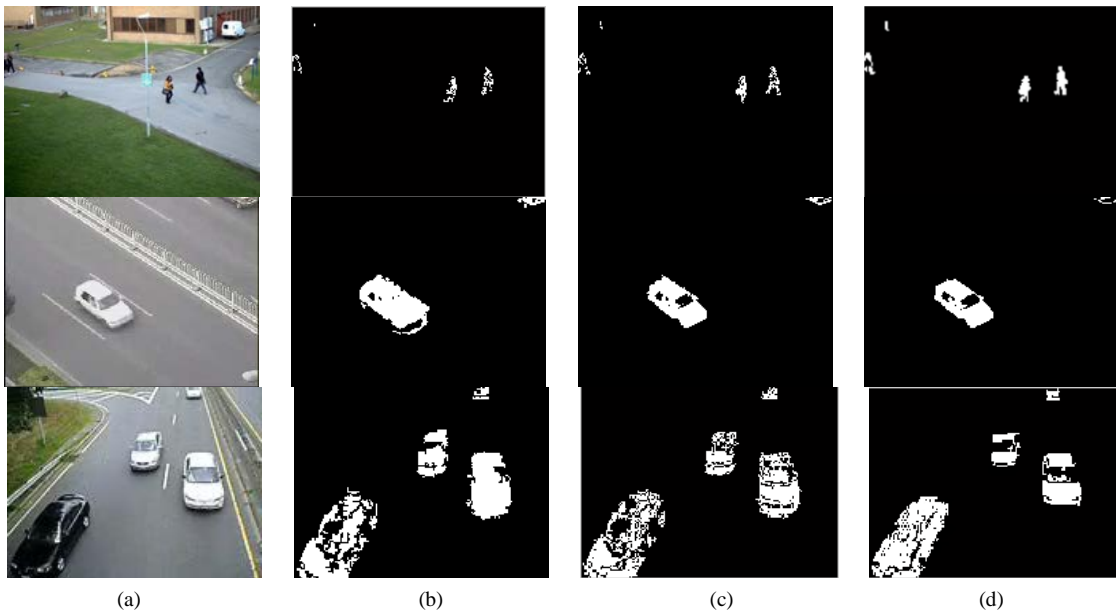


Figure 1. Single and multiple object. (a) The original image; (b) GMM algorithm; (c) Algorithm [12]; (d) Our algorithm.

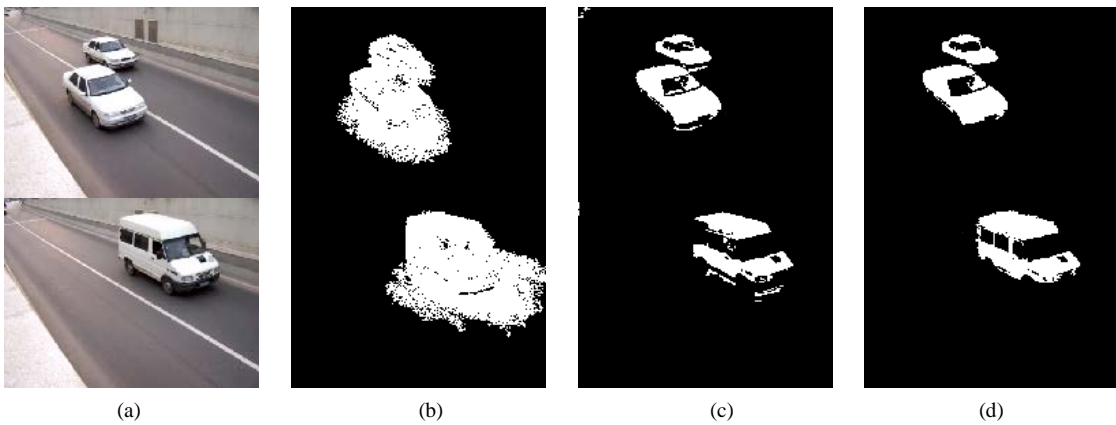


Figure 2. The result of shadow scenes. (a) The original image; (b) GMM algorithm; (c) Algorithm [12]; (d) Our algorithm.

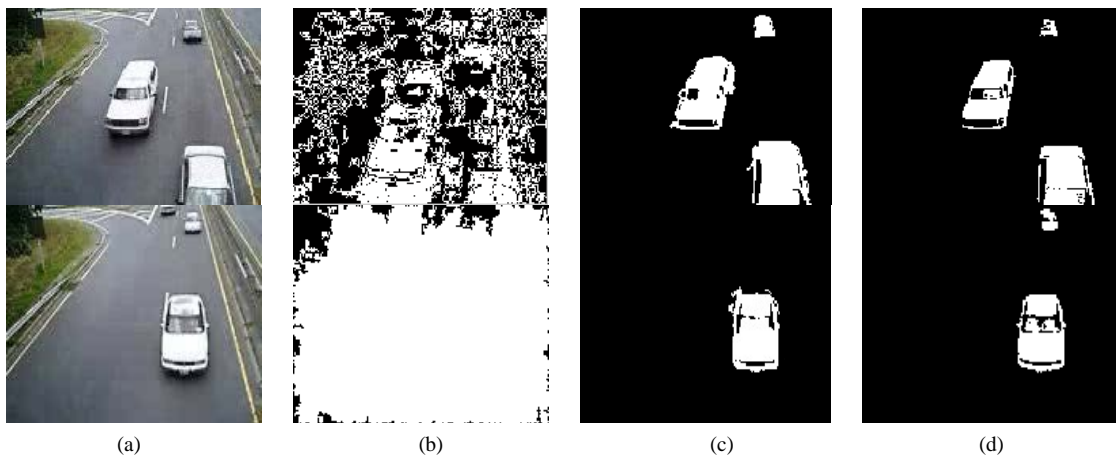


Figure 3. The result of the illumination change scene. (a) The original image; (b) GMM algorithm; (c) Algorithm [12]; (d) Our algorithm.



**Table 1.** The running time for single frame.

Algorithms	GMM algorithm	Algorithm [12]	Our algorithm
Running time(ms/f)	82	58	50

**Table 2.** A comparison of performance of various algorithms (%).

videos	GMM algorithm		Algorithm [12]		Our algorithm	
	Recall	Precision	Recall	Precision	Recall	Precision
a	84.73	82.52	91.55	90.32	93.25	91.33
b	85.31	83.72	90.26	87.67	94.15	89.81
c	80.42	75.74	86.58	85.45	89.37	87.51
d	72.43	70.58	87.49	86.56	90.82	89.71
e	74.93	72.25	88.47	85.52	91.55	90.12
f	90.34	40.55	84.15	82.37	92.82	91.54
g	96.65	32.23	88.72	86.38	94.18	92.39

detect the moving objects accurately and completely.

**Table 1** presents the processing speed of various algorithms. The image sequence is sampled to  $320 \times 240$ . As we can see, the proposed method has better real-time performance than the algorithm [12] and GMM algorithm. The improved three-frame difference method is used to quickly detect the motion regions, and it can avoid matching detection of pixels which belong to the background. In addition, we reduce the computational complexity of the algorithm by automatically choosing the number of components for each pixel.

In order to systematically evaluate various algorithms, it is useful to identify the following two important quality measures: *Recall* (also known as detection rate) and *Precision* (also known as positive prediction) [13]. They are written as follows

$$Recall = \frac{t_p}{t_p + f_n} \quad (19)$$

$$Precision = \frac{t_p}{t_p + f_p} \quad (20)$$

where  $t_p$  is the total number of true positives,  $f_n$  is the total number of false negatives,  $f_p$  is the total number of false positives. *Precision* reflects the false detection rate, and *Recall* reflects the accuracy of detection result.

We calculate *Recall* and *Precision* of various algorithms, and results are reported in **Table 2**. The ground truth for each frame is necessary, and we obtain it by segmenting the images with an accurate manual classification of points in foreground and background regions. From the table data, GMM algorithm is easily affected with the scenes change. The proposed algorithm reduces the interference of environmental impact and removes shadow regions. Compared with the algorithm [12], our algorithm has improved the rates of *Precision* and *Recall*.

## 5. Conclusion

This paper proposes an improved moving object detection algorithm based on adaptive Gaussian mixture model and three-frame difference method. The proposed algorithm can automatically select the number of components for each pixel. This modification dramatically improves the convergence and the accuracy of background subtraction whilst maintaining the same temporal adaptability. The three-frame difference method uses adaptive segmentation threshold that can adapt to illumination change scene. In addition, we preserve the complete edge information of moving object by using edge detection technology, and effectively remove the shadow by using HSV color space. Comprehensive analysis of experimental results shows that our algorithm can detect moving objects in the complex scenes effectively and has good robustness.

## References

- [1] Bouwmans, T. (2014) Traditional and Recent Approaches in Background Modeling for Foreground Detection: An

- Overview. *Computer Science Review*, **11-12**, 31-66.
- [2] Senst, T., Evangelio, R.H. and Sikora, T. (2011) Detecting People Carrying Objects Based on an Optical Flow Motion Model. 2011 *IEEE Workshop on Applications of Computer Vision (WACV)*, Kona, HI, 5-7 January 2011, 301-306.
- [3] Haritaoglu, I., Harwood, D. and Davis, L.S. (2000) W4: Real-Time Surveillance of People and Their Activities. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **22**, 809-830. <http://dx.doi.org/10.1109/34.868683>
- [4] Chiu, C.C., Ku, M.Y. and Liang, L.W. (2010) A Robust Object Segmentation System Using a Probability-Based Background Extraction Algorithm. *IEEE Transactions on Circuits & Systems for Video Technology*, **20**, 518-528. <http://dx.doi.org/10.1109/TCSVT.2009.2035843>
- [5] Stauffer, C. and Grimson, W.E.L. (1999) Adaptive Background Mixture Models for Real-Time Tracking. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Fort Collins, CO, 23-25 June 1999, 2246. <http://dx.doi.org/10.1109/cvpr.1999.784637>
- [6] Chen, Z. and Ellis, T. (2011) Self-Adaptive Gaussian Mixture Model for Urban Traffic Monitoring System. 2011 *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, Barcelona, 6-13 November 2011, 1769-1776.
- [7] Zivkovic, Z. (2004) Improved Adaptive Gaussian Mixture Model for Background Subtraction. *Proceedings of the 17th IEEE International Conference on Pattern Recognition (ICPR)*, Cambridge, 23-26 August 2004, 28-31.
- [8] Bin, Z. and Liu, Y. (2010) Robust Moving Object Detection and Shadow Removing Based on Improved Gaussian Model and Gradient Information. 2010 *IEEE International Conference on Multimedia Technology (ICMT)*, Ningbo, 29-31 October 2010, 1-5.
- [9] Canny, J. (1986) A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 679-698.
- [10] Porikli, F. and Thornton, J. (2005) Shadow Flow: A Recursive Method to Learn Moving Cast Shadows. *Tenth IEEE International Conference on Computer Vision (ICCV)*, Beijing, 17-21 October 2005, 891-898.
- [11] Buch, N., Velastin, S.A. and Orwell, J. (2011) A Review of Computer Vision Techniques for the Analysis of Urban Traffic. *IEEE Transactions on Intelligent Transportation Systems*, **12**, 920-939. <http://dx.doi.org/10.1109/TITS.2011.2119372>
- [12] Xu, X. and Jiang, D.B. (2013) Optimization of Gaussian Mixture Model Motion Detection Algorithm. *Application Research of Computers*, **30**, 2190-2194.
- [13] Lucia, M. and Alfredo, P. (2008) A Self-Organizing Approach to Background Subtraction for Visual Surveillance Applications. *IEEE Transactions on Image Processing*, **17**, 1168-1177.



Scientific Research Publishing

---

**Submit or recommend next manuscript to SCIRP and we will provide best service for you:**

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>