

# Moving Object Detection Based on Codebook Algorithm and Three-frame Difference

Shuai Li<sup>1,2</sup>, Peixun Liu<sup>1</sup> and Guangliang Han<sup>1</sup>

<sup>1</sup>. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, 130033, China.

<sup>2</sup>. University of Chinese Academy of Sciences, Beijing, 100049, China.  
[hangl@ciomp.ac.cn](mailto:hangl@ciomp.ac.cn)

## Abstract

*In this paper, we propose a moving object detection method based on codebook and three-frame difference. The method uses codebook to learn the background and the three-frame difference is employed to extract initial foreground object, then we adopt the processes based on Log edge detection and component filling. It associates the foreground object obtained by codebook with the object, which is obtained by improved three-frame difference to perform logic “AND” operation, and gets the final moving object. Experimental results show that our method can effectively eliminate ghosting, double shadow and cavities of the video sequence, and the proposed method has strong robustness under the variety changing background, as well as obtains the entire moving object.*

**Keywords:** moving object detection, codebook, three-frame difference, edge detection

## 1. Introduction

Moving object detection plays an important role in the field of computer vision, it is also essential to moving object tracking and action analysis, and it is widely applied to various areas, such as traffic, security, medical science and so on. In recent years, a great many algorithms researched by experts, such as optical flow method, background subtraction, frame difference, Gaussian Mixture Model (GMM), codebook and some improved algorithms. The optical flow method [1] is sensitive to varying light of background. It cannot present the foreground in real time, and the computational method is complicated. As far as background subtraction [2,3] is concerned, the computational method is easy, but it cannot get effective detection under the variety changing background. Frame difference [4] can present real-time detection, the method of which is simple to implement, however, it is difficult to detect unchanged pixels, thus the problem of double shadow and cavities may generate. GMM [5, 6, 7] can effectively overcome the adverse effects, such as sunlight changing, branch shaking. But GMM has the disadvantages of processing slowly, obvious shadow of the image and isolated noises in the complicated background. Codebook [8] is an effective algorithm to extract the entire foreground. It has good robustness in some cases of light changing, moreover, its processing speed is faster than GMM, but the algorithm based on codebook needs to learn the background, so there is a phenomenon of ghosting in the video sequence.

In the view of above problems, we propose an improved method of moving object detection based on codebook and three-frame difference. The improved method makes use of advantages of two algorithms. It not only eliminates the phenomenon of ghosting caused by codebook, but also solves the problem of double shadow and cavities caused by three-frame difference in the video sequence. Our performance can effectively abstract the entire moving object in the varying scenes.

The frame of article is shown as follows: In Section 2, we discussed in detail the contents and fundamental of adopted algorithm and important improvement to

conventional algorithms. And the comparing and analyzing experimental results are shown in Section 3. Conclusion and future work are presented in Section 4.

## 2. Moving Object Detection

### 2.1. Main Idea of Method

In this paper, we propose an improved method, which based on codebook, three-frame difference, edge detection, component filling, median filtering, morphological processing and logic operation. The frame of method is shown in Figure 1. Codebook algorithm is used for foreground-background segmentation firstly. Then, three-frame difference is employed in detecting the foreground, and then we adopt Log edge detection and component filling to optimize the foreground for eliminating cavities caused by three-frame difference. Finally, we associate the foreground object obtained by codebook with the object obtained by improved three-frame difference to do logic “AND” operation and the final foreground is acquired.

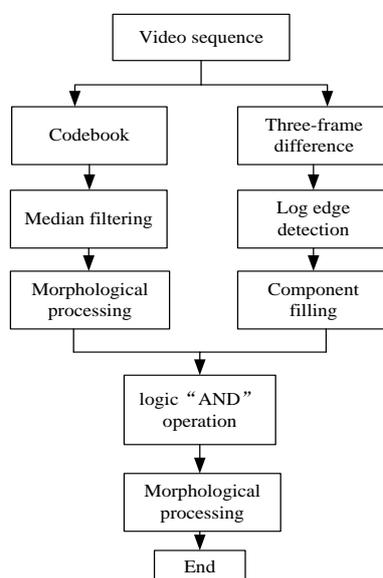


Figure 1. The Frame of Proposed Algorithm

### 2.2. Conventional Algorithms Introduction

#### 2.2.1. Codebook Algorithm

Codebook algorithm adopts a quantization technique [9]. Sample background values at each pixel are quantized into codebooks which represent a compressed form of background model for a long image sequence [8, 9]. Firstly, for each pixel, it need to build a codebook consisting of one or more codewords to describe the background. Samples at each pixel are clustered into the set of codewords based on a color distortion metric and brightness bounds [10]. Secondly, the image can be partitioned to foreground and background according to two conditions. On one hand, the color distortion is less than defined threshold, on the other hand, the brightness lies in brightness range of codeword. If the pixel meets these two conditions, it is classified as background. Otherwise, it is classified as foreground.

Algorithm for codebook construction as follows:

(1) Construction of initial codebook: For each pixel, it builds a codebook, which consists of RGB vector  $v_i$ , the min and max brightness  $\tilde{I}, \hat{I}$ , the frequency with which the codeword has

occurred  $f$ , the longest period that the codeword has not recurred  $\lambda$ , the first and last occurred times  $p, q$ .

(2) The algorithm estimates the color distortion metric and brightness through developing a color model. It has sampled pixel value at time  $t$   $x_t = (R, G, B)$ , codeword  $c_i$ , and color model  $v_i = (\bar{R}_i, \bar{G}_i, \bar{B}_i)$ .

$$\begin{aligned} \|x_t\|^2 &= R^2 + G^2 + B^2, \\ \|v_i\|^2 &= \bar{R}_i^2 + \bar{G}_i^2 + \bar{B}_i^2, \\ (x_t, v_i)^2 &= (\bar{R}_i R + \bar{G}_i G + \bar{B}_i B)^2 \end{aligned} \quad (1)$$

The color distortion  $\delta$  can be calculated by

$$\begin{aligned} p^2 &= \|x_t\|^2 \cos^2 \theta = \frac{(x_t, v_i)^2}{\|v_i\|^2}, \\ \text{colordist}(x_t, v_i) &= \delta = \sqrt{\|x_t\|^2 - p^2} \end{aligned} \quad (2)$$

In the codebook training period, it samples  $N$  frames image sequences for learning background. For all codewords in  $N$  frames, we should find out the codeword  $c_m$  matching to  $x$  according to two conditions:

$$\text{colordist}(x, c_m) \leq \varepsilon$$

$$\text{brightness}(I_t) = \begin{cases} \text{true}, & \theta_1 \hat{I}_i \leq I_t \leq \min\{\theta_2 \hat{I}_i, \tilde{I}_i\} \\ \text{false}, & \text{otherwise} \end{cases} \quad (3)$$

Update the matched codeword as in Step (3), otherwise, go to Step (4).

(3) Updating codebook  $c_k$ : the RGB vector  $v_k = (\frac{f_k \bar{R}_k + R}{f_k + 1}, \frac{f_k \bar{G}_k + G}{f_k + 1}, \frac{f_k \bar{B}_k + B}{f_k + 1})$ , the min and max brightness  $\min\{I, I_k\}$  and  $\max\{I, I_k\}$ , the frequency with which the codeword has occurred  $f_k + 1$ , the longest period that the codeword has not recurred  $\max\{\lambda_k, t - q_k\}$ , the first and last occurred times  $p = p_k$  and  $q = t$ .

(4) Creating new codebook  $c_k$ : the RGB vector  $v_k = (R, G, B)$ , the min and max brightness are both  $I$ , the frequency with which the codeword has occurred  $f = 1$ , the longest period that the codeword has not recurred  $\lambda = t - 1$ , the first and last occurred times  $p = t, q = t$ .

(5) Eliminating the redundant codebooks: The codeword is generated by noises or moving object, if the codeword has not been recurred for a long time. Therefore, for eliminating the redundant codebooks, it sets  $\lambda_i = \max\{\lambda_i, N - q_i + p_i - 1\}$ .

Although the codebook algorithm needs to learning background, it can handle the scenes containing varying light or shaking branches, and has good robustness for different types of video sequences.

### 2.2.2. Three-frame Difference

Frame difference is a moving object detection method on the basis of calculating differences between two frames and acquiring the absolute difference at every pixel position [11]. It takes less computation amounts, and gets the moving target at express speed. However it merely

stores the absolute difference, thus we cannot detect the overlapping region. The algorithm has disadvantages of “cavities” and “double shadow” phenomena. But it also has such advantages as excellent stability and simple algorithm. The three-frame difference [12,13] is an improved method based on frame difference. As is shown in Figure 2. At every pixel position, it considers consecutive video frames  $I_{k-1}$ ,  $I_k$ ,  $I_{k+1}$ . Firstly, the difference is calculated between the first two frames and stores the first difference.

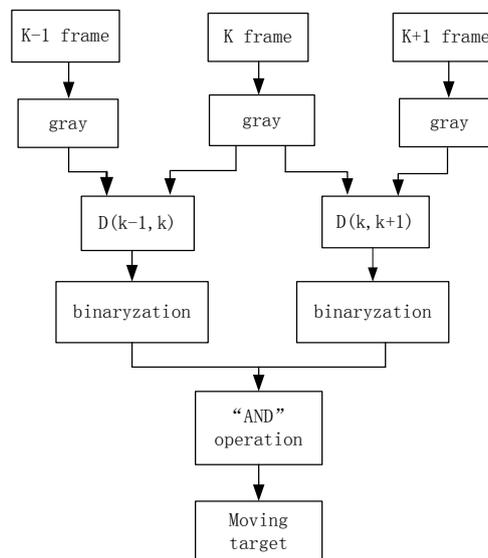
$$D(k-1, k) = \begin{cases} 1, |I_k(i, j) - I_{k-1}(i, j)| \geq T \\ 0, else \end{cases} \quad (4)$$

Secondly, it calculates the difference between the latter two frames and store the second difference.

$$D(k, k+1) = \begin{cases} 1, |I_{k+1}(i, j) - I_k(i, j)| \geq T \\ 0, else \end{cases} \quad (5)$$

Finally, logic “AND” operation is performed between the above two results and gets the real moving target profile.

$$D_k = \begin{cases} 1, D_{(k-1,k)}(i, j) \cap D_{(k,k+1)}(i, j) = 1 \\ 0, else \end{cases} \quad (6)$$



**Figure 2. The Algorithm Frame of Three-frame Difference**

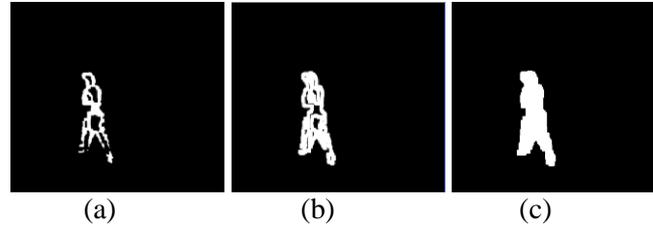
## 2.3. Improvement and Fusion Algorithms

### 2.3.1. Log Edge Detection and Component Filling

After performing the three-frame difference, an improved method based on Log edge detection and component filling is proposed to eliminate “cavities” phenomenon. The Log edge detection [14, 15] combines Gaussian Smoothing Filter with Laplacian Sharpening Filter. Its function lies in smoothing interference form noise and enhancing the detail and edge of image. The Log operator, namely Laplacian-of-Gaussian operator, is represented as

$$\nabla^2 h(r) = -\left[ \frac{r^2 - \sigma^2}{\sigma^4} \right] e^{-\frac{r}{2\sigma^2}} \quad (7)$$

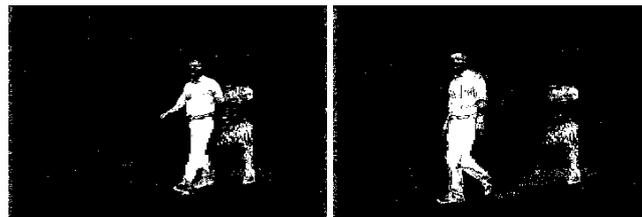
where  $\sigma$  is standard deviation of Gaussian model. The smaller value  $\sigma$  makes the enhancing effect better. The bigger value  $\sigma$  is, the better smoothing effect we get. According to the experiment, we set  $\sigma$  equals to 0.5. The experimental results are shown in Figure 3. After performing the Log edge detection, we adopt the component filling for eliminating “cavities” phenomenon, and extract the foreground.



**Figure 3. (a)Three-frame Difference; (b) Log Edge Detection; (c) Component Filling**

### 2.3.2. Fusion Algorithm

The codebook algorithm has several disadvantages during background learning phase. If the moving object appeared at the beginning, there will be a ghosting during the training background. If the moving object suddenly gets stopped in the detection process, the object will be still visible in the background. Figure 4 can illustrate this situation well.



**Figure 4. The Codebook Detection in the 33rd and 61st Frame with Ghosting Image**

The three-frame difference can get rid of ghosting problem. Therefore, we combine three-frame difference with codebook algorithm. We consider the processed image by codebook algorithm as  $B1_k$  and the processed image by eliminating cavities based on three-frame difference as  $B2_k$ . Next, logic “AND” operation is performed on both  $B1_k$  and  $B2_k$ .

$$B_k = \begin{cases} 1, & B1_k(i, j) \cap B2_k(i, j) = 1 \\ 0, & else \end{cases} \quad (8)$$

Our fusion method makes full use of advantages of codebook and three-frame difference method, and it also can eliminate disadvantages of these two methods. It avoids the “ghosting” phenomenon caused by codebook and the double shadow phenomenon by the three-frame difference. Therefore, the entire moving objects in the video are effectively detected.

### 2.3.3. Morphological Processing

We employ the morphological opening operation [16] so as to make image contour smooth. Opening operation is defined as the process of corrosion and expansion. The function of opening operation is disconnecting the narrow and eliminating burrs. A structural element  $S$  could be used to do opening operation on  $A$ , denoted by  $A \circ B$ .

$$A \circ S = (A \ominus S) \oplus S \quad (9)$$

A structural element  $S$  could be used to do corrosion operation on  $A$ . Defined as:

$$A \ominus B = \{ \| x \mid S_x \subseteq A \| \} \quad (10)$$

A structural element  $S$  could be used to do expansion operation on  $A$ . Defined as:

$$A \oplus S = \{ \| x \mid S_x \cap A \neq \emptyset \| \} \quad (11)$$

### 3. Experimental Results

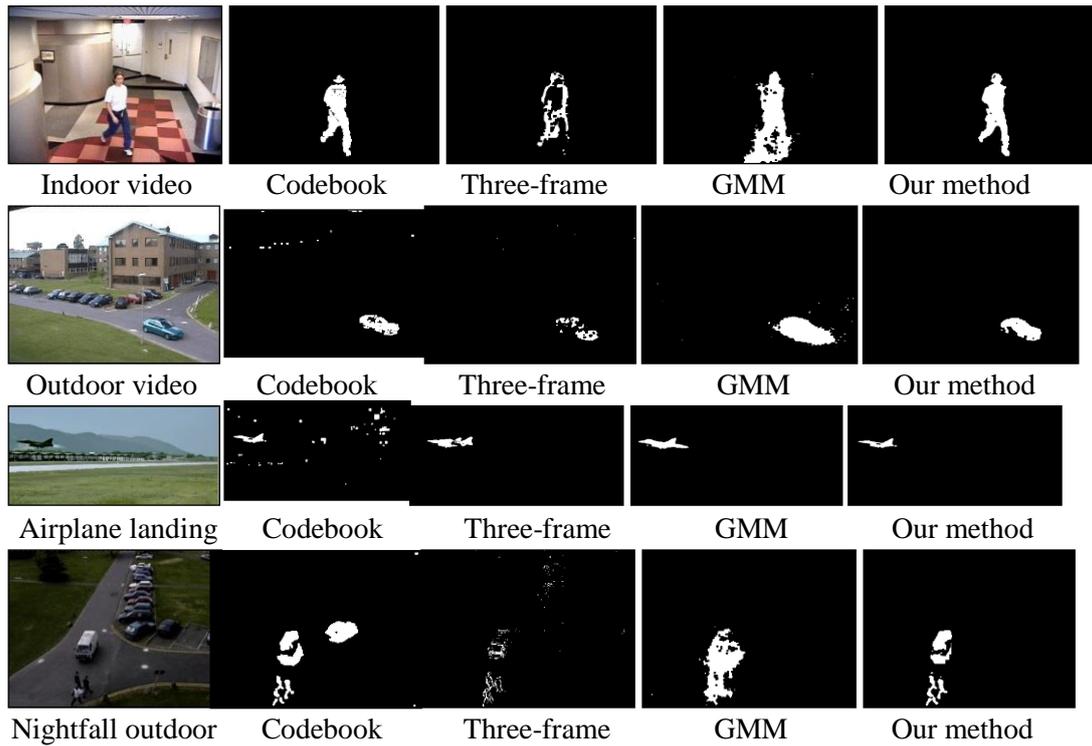
The experiment platform was implemented in C using OpenCV 1.0 library and Visual Studio 2010. In order to test and verify our method's efficiency and accuracy, it is performed on video sequences with variety changing scenes in an Intel Core i3-2120 3.30GHz PC. In the experiment, parameters are set as follows: for the codebook algorithm, the number of frames used for training background is 30. In the phase of matching codeword, the threshold  $\varepsilon$  is set to 8, and parameters  $\theta_1$  and  $\theta_2$  are set to 0.4 and 1.5. In the three-frame difference, the two thresholds  $T$  are both 20. In the phase of Log edge detection, the standard deviation  $\sigma$  is set to 0.5.

The real-time ability and integrity of detected targets in the algorithm are both considered in evaluating the performance of algorithm. We choose some typical videos to compare our present method with some traditional methods, which are shown in Table 1.

**Table 1. Type of Detection Video**

Video name	Moving object	Object number	Scene condition	Resolution ratio
Indoor video	Person	1	Medium noise level, little changing light	320×240
Outdoor video	Car	1	High noise level, varying light and complicated background	384×288
Airplane landing video	Airplane	1	Varying light, fast moving target	556×266
Nightfall outdoor video	Persons and car	4	High noise level, dark and complicated background	320×240

We evaluate the integrity of detected targets on PETS2001 test video and Internet video. System was tested over four types of video in different surroundings, and experimental results are illustrated in Figure 5. The first video is an indoor video that a woman goes across the hallway with little changing light. We find out the moving object has the evident cavities processed by the three-frame difference, and there are some shadow problems in the detection results of GMM. The second video is an outdoor video that a car passes through a housing estate with varying light and complicated background. It is obvious that varying light caused some interference noises in the process of codebook. And the three-frame difference also has evident cavities problem. The GMM brings about obvious shadow of the moving car and isolated noises phenomenon. The third video's surroundings are similar to the second video's, so their experimental results have common defects. However, airplane landing is a fast moving, so double shadow phenomenon is caused by the process of three-frame difference. The last video is a nightfall outdoor video, because of little varying light, the codebook has little interference noise. However, the codebook has phenomenon of ghosting in the video sequence. And the result of GMM hardly appears a clear outline of the target. Comparing experimental results of three traditional methods, in general we draw a conclusion that our method can effectively detect the entire and clear moving object in the various scenes.



**Figure 5. Algorithm Comparison**

Another norm that we evaluate the performance of algorithm is the real-time ability. As it is shown in Figure 6, We compare detection time of our method with 3 conventional algorithms in the various scenes. Resolution of Indoor video and Nightfall outdoor video are both  $320 \times 240$  pixels. Outdoor video resolution is  $384 \times 288$  pixels. Airplane landing video resolution is  $556 \times 266$ . Along with the increase of resolution, the detection time of 3 conventional algorithms increases steadily. Among them, the GMM takes the longest time to detect moving object, and the three-frame difference takes the shortest time.

As it is illustrated in Figure 6, the detection time of our method changes unsteadily. While the video resolution is  $320 \times 240$  pixels, the detection time of our method is shorter than the GMM. Whereas if the video resolution is larger, the detection time will increase significantly. It is supposed that the process of Log edge detection and component filling results in this weakness. We consider there are significant optimizations and improvements still to be made.

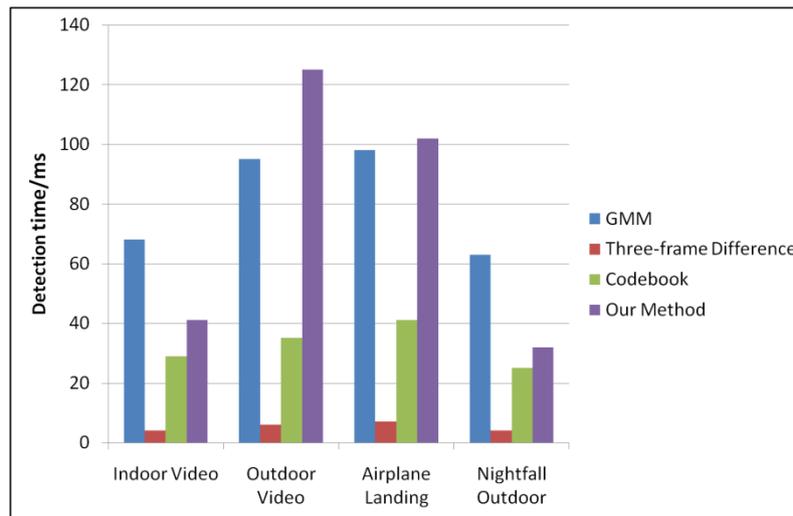


Figure 6. Detection Time Comparison

#### 4. Conclusion

In this paper, we present a simple and high quality moving object measure by combining codebook algorithm with improved three-frame difference. It utilizes hybrid detection algorithms to extract all moving objects in sequence of image frames with the processing intervals about real-time under many difficult conditions. Evaluation results show that our method outperforms other conventional algorithms, which eliminates the phenomenon of ghosting caused by codebook and the phenomena of double shadow and cavities caused by three-frame difference. It has been examined in several scenario and has wide applications. The high quality and simple operations make our method suitable for reporting events in traffic, security, medical science and so on, and this can be applied in intelligent residential security system and video surveillance system. Our system is fully automated for background update, and this makes it to be smart and adaptive to scene changes. However, the limitation of detection time still needs to be improved and optimized in the future. It would be interesting to explore our method to identify overlapping objects and enhance the system robustness to various complex videos.

#### References

- [1] Y. Yang, "Moving Objects Tracking Based on Improved Optical Flow Method", *Computer & Digital Engineering*, vol. 39, no. 9, (2011), pp. 108-111.
- [2] H. Xu, "Key Measure Frame Techniques Based on Image Difference", *Computer Engineering and Design*, vol. 31, no. 12, (2010), pp. 2849-2852.
- [3] Z. Lu, D. Kong, X. Li and J. Wang, "A Method for Moving Object Detection Based on Background Subtraction and Three-frame Differencing", *Computer Measurement and Control*, vol. 21, no. 12, (2013), pp. 3315-3318.
- [4] J. Chen, J. Zhang, S. Liu and X. Lu, "Improved Target Detection Algorithm Based on Background Modeling and Frame Difference", *Computer Engineering*, vol. 37, (2011), pp. 171-173.
- [5] X. Kang and J. Wu, "Object Detecting Technology Based on Gauss Background Modeling", *Chinese Journal of Liquid Crystals and Displays*, vol. 25, no. 3, (2010), pp. 454-459.
- [6] J. Liu and L. Wang, "An Improved Algorithm of Gaussian Mixture Model Background Method", *Computer Engineering and Applications*, vol. 46, no. 13, (2010), pp. 168-170.
- [7] Y. Hua and W. Liu, "Moving Object Detection Algorithm of Improved Gaussian Mixture Model", *Journal of Computer Applications*, vol. 34, no. 2, (2014), pp. 580-584.
- [8] K. Jiang, A. Li and Y. Su, "Moving Targets Detecting Algorithm in Video Based on Improved Codebook Model", *Journal of University of Electronic Science and Technology of China*, vol. 41, no. 6, (2012), pp. 932-936.
- [9] L. Xiong and W. Liu, "Foreground Detection Algorithm Based on Background Codebook Model", *Science Technology and Engineering*, vol. 10, no. 9, (2010), pp. 2117-2121.

- [10] C. Guo and P. Wang, "An Algorithm Based on Codebook Model to Moving Objects Detection", Journal of Image and Graphics, vol. 15, no. 7, (2010), pp. 1079-1083.
- [11] Z. Lu, D. Kong, X. Li and J. Wang, "A Method for Moving Object Detection Based on Background Subtraction and Three-Frame Differencing", Computer Measurement & Control, vol. 21, no. 12, (2013), pp. 3315-3318.
- [12] W. Zhang, R. Li and J. Zhu, "An Object Detection Algorithm Based on Gaussian Mixture Models and Three-frame Difference", Modern Electronics Technique, vol. 35, no. 8, (2012), pp. 57-60.
- [13] W. Xu, L. Wang and X. Li, "Three-image Difference Algorithm for Moving Target Detection Based on OpenCV", Computer and Digital Engineering, vol. 39, no. 11, (2011), pp. 141-144.
- [14] J. Yang, T. Cheng and Z. Zhong, "Edge Detection Technique Combined with Mathematics Morphology and LoG Operator", Computer Engineering and Applications, vol. 47, no. 36, (2011), pp. 177-179.
- [15] S. Das, "Comparison of Various Edge Detection Technique", International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 9, no. 2, (2016), pp. 143-158.
- [16] Y. Wu, "Application of Morphology in Image Processing", Computer and Modernization, vol. 5, (2013), pp. 90-94.

## Authors



**Shuai Li**, she was born in 1992. She received her B.E. degree in electronic information science and technology from Dalian Maritime University, China, in 2014. She is now a M.S student in Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. Her research interests are target tracking and machine learning.



**Peixun Liu**, he was born in 1986. He received his PhD degree in 2015 from College of Computer Science and Technology, Jilin University, Changchun, P. R. China. He is now a research assistant in Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His research interests include computer vision, image processing and pattern recognition.



**Guangliang Han**, he was born in 1968. He received his PhD degree in 2005 from Chinese Academy of Sciences. He is a leader researcher in Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His research interests include computer vision, image and video processing.

