International Journal of Applied Mathematics in Control Engineering 3 (2020) 49-56

Contents lists available at YXpublications

International Journal of Applied Mathematics in Control Engineering

Journal homepage: http://www.ijamce.com

# Research on Moving Target Detection and Self-Shadow Removal Based on Background Subtraction

Yongqiang Zhang*, Jiaze Li, Huiping Chen, Wanzhen Zhou

*College of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China*

ARTICLE INFO

Article history:
Received 12 February 2020
Accepted 15 April 2020
Available online 19 April 2020

Keywords:
Moving targets
Background subtraction
Adaptive Gaussian mixture model
HSV color space

ABSTRACT

In order to improve the detection speed of moving targets in the video and reduce the impact of self-shadow of moving targets on the detection accuracy, so it is necessary to propose a background subtraction method. The background subtraction method of Gaussian mixture model is researched and improved on this basis. The traditional Gaussian mixture model is improved by setting a different number of Gaussian components for each pixel, which reduces the complexity of the algorithm and improves the speed of detection. The threshold segmentation of the HSV color space can eliminate the self-shadow of the moving target and improve the accuracy of detection. Therefore, we proposed a new background subtraction method, which combine adaptive Gaussian mixture model and HSV color space. After verification, the speed and accuracy of detection improved by the new method. This research has important guiding significance for moving targets detection in video.



## 1. Introduction

As a means of image information acquisition and processing, intelligent video monitoring is widely used in different fields by combining traditional video technology and digital image processing technology. Such as robot positioning (e.g., Peng et al., 2018), human pose recognition (e.g., Zhao et al., 2019), etc. Moving target detection plays an important role, especially in the vehicle security check scenes of highway and traffic checkpoint. At present, there are mainly three methods: frame difference method, optical flow method and background subtraction method that can complete the detection of moving targets. Several methods can obtain the region of interest from the captured video (e.g., Liu et al., 2017). Moving target detection by background subtraction is in a scene where the camera position is fixed and there is a still background in video. First, the parametric model of the background approximates the pixel values of the background image. Then compare each frame of the image sequence is with the background model. Finally, it is distinguished whether the pixel is the motion area or the background area by setting threshold. Based on the existing background subtraction method, this paper designs a background subtraction method that can solve the problems of weak dynamic changes in background and self-shadow of moving targets in complex scenes. The improved algorithm has made an important contribution to scene-level video surveillance.

At present, a large number of scholars have conducted research on moving target detection and background segmentation modeling. For example, the literature (e.g., Ding et al., 2019) proposed the ViBe background model incorporating color and edge features. This method solves the problems of slow ‘ghost image’ elimination and inaccurate background description. Literature (e.g., Deng et al., 2011) proposed an algorithm for detection and removal of shadow by color difference. According to the background difference of the YCBCr color space, the statistical characteristics of the target area and the shadow area are analyzed. However, the above methods are not effective for detecting dynamic changes in a complex environment. Literature (e.g., He et al., 2019, Guo et al., 2017) proposed detection of moving target by HSV color space combined with edge extraction algorithm or Otsu threshold segmentation. Neither method can solve the problem of moving target detection in a weakly dynamic background. Therefore, in order to solve the problem of dynamic background, you can analyze the pixels at a specific position in each frame of the video and find that the pixel value changes followed normal distribution. So the Gaussian mixture model and the extension of the Gaussian model are applied. Literature (e.g., Lan et al., 2017) proposed a method of image block processing was to improve the detection efficiency of Gaussian mixture model, but the effect of the shadow interference was ignored. Literature (e.g., Zhao et al., 2017) proposed to extend the

* Corresponding author.
E-mail addresses: zyq@hebust.edu.cn (Y. Zhang)

Gaussian mixture model to the neighborhood and segment the foreground with Markov random fields, which increased the complexity of the algorithm.

Research on nonparametric kernel density estimation model, the description of the image background can be obtained by the recent historical image information. Through the study of the ViBe model, we can get the sample set to be represented by the value of pixels and random pixels in its neighborhood. Although the above methods are adaptable to the weak dynamic changes of the background in the video, it cannot solve the problem of inaccurate detection due to the self-shadow of the moving target itself. Therefore, this paper proposes a background subtraction of adaptive Gaussian mixture model with HSV color space through the research of existing background subtraction methods. This algorithm solves the problems of light change, dynamic change of background and self-shadow of moving targets in complex scene. It has good robustness.

## 2. Adaptive improvement of Gaussian mixture model

*2.1 Gaussian mixture model*

Gaussian distribution, also known as normal distribution, it is a description of the statistical distribution of a large number of unlabeled sample data. The distribution characteristics of the sample data are represented by curves. The characteristic of curve is that the middle is high and the sides are low. It is called a Gaussian curve. In the image sequence, the pixels in each frame of the image are independent of each other. The value of each pixel is constantly changing due to the influence of environmental factors and camera components. The distribution curve representing the trend is obtained by statistics on the value of pixel in the image sequence. The single Gaussian model as an example, probability statistics are used for pixel value changes (e.g., Cheng et al., 2018). The mean $\mu$ of the sample points is used to represent the center value and the variance $\sigma$ is used to represent the distribution range. The Gaussian fitting is performed on the distribution of pixel values according to statistical rules. The single Gaussian model distinguishes the foreground and background by the specific pixel distribution meets the Gaussian distribution. In other words, the brightness of the $(x,y)$ point of the background B satisfies $N(\mu,\sigma)$ is a Gaussian model established for the point of $(x,y)$. Establish a Gaussian model for the point of $(x,y)$. The established Gaussian model is shown in formula (1).

$$P(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

In the formula (1), $x$ is the brightness at $(x,y)$ point, $\mu$ is the mean of the Gaussian model, and $\sigma$ is the variance of the Gaussian model. The pixel value of $(x,y)$ point is greatly deviated from the previous mean value $\mu$ when there is a moving target in the video window. If the pixel value of the $(x,y)$ point exceeds the threshold $T_p$, that is, the formula (2) is satisfied, then the $(x,y)$ point is a front attraction. In the formula, $F(x,y)$ is the pixel value of $(x,y)$ point after a large deviation and $B(x,y)$ is the pixel value of $(x,y)$ point following the normal distribution previously identified as the background.

$$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(F(x,y)-B(x,y))^2}{2\sigma^2}}>T_p \tag{2}$$

The Gaussian mixture model is an extension of the Gaussian model. The Gaussian mixture model completes the infinite approximation of the pixel value curve in the image sequence by setting $K$ Gaussian components. Gaussian mixture model is weighted by multiple single Gaussian models. Because of the uncertainty of the environment in a complex scene, it is difficult to describe the probability distribution through a single Gaussian model, so the concept of a multi-Gaussian mixture model is introduced (e.g., Wang et al., 2017).

The multi-Gaussian mixture model is analyzed. For example, there are m classes: $\varpi_1,\varpi_2,...,\varpi_m$ and each class follows a normal distribution. The mean of each distribution is $\mu_1,\mu_2,...,\mu_m$, The variances are $\sigma_1,\sigma_2,...,\sigma_m$, The proportion of each class in all classes is $P(\varpi_1),P(\varpi_2),...,P(\varpi_m)$, and the proportion satisfies the formula (3).

$$\sum_{i=1}^{m}P(\varpi_i)=1 \tag{3}$$

At the same time, set n points as observation points: $x_1,x_2,...,x_n$. Among them, the probability is represented by $P$, and the probability density is represented by $p$. The probability density function is constructed as shown in formula (4).

$$\begin{aligned}p(x)&=N(\mu_1,\sigma_1)\cdot P(\varpi_1)+N(\mu_2,\sigma_2)\cdot P(\varpi_2)\\&+\cdots+N(\mu_m,\sigma_m)\cdot P(\varpi_m)\\&=\sum_{i=1}^{m}\frac{P(\varpi_i)}{\sqrt{(2\pi)^d\left|C\right|}}e^{-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)}\end{aligned} \tag{4}$$

where $d$ is the dimension and $|C|$ is the determinant.

There are two cases of using Gaussian mixture model for moving target detection, which are known parameters and unknown parameters. Analyze the known and unknown parameters separately.

1. When the sample classification is known.

You can directly use MLE (Maximum Likelihood Estimation) for parameter estimation:

The set of unknown quantities is shown in formula (5):

$$\lambda=(\mu_1,...,\mu_m,C_1,...,C_m,P(\varpi_1),...,P(\varpi_m)) \tag{5}$$

The criterion for measuring the advantage of the probability density function is shown in formula (6):

$$p(x|\lambda)=\prod_{k=1}^{n}\sum_{i=1}^{m}\frac{P(\varpi_i)}{\sqrt{(2\pi)^d\left|C\right|}}e^{-\frac{1}{2}(x_k-\mu_i)^T C^{-1}(x_k-\mu_i)} \tag{6}$$

The optimal pending parameter is obtained by finding the standard maximum position. In order to find the position of this maximum value, we need to use the derivative to find the extreme point. The specific solution process is as follows:

$$\ln^{p(x|\lambda)} = \ln^{\prod_{k=1}^{n} P(x_k|\lambda)} = \sum_{k=1}^{n} \ln^{P(x_k|\lambda)} = \sum_{k=1}^{n} \ln^{\sum_{i=1}^{m} N(x_k,\lambda_i)P(\varpi_i)} \tag{7}$$

Derivative of formula (7):

$$\frac{\partial}{\partial\lambda}\ln^{p(x|\lambda)} = \sum_{k=1}^{n}\frac{\partial}{\partial\lambda}\sum_{k=1}^{n}\ln^{\sum_{i=1}^{m} N(x_k,\lambda_i)P(\varpi_i)} = \sum_{\mathrm{k}=1}^{m}\frac{1}{\sum_{i=1}^{m} N(x_k,\lambda_i)P(\varpi_i)}\frac{\partial}{\partial\lambda}\{\sum_{i=1}^{m} N(x_k,\lambda_i)P(\varpi_i)\} = \sum_{k=1}^{m} p(x_k|\lambda)\frac{\partial}{\partial\lambda}\{\sum_{i=1}^{m}\frac{P(\varpi_i)}{\sqrt{(2\pi)^d|C|}}e^{-\frac{1}{2}(x_k-\mu_i)^T C^{-1}(x_k-\mu_i)}\} \tag{8}$$

Then we find the derivate of each parameter. Finally, the maximum value of the formula (6) is obtained, which is the optimal pending parameter.

2. When the parameters are unknown

EM algorithm can complete the estimation of Gaussian mixture model parameters. EM estimation algorithm flow:

1) Initialization:

Solution 1: Let covariance matrix $C_{j0}$ be the identity matrix. The prior probability of each model scale is set to $\alpha_{j0} = 1/M$ , and the mean $\mu_{j0}$ is a random number.

Solution 2: The K-means clustering algorithm is used to cluster the samples, and the average value of each cluster is set to $\mu_{j0}$. Calculate the proportion of $C_{j0}$ and $\alpha_{j0}$ samples to the total using the mean of each class.

2) Estimation steps:

Let the posterior probability of $\alpha_j$ be:

$$\beta_{ij} = \frac{\alpha_j N_j(x_i|\phi)}{\sum_{k=1}^{M}\alpha_k N_k(x_i|\phi)}, 1 \le i \le n, 1 \le j \le M \tag{9}$$

3) Maximization steps:

Update weight:

$$\alpha_j = \frac{\sum_{i=1}^{n}\beta_{ij}}{n} \tag{10}$$

Update Mean:

$$\mu_j = \frac{\sum_{i=1}^{n} x_i\beta_{ij}}{\sum_{i=1}^{n}\beta_{ij}} \tag{11}$$

Update the variance matrix:

$$C_j = \frac{\sum_{i=1}^{n}\beta_{ij}(x_i-\mu_i)(x_i-\mu_i)^T}{\sum_{i=1}^{n}\beta_{ij}} \tag{12}$$

4) Convergence conditions:

Steps 2) and 3) are continuously iterated, and the above value is repeatedly updated until the iteration is terminated by satisfying formula (13). In most cases, the value of $\varepsilon$ is $10^{-3}$.

$$\left| p(X|\phi) - p'(X|\phi) \right| < \varepsilon \tag{13}$$

Therefore, the EM algorithm can be used to estimate the parameters of the Gaussian mixture model. Then analyze the Gaussian mixture model.

Gaussian mixture model is a probability density distribution model. It is a multiple probability distribution model obtained by linearly superposing multiple different weighted Gaussian distribution functions. Through the mixture and superposition of multiple Gaussian models, the curve of pixel value change of each pixel in the image sequence can be fitted.

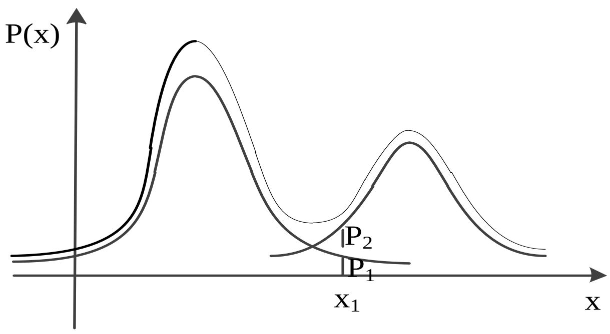



**Fig. 1.** $K = 2$ Gaussian mixed one-dimensional graph

Geometrically, the Gaussian mixture model is a weighted average of multiple Gaussian distributions. Fig. 1. is a one-dimensional graph of the Gaussian mixture model with $K = 2$ and the function expression is shown in formula (14).

$$P(x) = \sum_{\mathrm{k}=1}^{K}\partial_k N(x|\mu_k, \textstyle\sum_k), \sum_{\mathrm{k}=1}^{K}\partial_k = 1 \tag{14}$$

Where $\partial_k$ is the weight, and the weighted average of points $P_1$ and $P_2$ at $x_1$ gives the total Gaussian distribution P(x). $\mu_k$ is the mean value of the Gaussian mixture model, $\sum_k$ is the covariance matrix of the Gaussian mixture model, $N(x|\mu_k, \sum_k)$ is the Gaussian probability density function. The Gaussian probability density function is expressed as:

$$N(x|\mu_k, \textstyle\sum_k) = \frac{1}{(2\pi)^{\frac{n}{2}}\left|\sum_k\right|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t-\mu_k)^T \sum_k^{-1}(X_t-\mu_k)} \tag{15}$$

*2.2 Adaptive improvement*

The Gaussian mixture model fits the pixel value distribution of pixels by setting $K$ Gaussian components. $K$ fixed number of Gaussian components are established for all pixels in the video frame. However, the pixel value distribution of pixels in different regions conforms to different Gaussian distributions, so setting a fixed number of Gaussian component causes the algorithm complexity of the system to increase, and it will consume a lot of system resources when the computer processes it (e.g., Li et al., 2013).

Then the Gaussian mixture model is improved to obtain the adaptive Gaussian mixture model. The adaptive Gaussian mixture model assigns different number of Gaussian components to each pixel according to the change of the pixel value in the image sequence. For example, a single Gaussian model fits a stable area with small changes and a multi-peak Gaussian model fits a busy area with large changes. Algorithm flow of adaptive Gaussian mixture model:

1) The pixel value of each pixel in the first frame of the video is obtained and used as the mean $u$. The threshold of the variance $\sigma$ is set. The initialization of the single Gaussian model is complete as a formula (16).

$$\begin{aligned} \mu(x,y) &= P(x,y) \\ \sigma^2(x,y) &= threshold \end{aligned} \tag{16}$$

2) When a new image is acquired, the pixel $P_t(x,y)$ at time $t$ is sequentially matched with the existing Gaussian models. The matching conditions are shown in formula (17), where $\lambda$ is the set coefficient.

$$\left|P_t(x,y)-\mu_{p,t-1}(x,y)\right| < \lambda\sigma_{p,t-1} \tag{17}$$

3) If the matching conditions are satisfied, the pixel value is successfully matched with the Gaussian model, otherwise the matching fails. It is determined whether the number of models satisfying the updated pixel value $k$ is less than the maximum number of models of the pixel values that failed to match. If it is less, add a new Gaussian model; when the number is equal, the new Gaussian model is used to replace the model with the lowest priority. The priority is determined as formula (18).

$$priority = \varpi_{\mathrm{p}} / \sigma \tag{18}$$

Where $\varpi_p$ is the weight. Initialize the current new Gaussian model, including setting the mean, covariance and weight of the new model. The initialization of the new model is shown in formula (19), where $\alpha$ represents the learning rate.

$$\begin{aligned} \mu_p &= P(x,y) \\ \sigma_p &= threshold \\ \omega_i &= \alpha \end{aligned} \tag{19}$$

4) The mean and variance of the unmatched model remain unchanged, and the p-th Gaussian model parameter of the matched model is updated:

$$\begin{aligned} \mu_{\mathrm{p},t} &= (1-\alpha)\mu_{\mathrm{p},t-1} + \alpha P_t(x,y) \\ \sigma^2_{\mathrm{p},t} &= (1-\alpha)\sigma^2{}_{\mathrm{p},t-1} + \alpha(P_t(x,y)-\mu_{\mathrm{p},t-1})^2 \\ \omega_{\mathrm{p},t} &= (1-\alpha)\omega_{\mathrm{p},t-1} + \alpha \end{aligned} \tag{20}$$

5) After the parameters of the Gaussian model are updated, the Gaussian model of each pixel is sorted in descending order of priority. The priority is determined by referring to formula (18). The best description of the background pixels is to take the top $B$ Gaussian models:

$$B = \arg\min_k\left(\sum_{k=1}^{M}\omega_i > t\right), 0.5 < T < 1 \tag{21}$$

6) Continue to test the match between $P_t(x,y)$ and the above $B$ Gaussian models. If $P_t(x,y)$ matches any of the previous $B$ Gaussian models, the pixel point is the background point; otherwise, it is the front point.

7) Repeat steps 2)-6) until the video ends.

The adaptive Gaussian mixture background model has great advantages in the simple scene with little background change. Through the analysis of 789 frames in traffic.flv video data, the stable (10,10) point and active (220,10) point are selected to detect the first 100 frames of image sequence. The detection results of the RGB values at two points are shown in Tab. 1. and Tab. 2. The change curves of the pixel values at two points are shown in Fig. 2. and Fig. 3.

Tab. 1. RGB color distribution of (10,10) points

| Frame | 1 | 20 | 30 | 40 | 50 | 60 | 80 | 90 | 100 |
|---|---|---|---|---|---|---|---|---|---|
| R | 179 | 180 | 175 | 177 | 178 | 176 | 178 | 177 | 178 |
| G | 178 | 179 | 175 | 180 | 177 | 179 | 179 | 178 | 177 |
| B | 184 | 184 | 183 | 185 | 185 | 184 | 184 | 183 | 182 |

Tab. 2. RGB color distribution of (210,10) points

| Frame | 1 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| R | 63 | 63 | 63 | 64 | 63 | 63 | 64 | 64 | 63 | 64 |
| G | 66 | 66 | 66 | 67 | 66 | 66 | 67 | 67 | 66 | 67 |
| B | 73 | 73 | 73 | 74 | 73 | 73 | 74 | 74 | 73 | 74 |

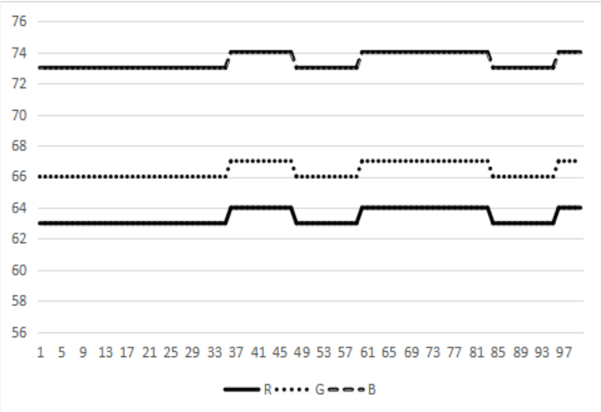



**Fig. 2.** RGB component distribution curve at coordinates (10,10)

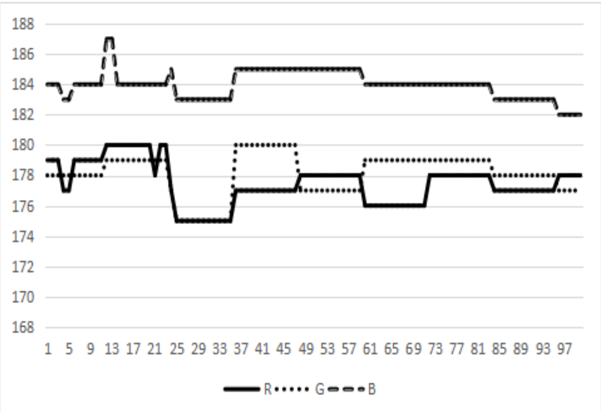
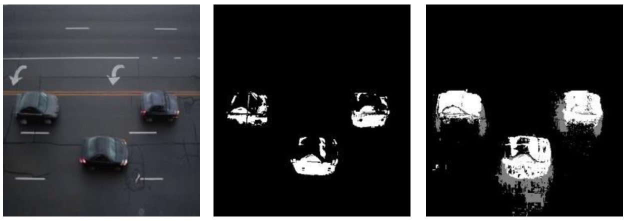



**Fig. 3.** RGB component distribution curve at coordinates (210,10)

According to the test results, the pixel value changes of the pixels in different regions of the video are different. The adaptive Gaussian mixture model can effectively reduce system operations by establishing different numbers of Gaussian components for different pixels. The difference between the adaptive Gaussian mixture model and the Gaussian mixture model is that the pixels satisfying the existing Gaussian model are no longer fitted by the multimodal Gaussian mixture model.

*2.3 Comparison of Algorithm Improvement Effects*

Gaussian mixture model (GMM) and adaptive Gaussian mixture model (AGMM) are used to model and analyze the video. The traffic.flv video is selected, which duration is 52 seconds, the frame rate of 15fps and the frame size is 480*360 pixels. The number of Gaussian components is 5, and the moving target is extracted. Display the 161st frame of the detection result. The original image is shown in Fig. 4. (a) and the detection results of the two foreground modeling methods on the moving targets are shown in Figures (b) and (c). The processing time of the two methods is shown in Tab. 3.



(a) Original figure (b) GMM model (c) AGMM model

**Fig. 4.** Comparison of processing effects of the two models

Tab. 3. Comparison of processing time between the two models

| Method | Duration/s |
|---|---|
| GMM | 87.3398048877716 |
| AGMM | 85.37715768814087 |

The detection results show that the Gaussian mixture model (GMM) and the adaptive Gaussian mixture model (AGMM) can complete the detection of moving targets in the video. However, GMM have to do complex operations on each frame pixel, so it is difficult to achieve the real-time requirements. In comparison, AGMM have strong anti-light interference ability, which can detect the moving targets with shadow. Meanwhile, the algorithm complexity of AGMM is low, so it can quickly and accurately detect moving target.

## 3. Self-shadow removal in HSV color space

*3.1 Principle of HSV color space*

The parameters of color in HSV color space are hue, brightness and saturation. HSV color space exhibits different spectral characteristics than RGB color space. Besides HSV can more accurately reflect the grayscale information and color information of the image than RGB (e.g., Wang et al.). The color component transformation from RGB to HSV space is realized by normalized mapping to range of [0,1]. The component values of $H$, $S$ and $V$ converted to the HSV color space can be calculated by the following formula (e.g., Peng et al., 2018):

$$H=\begin{cases} \theta \quad , \quad G \geq B \\ 2\pi\text{-}\theta \quad , \quad G < B \end{cases} \tag{22}$$

$$\theta=\cos^{-1}\{\frac{[(R-G)+(R-B)]/2}{\sqrt{[(R-G)^2+(R-B)*(R-G)]}}\}$$

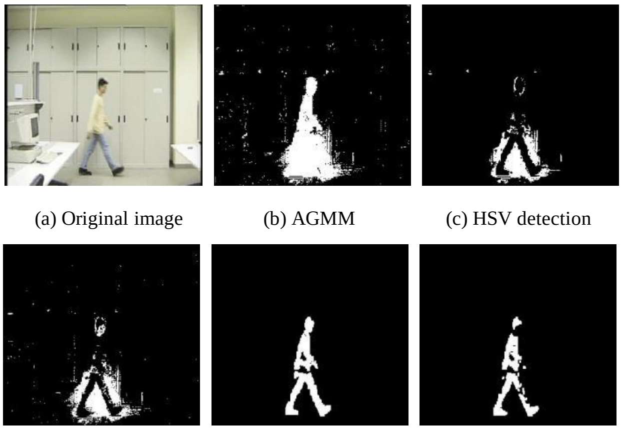
$$S=1-\frac{3\min(R,G,B)}{R+G+B} \tag{23}$$

$$V=\frac{(R+G+B)}{\sqrt{3}} \tag{24}$$

*3.2 Comparison of self-shadow removal effect*

The adaptive Gaussian mixture model can complete the detection of shadows. The effect of detecting and removing shadows by setting the shadow threshold in the RGB color space is common. The reason is that the primary colors of red, blue and green are mixed into different colors in RGB color space, which focuses on the display of colors. However, the model of HSV focuses on the expression of color, shade and brightness, which more accurately reflect the information of gray and color (e.g., Wang et al., 2004). Therefore, the detection effect of video shadows is better in HSV color space detection.

The public video data of Laboratory_raw.AVI is selected for experiments. First, the moving target is extracted through an adaptive Gaussian mixture model. The 96th frame image of the image sequence is selected for display. The 96th frame image is shown in (a) of Fig. 5. and the result of extracting the moving target by AGMM is shown in (b). The results of shadow detection through HSV and RGB color spaces are shown in (c) and (d). The results of moving targets detection are shown in (e) and (f) after the shadow removal and erosion expansion in HSV and RGB spaces.



(a) Original image (b) AGMM (c) HSV detection

(d) RGB detection (e) HSV removal (f) RGB removal

**Fig. 5.** Comparison of shadow detection removal effect

Analysis of experimental results, the threshold segmentation method in RGB and HSV color spaces can complete the detection and removal of shadows. The effect of detection in RGB and HSV spaces are different under the same experimental conditions. Some pixels of the moving target were mistakenly detected as shadow pixels. Because of the range of shadow threshold in RGB color space is large. The color performance of shadows in the HSV color space is more prominent and the threshold range of shadows is smaller, so shadow detection and removal can be complete more accurately.

## 4. Moving target detection combining AGMM and HSV

*4.1 AGMM-HSV algorithm principle*

AGMM-HSV is an algorithm obtained by combining the adaptive Gaussian mixture model with the HSV color space. AGMM reduces the redundancy of the algorithm by automatically assigning Gaussian components to each pixel. The detection result is converted to HSV space to eliminate shadows and shadow removal effect is better.

The combination of the two algorithms can overcome the difficulty of detecting moving targets due to weak dynamic changes in complex backgrounds, different lighting conditions and self-shadows of moving targets. The algorithm can achieve fast and accurate detection of moving targets and has good robustness, which is suitable for various scenarios. The flow of algorithm is shown in Fig. 6.

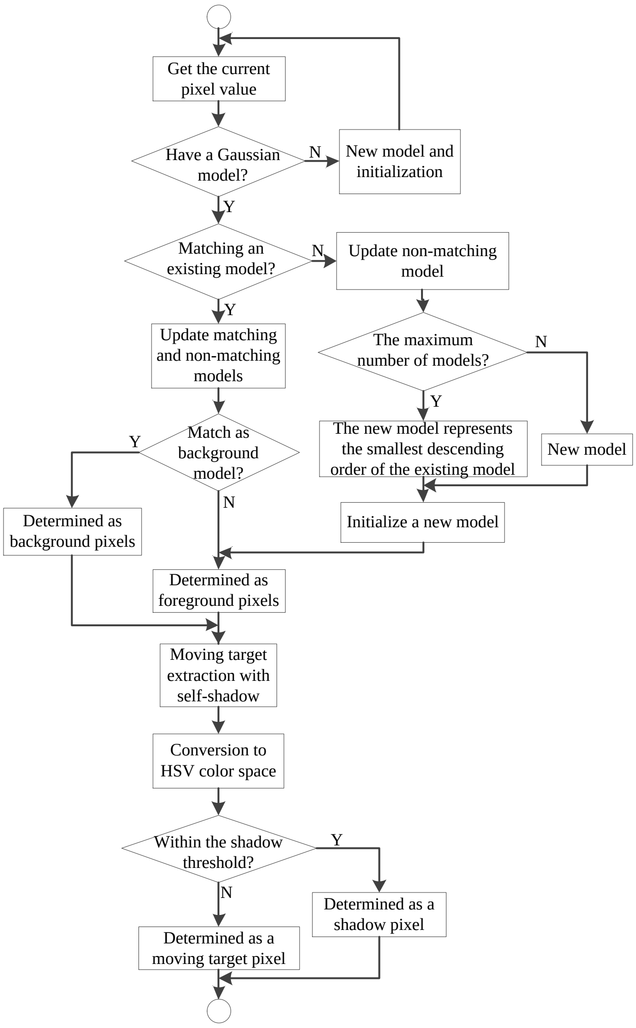
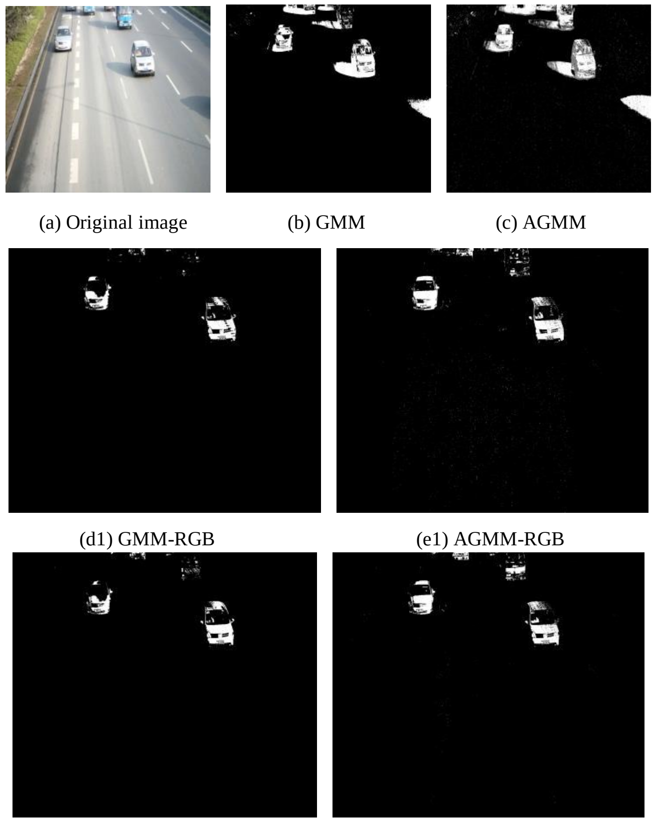



**Fig. 6.** AGMM-HSV moving target detection process

*4.2 Overall evaluation of the AGMM-HSV algorithm*

The effectiveness of the algorithm is proven and verified by experiments. The software environment of the experiment is the operating system of Windows 10 64-bit. The development platform of PyCharm and OpenCV 3.4.2 is deployed. The hardware environment of the experiment is that the CPU is 2.8GHz frequency, 8.0GB memory and the hard disk is a 256G solid state hard disk.

The experimental data was selected from the publicly available vehicle video data car.avi in a complex scene. The resolution of each frame of the video is 1920*1080 pixels. The Gaussian distribution weights are set to 3 to simulate scene changes. Then the relevant parameters are set, including the number of training frames of the background model, the threshold of background variance, and the threshold of shadow. After the relevant parameters are set and the model is initialized, the initialized model is matched by each pixel of each frame. Pixels are judged as moving target pixels or background pixels through a set threshold. Finally, the point set of moving target pixels in each frame of image is obtained, which is the foreground target. Research and experiment on Gaussian mixture model (GMM), adaptive Gaussian mixture model (AGMM), RGB color space and HSV color space. We get the experiment results. The 154th frame of the image sequence of the experimental results is selected for display. The horizontal comparison of the algorithm is shown in Fig. 7. , the picture of (a) is the original image of the experiment, the picture of (b) is the moving target extraction of the GMM model, the picture of (c) is the moving target extraction of the AGMM model. The picture of (d1) and (d2) are the processing results of the GMM model combining two color spaces and the picture of (e1) and (e2) are the processing results of the AGMM model combining two color spaces. The length of time for the algorithm to process an image is shown in Tab. 4.



(a) Original image (b) GMM (c) AGMM

(d1) GMM-RGB (e1) AGMM-RGB

(d2) GMM-HSV (e2) AGMM-HSV

**Fig. 7.** Improved algorithm effect comparison

Tab. 4. Compare the duration of processing a frame of image

| | Duration/ms | |
|---|---|---|
| Method | RGB | HSV |
| GMM | 119.20709 | 126.642903 |
| AGMM | 113.887855 | 116.169942 |

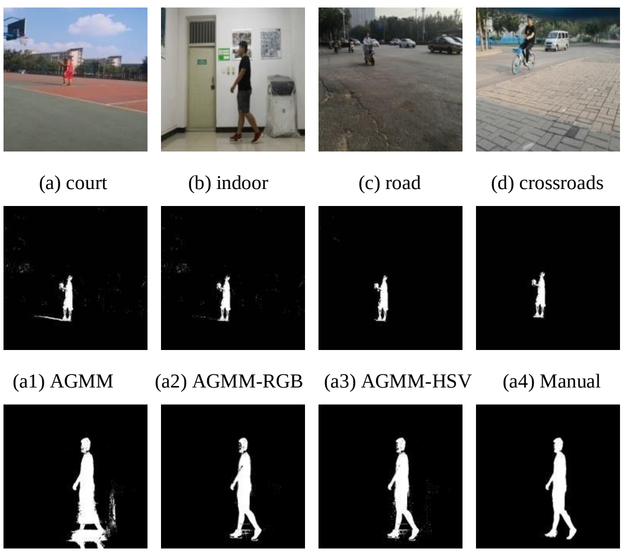
Comparative analysis of the background subtraction methods of GMM and AGMM combined with RGB and HSV color spaces, it can be obtained that the background subtraction method of AGMM has a good detection effect, and the algorithm has low complexity and fast detection speed. The improved algorithm can quickly and effectively detect and extract moving targets and removal self-shadowing. Shadow removal can be done in RGB color space and HSV color space by setting different thresholds. In the RGB color space, the shadows are removed by setting a threshold on the color parameters of the three primary color channels. In the HSV color space, the shadow and non-shadow areas are significantly different in the hue and saturation channels. Therefore, it is easier to complete the detection and removal of shadows in the HSV color space by setting a threshold. Through a comparative analysis of the experimental results, it can be obtained that RGB shadow detection has a shorter than HSV shadow detection, but the time of both meets the real-time specification at the millisecond level. The sensitivity and detection effect of the HSV color space on shadows are much better than RGB color space, so AGMM-HSV can accurately extract moving targets and meet the real-time requirements.

*4.3 Overall Evaluation of the AGMM-HSV Algorithm*

In order to prove the effectiveness and accuracy of the algorithm for detecting moving targets in different environmental conditions, multi-scale and multi-type experimental data were selected for testing. According to different experimental scenarios, the threshold parameters were adjusted. The experimental results were evaluated by qualitative and quantitative analysis.

The video data used in the experiment was obtained in a real scene though a CMOS camera. The picture of (a) is acquired in a scene where the camera is far away from a single moving target under strong lighting conditions in outdoor. The video data is named court.mp4; The picture of (b) is acquired in a scene with lights and relatively fixed background in indoor. The video data is named indoor.mp4; The picture of (c) is obtained under the conditions of backlighting in outdoor, the video background is complex and multiple moving targets are blocking each other. The video data is named road.mp4; The picture of (d) is acquired under the conditions of backlighting in outdoor, the video background is relatively fixed and multiple moving targets without mutual blocking. The video data is named crossroads.mp4.

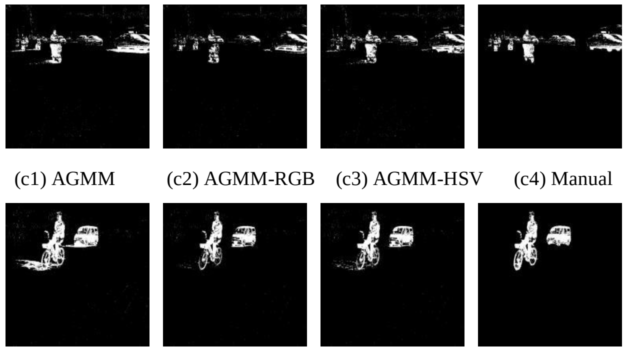
We processed multiple video data of different scales and types and displayed some of the processing result. The result is shown in Fig.8. The picture of (a) is the 251st frame of the court video processing result; The picture of (b) is the 228th frame of the indoor video processing result; The picture of (c) is the 229th frame of the road video processing result; The picture of (d) is the 218th frame of the results of the crossroads video processing. From left to right, the experimental results are AGMM moving target extraction, AGMM-RGB moving target extraction and self-shadow removal, AGMM-HSV moving target extraction and self-shadow removal, and manual processing. Because the moving target is small in picture (a), a part of the view is taken for display.



(a) court (b) indoor (c) road (d) crossroads

(a1) AGMM (a2) AGMM-RGB (a3) AGMM-HSV (a4) Manual

(b1) AGMM (b2) AGMM-RGB (b3) AGMM-HSV (b4) Manual



(c1) AGMM (c2) AGMM-RGB (c3) AGMM-HSV (c4) Manual

(d1) AGMM (d2) AGMM-RGB (d3) AGMM-HSV (d4) Manual

**Fig.8. Test image processing results**

The results are qualitatively evaluated. The court video is processing to obtain AGMM-RGB and AGMM-HSV can be used to extract small moving targets at long distances. However, the threshold method of RGB is less effective than the threshold method of HSV for shadow boundary detection. Though the processing of indoor video, it can be concluded that the shadow threshold has a large range in RGB channel. It leads to the false detection of moving target is partial pixel value within the set shadow threshold. The effect of moving target detection is better, but some shadow pixels are missed in the HSV channel. Through the processing of road and crossroad videos, it is obtained that the detection effect of AGMM-RGB and AGMM-HSV on multiple moving targets in complex environments is normal. The threshold method of RGB can remove shadows and has a higher detection rate for shadows. However, this method can easily detect the shadow as a moving target. The threshold method of HSV has a good extraction effect on moving targets and shadow areas and the recognition rate of shadow is higher.

The two methods cannot be accurately evaluated though qualitative analysis, so the results processed by AGMM-RGB and AGMM-HSV are quantitatively evaluated. The results of manual processing are standard. Then the results after AGMM-RGB and AGMM-HSV are compared. The index of the quantitative analysis is the formula (25). The false detection point is that both the moving target pixel is detected as the background pixel and the background pixel is detected as the moving pixel. The number of false detection points is $FP$. The number of moving target pixels obtained through manual processing is $TP$. The number of pixels of moving targets detected by AGMM-RGB and AGMM-HSV is $MP$. The accuracy of moving target extraction is $R$. The results of evaluation are shown in Tab. 5.

$$R=\frac{MP-FP}{TP} \tag{25}$$

Tab.5. Comparison of accuracy of moving target detection

| | | R/% | |
|---|---|---|---|
| Video | Frame | AGMM-RGB | AGMM-HSV |
| court | 251 | 91.42 | 95.42 |
| indoor | 228 | 96.43 | 96.85 |
| road | 229 | 69.05 | 77.93 |
| crossroads | 218 | 86.75 | 86.86 |

It can be obtained that both RGB and HSV color spaces can complete the detection and removal of shadows through qualitative and quantitative analysis. However, the detection effect of RGB and HSV is different in different scenes. Moving targets with their own shadows can be quickly and accurately detected and extracted by AGMM. The experimental results are obtained by combining AGMM with RGB and HSV color spaces. The results show a better shadow detection rate in RGB color space. The shadow does not stand out in the RGB channel color expression, which easily leads to the false detection of moving target pixels as background pixels. In the HSV color space, the color expression of shadows and moving target areas are significantly different, so threshold segmentation can be easily performed in the HSV color space. Besides, the recognition rate of shadows is higher in HSV color space. Therefore, the model of AGMM-HSV can be effectively extract moving targets, which has robustness.

## 5. Conclusion

In this paper, the adaptive Gaussian mixture background model is obtained by adaptively improving the background subtraction method of the Gaussian mixture model. The background subtraction method of the adaptive Gaussian mixture model can quickly and efficiently detect and extract moving targets though experimental comparison and analysis. In terms of improving the detection accuracy, the shadow threshold segmentation in HSV and RGB channels can detect and remove the self-shadow of the moving target to improve the detection accuracy. However, the span of the shadow threshold in the RGB color space is large, so the accuracy of detecting the shadow is not high and it is prone to misdetection. The threshold range of shadows in HSV color space is small. The small span of the shadow is not easily to false detection. Therefore, the shadow detection efficiency in the HSV color space is higher than that in the RGB color space.

The experimental results show that the adaptive Gaussian mixture background model combined with HSV color space can effectively overcome the effects of weak dynamic changes in the background and the self-shadow of the moving target in complex scenes. The background subtraction method based on the AGMM-HSV model can quickly and accurately perform segmentation extraction and self-shadow removal of moving targets in a video. Improved model has higher accuracy and stability than previous model detection.

## Acknowledgements

The work has been partly supported by Natural Science Foundation of Hebei Province (F2018208116); Hebei Science and Technology Support Plan Project (16210312D); Soft Science Funding Project of Shijiazhuang Science and Technology Bureau (195790055A).

## References

Peng, R., et al., 2018. Design and Implementation of Monocular Vision SLAM for Mobile Robot Based on ROS. International Journal of Applied Mathematics in Control Engineering, 1(1): p. 85-91.

Zhao, X., Du H., 2019. Research on Human Walking Feature Extraction and Identification Recognition System. International Journal of Applied Mathematics in Control Engineering. 2(1): p. 81-86.

Liu Z.M., He S.J., Hu W.J., et al, 2017. Moving target detection of video sequences based on background subtraction. J. Computer Application,37(06):1777-1781.

Ding Z., Lu W., 2019. Moving object detection algorithm based on Vibe background modeling. J. Computer System Applications,28(04):183-187.

Deng Y.L., Wu L.F., Li Y.T., 2011. An effective automatic image shadow removal algorithm. J. Signal Processing, 27 (11): 1724-1728.

He Z.L., Xiao Z.J., Yan Z.G., 2019. Vehicle License Plate Location and Character Segmentation Based on HSV and Edge Information. J/OL.. Journal of Qilu University of Technology, (03): 44-48.

Guo H.L., Lin J., Zhang X., 2017. Segmentation of background images of regenerated rice plants and soil based on HSV space. J. Research of Agricultural Mechanization, 39 (07): 169-174.

Yan Y.Y., Mu P.A., 2017. An improved Gaussian mixture model for moving target detection algorithm. J. Data Communications, (05): 25-28.

Zhao Z, Bouwmans T, Zhang X, et al,2012. A Fuzzy Background Modeling Approach for Motion Detection in Dynamic Backgrounds[M]// Multimedia and Signal Processing. Springer Berlin Heidelberg.

Cheng J., Wang D.W., Yang L.K., et al, 2018. An Improved Gaussian Mixture Model for Coal Gangue Video Detection. J. Journal of Central South University (Science and Technology), 49 (01): 118-123.

Wang Y.N., Zhang H., 2017. Shadow removal algorithm based on improved Gaussian mixture model and texture. Computer Engineering and Design. 38(7): 1879-1884.

Yang C.D., Guo S., 2018. Improved shadow detection algorithm based on HSV space. J. Computer Engineering and Design.

Li B.H., Yang G., 2013. Adaptive foreground extraction of mixed Gaussian model. J. Journal of Image and Graphics, 18 (12): 1620-1627.

Wang X.Y., Jia K.B., Sun Z.H.. Target shadow removal method combining image region information in HSV color space. C. // The Ninth National Conference on Signal and Intelligent Information Processing and Applications.

Peng M.Y., Wang J.H., Wen X.X.,et al, 2018. Water surface line detection based on HSV space image features. J.Journal of Image and Graphics,23 (04): 526-533.

Wang Z, Bovik A C, Sheikh H R, et al,2004. Image Quality Assessment: From Error Visibility to Structural Similarity. J. IEEE Transactions on Image Processing, 13(4)

***Zhang Yongqiang*** born in 1981, is a PhD candidate, an associate professor, and a master's supervisor. He was graduated from Anhui University of Technology and received master's degree of Computer Application Technology. Main research areas: artificial intelligence, Internet of things technology.

***Li Jiaze*** born in 1995, graduated from Hebei University of Science and Technology with a major in Computer Technology. The main research directions are: digital image processing technology, computer vision, artificial intelligence.

***Chen Huiping*** born in 1993, she obtained a master's degree in Computer Science and Technology at Hebei University of Science and Technology. Her main research interests are pattern recognition and image processing.

***Zhou Wanzhen*** born in 1966, PhD, professor. He was graduated from the Ordnance Engineering College of the Chinese People's Liberation Army in 2009 with a doctorate degree in engineering. He studied at Harbin Institute of Technology with a master of science degree. His main research directions are: network and database, Internet of Things.