
*British Journal of Applied Science & Technology*
*14(6): 1-8, 2016, Article no.BJAST.24249*
*ISSN: 2231-0843, NLM ID: 101664541*

**SCIENCEDOMAIN** *international*
www.sciencedomain.org

# Moving Object Detection Algorithm Based on Gaussian Mixture Model and HSV Space

**Xuegang Hu[1] and Cheng He[1*]**

[1]*Chongqing University of Posts and Telecommunications, Chongqing, China.*

***Authors' contributions***

*This work was carried out in collaboration between both authors. Authors XH and CH designed the study, performed the statistical analysis, wrote the protocol, wrote the first draft of the manuscript and managed literature searches. Author CH managed the analyses of the study and literature searches. Both authors read and approved the final manuscript.*

***Article Information***

DOI: 10.9734/BJAST/2016/24249
Editor(s):
(1) Samir Kumar Bandyopadhyay, Department of Computer Science and Engineering, University of Calcutta, India.
Reviewers:
(1) V. Usha Shree, Vidya Jyothi Institute of Technology (Autonomous), Hyderabad, India.
(2) Jen-Yu Shieh, National Formosa University, Taiwan.
(3) Anonymous, Ohio University, USA.
Complete Peer review History: http://sciencedomain.org/review-history/13569

***Method Article***

***Received 10th January 2016***
***Accepted 18th February 2016***
***Published 5th March 2016***

## ABSTRACT

Aiming at the traditional Gaussian mixture model has poor adaptability to the complex scenes, we proposes an improved moving object detection algorithm based on Gaussian mixture model and HSV space. The motion region is first extracted by the improved three-frame difference method. With the matching results, region segmentation of current frame is realized. Then different regions adopt different update strategy that improves the ability to reflect the illumination and scenes change. Next, utilizing characteristics of HSV color space and image first-order gradient achieve shadow detection. It effectively reduces interference of shadows, especially the pixels of foreground which has similar brightness properties with background. Experimental results show that the algorithm has good robustness and real-time performance.

*Keywords: Gaussian mixture model; moving object detection; three-frame difference method; HSV space; shadow detection.*

---

**Corresponding author: E-mail: hc01357@163.com;*

## 1. INTRODUCTION

Moving object detection is an important research direction in intelligent video processing. It's also the basis of object identification, tracking and analysis [1]. The detection results directly impact on the effectiveness of monitoring system. At present, optical flow method, frame difference method and background subtraction method are commonly used to detect moving objects [2-4]. Optical flow method is to estimate motion field of image and to merge the similar motion vectors of the image. It can be applied to dynamic scenes. Whereas this approach is sensitive to illumination change and the calculation is complex, so it is difficult to meet the requirement of real-time detection. The frame difference method extracts the motion region by calculating the difference between two adjacent frames and image threshold. It's simple and easy to implement. But holes phenomenon frequently occur in the process, consequently detected result is inaccurate. Background subtraction method uses the difference between current image and background image to extract the moving objects. The focus is to establish accurate background image and update the image in real-time. Recently, pixel-based statistical approach for background modeling has been greatly developed for its good adaptability to complex scenes.

Stauffer et al. [5] proposed the Gaussian mixture model (GMM) it creates *K* Gaussian distributions for each pixel of the image to describe its characteristics. With the analysis of model parameters, it can determine that the pixel belongs to the background or foreground. Therefore, this approach can adapt to part of scenes change, such as light dimmed, branches shake. In the case of object moves slowly and illumination mutation, the detected foreground will appear large false detection rate, which affects the integrity of the segmentation result [6,7]. To solve these problems, many researchers have presented many improved algorithms. In [8,9], each pixel adaptively selects the number of Gaussian distributions and different regions adopt different update strategies. It can effectively adapt to the observed scenes and eliminate the noise interference. In [10], it re-initialized the background image on the condition of illumination mutation, which avoided introducing noise. However, it can't deal with the shadow due to the light blocking. In [11-13], the approach used the image gradient and texture features to detect shadows, but it can't guarantee stable convergence states for foreground, especially the pixel values are closed to the background.

In this paper, a motion detection algorithm based on GMM and HSV space is introduced, which could handle the scenes variation. The proposed algorithm is explained in detail in Section 2. Detection result are shown and compared in Section 3, while conclusions are drawn in Section 4.

## 2. PROPOSED ALGORITHM

In this paper, the proposed algorithm combines GMM algorithm and shadow removal in HSV space. First of all, the algorithm introduces the principle of regional block for background modeling. Combining three-frame differencing and neighborhood information achieves the coarse segmentation of motion region and overcomes the holes phenomenon of frame differencing. The next step, the matching detection results determine the background exposure region. In the process of background updating, different regions adopt a different update strategy, which improves the adaptability to the scenes change. Last, using background subtraction extracts foreground. In addition, shadow removal algorithm combines the HSV space and image first-order gradient that effectively reduce false detection rate. The algorithm flow chart is shown in Fig. 1.

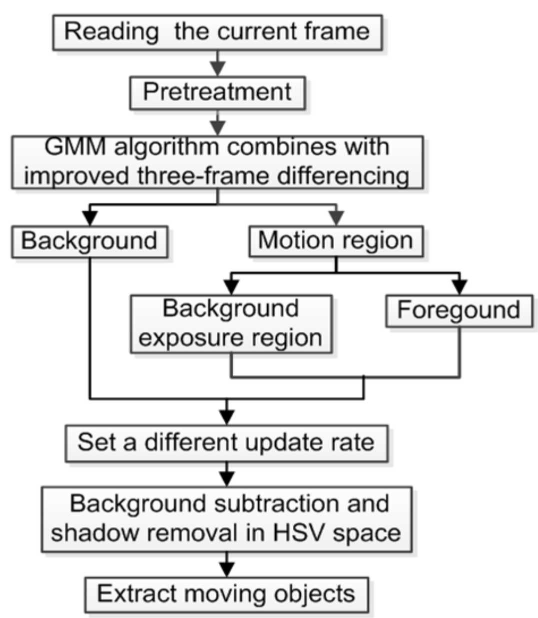



**Fig. 1. Algorithm flow chart**

### 2.1 Gaussian Mixture Model

GMM algorithm uses a mixture of $K$ Gaussian components to simulate the background image. At time $t$, the probability of current pixel value $x_t$ can be written as:

$$P(x_t)=\sum_{i=1}^{K} w_{i,t}\times\frac{1}{(2\pi)^{n/2}|\varphi_{i,t}|}e^{-\frac{1}{2}(x_t-\mu_{i,t})^{\mathrm{T}}\sum_{i,t}^{-1}(x_t-\mu_{i,t})} \quad (1)$$

where $K$ is the number of distributions, generally taking 3 to 5, $w_{i,t}$ is an estimate of the weight of the $i^{th}$ Gaussian distribution at time $t$, $\mu_{i,t}$ is the mean value of the $i^{th}$ Gaussian distribution at time $t$, $\varphi_{i,t}$ is the covariance matrix of the $i^{th}$ Gaussian distribution at time $t$, it meets $\varphi_{i,t}=\sigma_{i,t}^2 I$ , and the sum of $K$ weight value is 1. The process of GMM algorithm is as follows:

1) Model initialization. Reading the first frame image and regarding the pixel values as the mean parameter $\mu$ of the first background image, and initialized with a large standard deviation $\sigma_0$ and small weight $w_0$.
2) Updating the model parameters. Each pixel value $x_t$ is checked against the existing $K$ Gaussian distributions according to the priority of $\rho_{i,t}$. A match is defined as pixel value within 2.5 standard deviation of distribution. At time $t$, weights of the $K$ distributions are updated as follows:

$$w_{\mathrm{i},t}=(1-\alpha)*w_{\mathrm{i},t-1}+\alpha*M_{\mathrm{i},t} \quad (2)$$

where $\alpha$ is the learning rate and $0\leq \alpha \leq 1$. The value of $\alpha$ determines the speed of background update. $M_{i,t}$ is 1 for the model which matched and 0 for remaining models. After this approximation, the weight of each Gauss distribution is renormalized. The $\mu$ and $\sigma$ parameters for unmatched distributions remain the same. The parameters of the distribution which matched the new observation are updated as follows:

$$\mu_{i,t}=(1-\rho)\mu_{i,t-1}+\rho*x_t \quad (3)$$

$$\sigma_{i,t}{}^2=(1-\rho)*\sigma_{\mathrm{i},t-1}{}^2+\rho*(x_t-\mu_{i,t-1})^{\mathrm{T}}(x_t-\mu_{i,t-1}) \quad (4)$$

where $\rho$ is the learning rate, and $\rho=\alpha/w_{i,t}$. If none of the $K$ Gaussian distributions match the current pixel value, the least probable distribution is replaced by a new distribution using the current value as its mean and initialized with a large standard deviation $\sigma_0$ and small weight $w_0$.
3) Background estimation. The Gaussian distributions are arranged according to the priority of $\rho_{i,t}$ from highest to lowest. The former $B$ distributions are chosen for the background estimation, where $T$ is a threshold measuring of the minimum portion of data that should be accounted for. The value of $T$ is 0.75. It is described as follow:

$$B=\arg\min_{b}\left(\sum_{k=1}^{b} w_k > T\right) \quad (5)$$

Each pixel which matched with any of the former $B$ Gaussian distributions will be marked as background. Otherwise, it will be marked as foreground. Finally, we can extract the moving objects with background subtraction.

### 2.2 Improved Three-frame Difference Method

In the process of background modeling, each pixel is considered as a random variable and made matching detection independently. This approach ensures that each pixel can real-time update its parameters. However, the stability of pixel values in the background is not used. The traversal of all pixels produces redundancy and increase the running time. Therefore, the principle of regional block is used. By extracting the motion region firstly, it is useful to improve the efficiency of background modeling.

The three-frame difference method is easy to implement, it can quickly detect the motion region. However, it will appear holes phenomenon in the moving objects interior. In order to eliminate the impact, the improved algorithm combines three-frame differencing with changed neighborhood information. Three-frame differencing choose three consecutive frames $I_{i-1}$, $I_i$, $I_{i+1}$ in image sequences and respectively calculate the difference value image between two frames. Binary image is obtained by the segmentation of $T_1$. Next, the difference binary images do “and” operation that get result $P_i$. It can be written as follows:

$$D_{i,i-1}(x,y)=\begin{cases}1, |I_i(x,y)-I_{i-1}(x,y)|>T_1\\ 0, others.\end{cases} \tag{6}$$

$$D_{i+1,i}(x,y)=\begin{cases}1, |I_{i+1}(x,y)-I_i(x,y)|>T_1\\ 0, others\end{cases} \tag{7}$$

$$P_i(x,y)=\begin{cases}1, D_{i,i-1}(x,y)=1 \, and \, D_{i+1,i}(x,y)=1\\ 0, others.\end{cases} \tag{8}$$

Threshold of $T_1$ is very important. The value is set based on the principle of Gaussian distribution. In the frame difference result, the variation of pixel value in the foreground is not meet the Gaussian distribution. On the contrary, pixel of background changes slightly and is affected by the noise, so it can meet the Gaussian distribution. The $3\sigma$ criterion is used, and $T_1=3\sigma+\mu$. In another aspect, with the variations of frame difference result $D_{i,i-1}(x,y)$, the addition of its eight neighborhood values is used to make up the internal holes. The result $J_{i,i-1}(x,y)$ is described as follow:

$$\begin{aligned}J_{i,i-1}(x,y) = \; & D_{i,i-1}(x+1,y+1)+D_{i,i-1}(x+1,y) + \\ & D_{i,i-1}(x+1,y-1)+D_{i,i-1}(x,y+1)+D_{i,i-1}(x,y-1)+ \\ & D_{i,i-1}(x-1,y+1)+D_{i,i-1}(x-1,y)+D_{i,i-1}(x-1,y-1)\end{aligned} \tag{9}$$

Binary image was obtained with threshold of $T_2$, and $T_2=6.5T_1$. The result and $P_i(x,y)$ do "or" operation as follow:

$$S_i(x,y)=P_i(x,y) \,|\, J_{i,i-1}(x,y)>T_2 \tag{10}$$

Finally, $S_i(x,y)$ does the "fill" operation to get motion region $A_c$ completely.

## 2.3 Background Extraction

Due to the outside illumination variation, noise interference and others factors, the GMM algorithm with fixed update rate is difficult to adapt to the complex scenes. It frequently brings the problem that the background image updating is too fast or too slow, which introduces the background noise interference. For the above problem, the improved algorithm takes a differential update strategy. After the update is complete, the extracted background image can be more close to the real background.

The focus of this process is how to distinguish the region property for each pixel. Based on the principle of regional block, pixels not belong to the motion region are identified as the background point. Pixels of motion region are matched with the $K$ Gaussian distributions according to the priority from highest to lowest. The matching condition is defined as a pixel value within 2.5 standard deviation of distribution. The matching classification for any pixel is as follows:

$$\begin{cases}|x_t(x,y)-\mu_{i,t-1}(x,y)|\le 2.5\sigma_{i,t-1}, \; x_t(x,y)\in A_u\\ |x_t(x,y)-\mu_{i,t-1}(x,y)|> 2.5\sigma_{i,t-1}, \; x_t(x,y)\in A_m\end{cases} \tag{11}$$

where $0<i<K$, $\mu_{i,t-1}$ is the mean value of the $i^{th}$ Gaussian distribution at time $t$. $\sigma_{i,t-1}$ is the standard deviation of the $i^{th}$ Gaussian distribution at time $t$. If the matching is success, the pixel belongs to the background exposure region ($A_u$). If none of the $K$ distributions match the current pixel value, the pixel belongs to the foreground ($A_m$).

The pixel value of the background is relatively stable, and the changed amplitude is small. Therefore, background sets minor update rate, maintain the stability of the background update and avoid introducing noise. Foreground uses smaller update rate, reducing the influence of moving object to the background image. On the contrary, the background exposure region should speed up its update rate. With the large update rate, the Gaussian distributions obtain large mean value and small variance, which promote the corresponding priority. In particularly, the matched distributions are more possible to describe the real background image. In addition, when the vehicle starts movement from static, it will appear "double" shadow phenomenon in the stayed region. That represent pixel in the parking area is false detected as foreground. So the large update rate is needed to restore the exposure background region, which can deal with "double" shadow phenomenon. The current frame is separated into background exposure region $A_u$, foreground $A_m$ and background $A_s$, the update rate for different regions is $\alpha_u$, $\alpha_m$ and $\alpha_c$. It is described as follow:

$$\alpha(x,y)=\begin{cases}\alpha_m \quad (x_t \in A_m)\\ \alpha_c \quad (x_t \in A_c)\\ \alpha_u \quad (x_t \in A_u)\end{cases}, \alpha_m < \alpha_c < \alpha_u \tag{12}$$

In this paper, $\alpha_m$=0.005, $\alpha_c$=0.01 and $\alpha_u$=0.05. After updating, all Gaussian components will be arranged to the priority of $\rho_{i,t}$ from highest to lowest. According to formula (5), we can extract the background image.

### 2.4 Shadow Detection

Shadow detection remains an extremely challenging problem. The detected foreground contains portion of shadows due to the light blocking. In HSV space, the three independent components of brightness, hue and saturation will change following the illumination variation. Comparing the variation of same pixel under the condition of light blocking with normal illumination, there is a result that brightness has changed obviously, saturation a little, hue, however, hardly ever. Therefore, we use the characteristic of the variation to detect shadows. The formula can be described as follow:

$$S_h(x,y)=\begin{cases}1, & \alpha \le \dfrac{I_{V(}x,y)}{B_V(x,y)} \le \beta \ \&\& \\ & |I_S(x,y)-B_S(x,y)| \le T_s \ \&\& \\ & |I_H(x,y)-B_H(x,y)| \le T_H \\ 0, & others.\end{cases} \tag{13}$$

where $I$ is current frame, $B$ is the background frame, $H$, $S$, $V$ represent the independent component of brightness, hue and saturation in $HSV$ space, $\alpha$ and $\beta$ are set based on the light intensity, actually $0<\alpha<\beta<1$. $T_S$ and $T_H$, small constant, are thresholds for the hue and saturation. In this paper, $T_S$ =0.1, $T_H$ =0.15.

The above approach can effectively remove part of shadows. If we adjust to expand the range of $\alpha$, $\beta$, it could remove most of shadows. At the same time, however, it can detect part of object pixels as shadow. Especially the foreground object has similar brightness properties with the background. To solve previous question, the shadow removal algorithm combines HSV space with image gradient feature to detect the shadows. Direction of the gradient of foreground pixel is invariant between illumination change before and after, so using "*sobel*" operator to detect the gradient value of each pixel in the Horizontal $x$ and vertical $y$ directions, which represents as $I_{ix}(x,y)$ and $I_{iy}(x,y)$ ($i=H$, $S$, $V$). In $x$ and $y$ directions, the gradient detection operator $\Delta_x f$ and $\Delta_y f$ are described as follows:

$$\Delta_x f=\begin{bmatrix}1 & 2 & 1\\ 0 & 0 & 0\\ -1 & -2 & -1\end{bmatrix} \text{ and } \Delta_y f=\begin{bmatrix}1 & 0 & -1\\ 2 & 0 & -2\\ 1 & 0 & -1\end{bmatrix}$$

The gradient value of pixel in the background image is expressed as $B_{ix}(x,y)$ and $B_{iy}(x,y)$. As the difference of image gradient information, the pixel whether belongs to the foreground or background is determined as follows:

$$S_t(x,y)=\begin{cases}1, & \sum\limits_{i=H,S,V}\left(\sqrt{(I_{ix}(x,y)-B_{ix}(x,y))^2+(I_{iy}(x,y)-B_{iy}(x,y))^2}-3T_\sigma\right)>0\\ 0, & others\end{cases} \tag{14}$$

where $T_\sigma$ is a threshold and depends on the scenes. When $S_t(x,y)$ is 1, the pixel is a foreground point, otherwise, it belongs to shadow.

## 3. EXPERIMENTAL RESULTS AND ANALYSIS

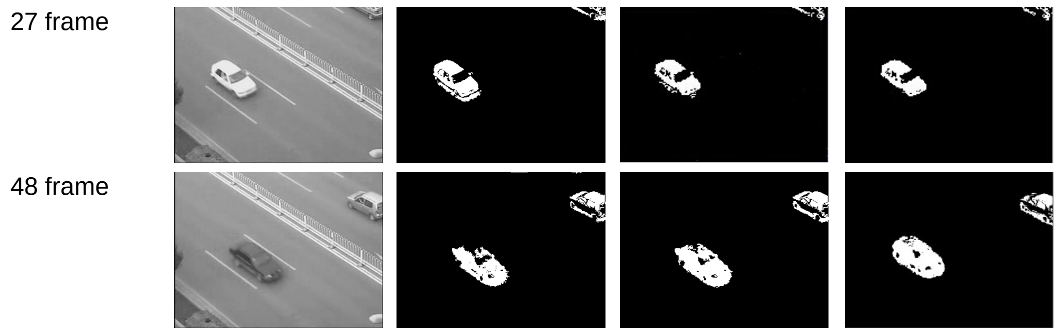
In this paper, we use vs2010 as the development tool and choose the computer configuration for the Intel Celeron E3400 2.6GHz, 3GB of memory, windows 7 operating system. Experimental test videos are on the condition of normal light and dim light, respectively corresponding to video 1,3 and video 2. The results are compared with GMM algorithm and the algorithm in [14] in subjective visual and objective parameters statistics.

From Fig. 2, Fig. 3 and Fig. 4, the test results show that GMM algorithm has poor adaptability with the fixed update rate. With different update rate for different regions, the proposed algorithm has a strong ability to quickly reflect scenes change. What's more, it effectively reduces the noise interference and avoid "double" shadow phenomenon. The algorithm [14] can only remove part of shadows. In contrast, by selecting appropriate brightness ratio and combining with the first-order gradient, the proposed algorithm can remove most of the shadows and decrease the influence in the pixels of object region, which has similar brightness properties with the background. Detection results show that the robustness has been improved in complex scenes.

**Table 1. Compare the running time for a single frame**

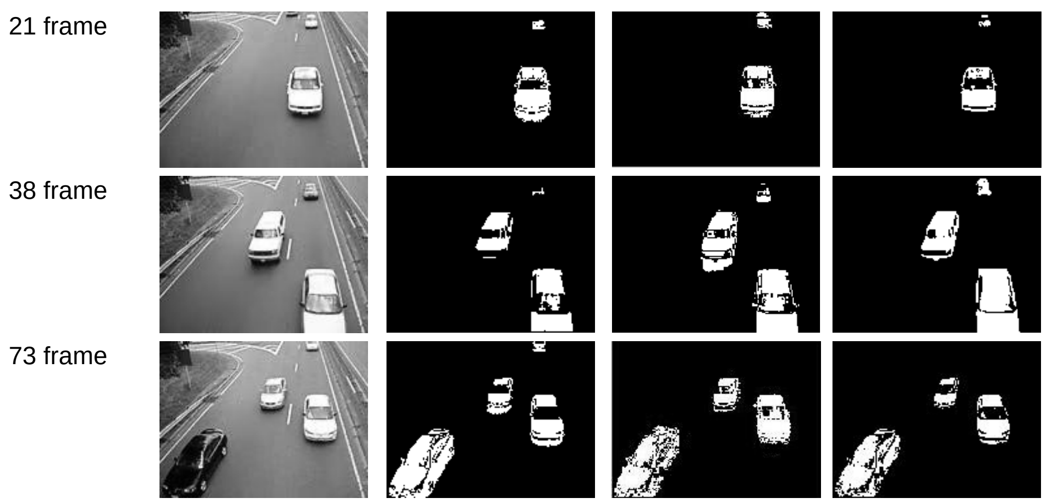
| Algorithms | GMM algorithm | Algorithm [14] | The proposed algorithm |
|---|---|---|---|
| running time (ms/f) | 157 | 75 | 66 |

Table 1 shows the real-time comparison of various algorithms, statistics results are tested on videos with a resolution of $320 \times 240$ pixels. The running time for a single frame of the proposed algorithm is lower than GMM algorithm. It has better real-time performance than the Algorithm [14]. The reason why time is lower is that the principle of regional block is introduced. By extracting the motion region, the matching detection of pixels of background region is avoided. As a result, the computation time was decreased.



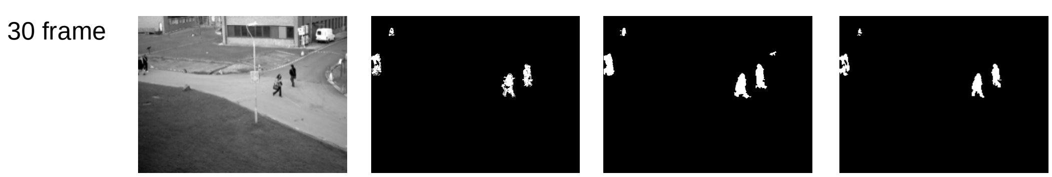

**(a) The original image (b) GMM algorithm (c) Algorithm [14] (d) The proposed algorithm**

**Fig. 2. Test results of Video 1**




**(a)The original image (b) GMM algorithm (c) Algorithm [14] (d) The proposed algorithm**

**Fig. 3. Test results of video 2**




**(a) The original image (b) GMM algorithm (c) Algorithm [14] (d) The proposed algorithm**

**Fig. 4. Test results of video 3**

**Table 2. detect results of precision and recall rate for each algorithm (%)**

| Videos | GMM algorithm | | Algorithm [14] | | The proposed algorithm | |
|---|---|---|---|---|---|---|
| | Precision | Recall | Precision | Recall | Precision | Recall |
| a | 82.43 | 84.32 | 85.31 | 91.24 | 86.46 | 93.03 |
| b | 79.28 | 85.42 | 84.36 | 88.16 | 90.73 | 93.55 |
| c | 78.17 | 81.37 | 85.62 | 90.43 | 85.74 | 92.62 |
| d | 80.35 | 80.23 | 86.72 | 91.24 | 92.53 | 96.34 |
| e | 75.71 | 82.26 | 87.29 | 88.72 | 88.71 | 89.36 |
| f | 82.56 | 86.35 | 90.62 | 92.28 | 93.46 | 94.73 |

In order to quantitatively evaluate the detection results for various algorithms, the method of literature [15] was used to count precision and recall, defined as follows:

$$precision = \frac{t_p}{f_p + t_p}, \qquad recall = \frac{t_p}{t_p + f_n},$$

where $t_p$ represents the number of pixels of foreground correctly detect, $f_n$ represents the number of pixels of foreground false detect as background, $f_p$ represents the number of pixels of background false detect as foreground. Precision reflects the false detection rate, and recall reflects the accuracy of detection result. Table 2 represents statistical parameters of various algorithms with different test videos. Among them, the real segmentation of foreground objects is obtained by artificial. From the table data, GMM algorithm based on the fixed update rate is easily affected by the scenes change. When the background image was deeply interfered, the recall rate is much smaller than other algorithms. The proposed algorithm reduces the interference of environmental impact and removes shadows. Compared with the algorithm [14], the precision and recall rates have certain improvement.

## 4. CONCLUSIONS

On the basis of in-depth study of background modeling and shadow suppression method, we propose an algorithm combining GMM and shadow removal in HSV space. This approach first extracts the motion region and adopts a differentiated background update strategy, which decreases the number of pixels in the matching detection, and improves the accuracy of background estimation. When pixel value of foreground is closed to the background, image first-order gradient is used, which reduces the false detection rate. Experimental results show that the algorithm has good robustness and real-time performance.

## COMPETING INTERESTS

Authors have declared that no competing interests exist.

## REFERENCES


1. WANG Q, ZHAO S. Comparison and analysis of video-based moving object detection algorithms [J]. Command Control & Simulation. 2012;6:013.
2. Liu X, Liu H, Qiang Z, et al. Adaptive background modeling based on mixture gaussian model and frame subtraction [J]. Journal of Image and Graphics. 2008; 13(4):729-734.
3. Senst T, Evangelio RH, Sikora T. Detecting people carrying objects based on an optical flow motion model [C]. Applications of Computer Vision (WACV), 2011 IEEE Workshop on. IEEE. 2011;301-306.
4. Vosters L, Shan C, Gritti T. Real-time robust background subtraction under rapidly changing illumination conditions [J]. Image and Vision Computing. 2012;30(12): 1004-1015.
5. Stauffer C, Grimson WEL. Adaptive background mixture models for real-time tracking [C]. Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on. IEEE. 1999;2.
6. Bouwmans T, El Baf F, Vachon B. Background modeling using mixture of Gaussians for foreground detection-a survey [J]. Recent Patents on Computer Science. 2008;1(3):219-237.
7. Zhao ZH, Tian XH, Gao YF. Real-time motion detection in complex scenes [J]. Journal of Chinese Computer Systems. 2013;34(4):915-920.
8. Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction [C]. /Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th

International Conference on. IEEE. 2004;2: 28-31.
9. Nonaka Y, Shimada A, Nagahara H, et al. Evaluation report of integrated background modeling based on spatio-temporal features [C]. Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on. IEEE. 2012;9-14.
10. Ying-hong L, Hong-fang T, Yan Z. An improved Gaussian mixture background model with real-time adjustment of learning rate [C]. Information Networking and Automation (ICINA), 2010 International Conference on. IEEE. 2010;1:V1-512-V1-515.
11. Leone A, Distante C. Shadow detection for moving objects based on texture analysis [J]. Pattern Recognition. 2007;40(4):1222-1233.
12. Martel-Brisson N, Zaccarin A. Moving cast shadow detection from a Gaussian mixture shadow model [C]. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE. 2005;2:643-648.
13. Surkutlawar S, Kulkarni RK. Shadow suppression using RGB and HSV color space in moving object detection [J]. International Journal of Advanced Computer Science and Applications. 2013;4(1).
14. Qing L, Zhu X, Shitong W. Moving objects detection of adaptive gaussian mixture models on HSV [J]. Computer Science. 2010;37(10):254-256.
15. Maddalena L, Petrosino A. A self-organizing approach to background subtraction for visual surveillance applications [J]. Image Processing, IEEE Transactions on. 2008;17(7):1168-1177.

*© 2016 Hu and He; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.*

*Peer-review history:*
*The peer review history for this paper can be accessed here:*
*http://sciencedomain.org/review-history/13569*