



Journal of  
**Software  
Engineering**

ISSN 1819-4311



Academic  
Journals Inc.

[www.academicjournals.com](http://www.academicjournals.com)

## Moving Object Detection Research Based on Background Image Set and Sparse Analysis

Huang Ke

School of Electrical and Information Engineering, Hunan International Economics University, Changsha, Postcode, 410205, China

### ABSTRACT

A moving target detection method is proposed based on the background image set and sparse representation in this study. The method combined the Robust Principal Component Analysis (RPCA) and the image block analysis method based on sparse representation. The authors got a series of background images from a video sequence by RPCA and combined these background images as the background image set, treating image block as basic unit and moving target was extracted from input frame by image block analysis method based on sparse representation. The simulation results indicate that when the background illumination mutates, the proposed method can effectively eliminate the impact of environment noise and reduce the false detection rate of target detection.

**Key words:** Image processing, target detection, background image set, robust principal component analysis (RPCA), sparse representation

### INTRODUCTION

Image background Gaussian noise, jitter foliage, light and other disturbances can cause changes in the pixel values of uncertainties, these result in misjudgment between foreground pixels and background pixels. If the misjudgment of the foreground pixels and background pixels is solved, which is caused by the various disturbances, background modeling methods become mainstream target detection method, the idea is that the background characteristics is extracted in the video sequence, a mathematical model is built to describe the background. Representative methods are the background modeling method based on Gaussian mixture model (Stauffer and Grimson, 2000), the K Gaussian distributions are established for each pixel in the hybrid model, the top N distributions are selected as a background model, the moving target detection can be satisfied under the Gaussian noise environments. In addition, Kim *et al.* (2004) proposed moving object detection method based on the model of the code background (Kim *et al.*, 2004), the color and brightness information of each pixel background are used in the image sequence to build a codebook, whether the input image pixel belongs to the background codebook is determined, goals testing is realized, this method can solve foliage jitter impact on the target detection.

However, the above modeling methods are based on statistics of single-pixel changes over time. Because the eigenvalues of a single pixel is sensitive to light, when the image background illumination mutation occurs, the local pixel values are mutated, pixel characteristic mutation values do not satisfy the mathematical model distribution (such as a Gaussian distribution) or depart from certain quantization range (e.g., codebook model). Therefore, due to the interference of light mutation, foreground and background pixels misjudgment problem is not well resolved.

Based modeling method for a single pixel, the correct moving target can not be extracted in the light mutation in this study, a moving target detection method is presented based on background image collection and block sparse analysis. In the method, a background image from a set series of video sequences is obtained by Robust Principal Component Analysis (RPCA) (Candes *et al.*, 2010), these background images are combined as a background set, image blocks are used as the basic unit, the image block analysis and processing are made based on sparse representation, the precise segmentation are achieved in the movement target area.

## **MATERIALS AND METHODS**

**Detection framework based on the background set:** When the dynamic changes in the background image, the problems mainly are in follow based on a single-pixel modeling approach:

- Dynamic changes of the background image can not accurately be described in single-pixel modeling approach
- Individual pixels are sensitive to changes in illumination, it is easy in the target detection that there is misjudgment between foreground pixels and background pixels

In fact, when there is a dynamic change for a video sequence, according to the time correlation and the spatial correlation between pixels of adjacent frames, the dynamic information can be found from the adjacent image frames. Thus, the dynamic change of the image background is effectively described by way of building background set. Meanwhile, when the image background illumination mutations occur on a single pixel, the pixel value variation is caused by such mutations, it will be extremely strong, however, the image block is composed of a plurality of pixels, the intensity of the changes are significantly less than the strength of a single pixel, the resistant tiles background illumination mutation is stronger than a single pixel. In order to reflect the true background information of the image, the background image is accurately described, a method is proposed to build a background image set and in turn, collections are composed of the background image tiles. Image blocks are used to describe the background in the light mutations, it can still be a valid representation of the background image and then it is able to determine accurately the background and foreground from the input image, the foreground are extracted effectively to detect targets.

Based on background set, the background pixels and foreground pixels are separated in target detection, the separation process are converted for the separation process foreground block and background block. Thus, target detection problem is whether an input image block is background block set, the detection error is effectively prevented which is generated by the pixel perturbation, the principles are shown in Fig. 1.

Target detection method is used based on the collection of background, the background set and the separation are the core between the foreground bock and the background block.

**Problem description:** An image is composed of the relatively stable background image and relatively sparse foreground image. If the video frame series of the turn cascade are composed of a column vector matrix, the background image recovery of these video frames can be attributed to the problem, the low-rank matrix is recovered from the observation matrix, therefore, the building processes of background set can be defined as follows:

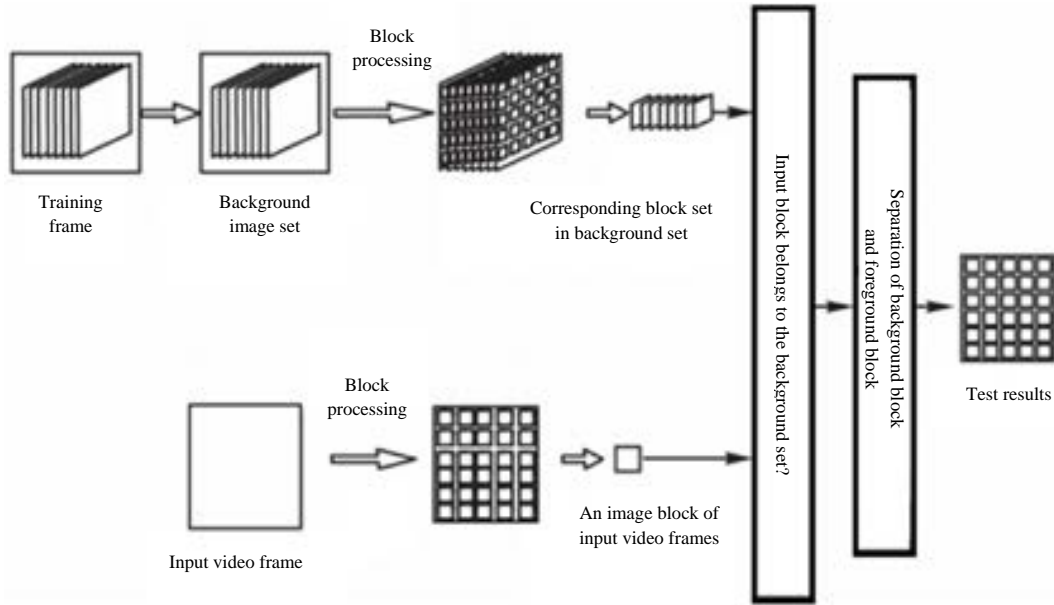


Fig. 1: Schematic of target detection method based on background image collection

- Definition 1: Suppose an image  $f$  can be expressed as the superposition of the background  $b$  and foreground  $t$ , i.e.,  $f = b + t$ , wherein,  $f, b, t$  are a column vector, which is obtained by line scanning of the image matrix, then, the matrix  $F = [f_1, f_2, \dots, f_n]$ ,  $B = [b_1, b_2, \dots, b_n]$ ,  $T = [t_1, t_2, \dots, t_n]$  are composed of  $n$  vectors  $f, b, t$ , they must satisfy  $F = B + T$ . According to the stability of the image background, a low-rank matrix  $B$  may be launched, so that the recovery of the background matrix  $B$  of the image matrix  $F$  can be attributed to recover low rank matrix of an observation matrix

Image block is the smallest unit of the described background image in background set. Whether the input image block belongs to the background set, it can be expressed as follow:

For a collection of elements, if a new element belongs to the set, it may be a linear combination of the elements in the set, the elements of the background block can be represented by a linear combination of the background block set, this process can be defined as follows:

- Definition 2: A set of background blocks  $X = [x_1, x_2, \dots, x_n]$  is given, where  $x_1, x_2, \dots, x_n$  is the background block. If the new input block  $y$  belongs to the set  $X$ , it can be represented by a linear combination of:  $y = a_1 * x_1 + a_2 * x_2 + \dots + a_n * x_n$ , there is the problem for judging whether image block belongs to the background set, it is solved by solving the equation  $y = Ax$

**Construction of background set based on robust principal component analysis:** According to the aforementioned problems about the background model, which are obtained from the sample video sequence,  $k_t$  can be transformed to recover a low rank matrix from the observation matrix, a question assume is given, it is the  $m \times n$  observation matrix  $M$ , which can be broken down into the Eq. 1:

$$M = L + S \quad (1)$$

where,  $L$  is an  $m \times n$  matrix with low rank,  $S$  is a sparse matrix of the  $m \times n$ . If the observation matrix  $M$  is known only, a low rank matrix  $L$  of the observation matrix  $M$  is recovered by solving the Eq. 2:

$$\min \text{rank}(L) + \gamma \|S\|_0; M = L + S \quad (2)$$

where,  $\text{rank}(L)$  is on behalf of the rank of the matrix  $L$ ,  $S_0$  represents zero-norm of matrix  $S$ . Since the problem is related to the matrix rank and matrix zero-norm, it is not the only solution.

Robust Principal Component Analysis (RPCA) was proposed by Candès and Tao (2006), which is the analysis and solving method of the above problem. The RPCA algorithm is convex optimization method, the above problem (2) is converted to solve the Eq. 3:

$$\min \|L\|_* + \lambda \|S\|_1; M = L + S \quad (3)$$

where,  $\|L\|_*$  is the trace norm of matrix  $L$ , it is the singular value sum of the matrix  $L$ ;  $S_1$  is the  $l_1$ -norm of matrix  $S$ . Thus, the matrix trace is obtained by solving the matrix rank convex optimization problem, so that the problem can be the only solution.

For the solution of the Eq. 3, Augmented Lagrange Multiplier (ALM) method is used in this study (Lin *et al.*, 2009), a new scalar unknowns and the Lagrange multiplier matrix  $Y$  is introduced, the Eq. 3 is converted to the Eq. 4:

$$l(L, S, Y) = \|L\|_* + \lambda \|S\|_1 + \langle Y, M - L - S \rangle + \frac{\mu}{2} \|M - L - S\|_F^2 \quad (4)$$

where,  $\|M - L - S\|_F$  is Frobenius norm, which is the square root of the matrix element squared sum. For low rank and sparse matrix factorization problem, there is a relatively simple and effective solution in solving:

$$\min_L x(L, S, Y)$$

and:

$$\min_S l(L, S, Y)$$

A contraction operator function  $S_\tau(t) = \text{sign}(t) \max(|t| - \tau, 0)$  is defined, it roles each element in the matrix, Eq. 5 can be obtained:

$$\arg \min_S l(L, S, Y) = S_{\lambda/\mu}(M - L + \mu^{-1}Y) \quad (5)$$

Similarly, a matrix arithmetic function  $D_\tau(X) = US_\tau(\Sigma)V^*$  is defined,  $X = U\Sigma V^*$ , Eq. 6 can be obtained:

$$\arg \min_S l(L, S, Y) = D_\mu(M - S - \mu^{-1}Y) \quad (6)$$

In the case, S is fixed, the minimum value of the function l is obtained on L, then L is fixed, the minimum value of the function l is obtained about S. According to the M-L-S, the Lagrange multiplier matrix Y is updated until it meets the iteration convergence conditions, a low-rank matrix L is obtained.

**Image block analysis based on sparse representation:** Ultra complete base is a set of base, the number of base is larger than dimension of based elements. Based on sparse representation theory, the ultra perfect base is used to represent the image, which can be the most likely sparse image representation and higher resolution information can be obtained than traditional non-adaptive methods (Chen *et al.*, 1998).

Linear equations  $y = Ax$  can be solved to describe sparse representation of ultra-complete base, where A is  $m \times n$  ( $m < n$ ) matrix, x is a n-dimensional vector, y is a m-dimensional vector. The idea of the sparse representation makes the equation x sparse as possible,  $\|x\|_0$  is as small as possible, it is expressed as follow with a mathematical expression:

$$\min \|x\|_0; \text{ s.t. } y = Ax \tag{7}$$

Due to this problem involves a matrix zero-norm, non-convexity of zero-norm makes Eq. 7 become solving NP-hard combinatorial optimization problems, solving is quite difficult. Candes and Tao (2006) proved that under RIP conditions,  $l_0$  norm optimization problem and  $l_1$  norm optimization problems have the same solution, that Eq. 7 can be converted to solve the Eq. 8:

$$\min \|x\|_1; \text{ s.t. } y = Ax \tag{8}$$

Equation 8 is a convex optimization problem, which can be solved by linear programming algorithm.

The obtained n frame background images are decomposed into the  $N \times N$  image blocks, which are not un-overlapping, the same  $N \times N$  block position in all background blocks of n frames is changed to n column vectors by the raster scan mode and the n vectors are cascaded into matrix  $A_i$  (i is the block label). For a new input image, it will also be decomposed into non-overlapping image blocks of  $N \times N$  and each image block is scanned into a column vector  $y_i$  by raster, then ideally, the background block matrix  $A_i$  is corresponding new input  $y_i$ , relationship between two images can be expressed by  $y_i = A_i x$  equation. In practice, since the noise will cause some errors, therefore  $y_i = A_i x$  is rewrote as follows (Mallat and Zhang, 1993):

$$y_i = A_i x + \varepsilon \tag{9}$$

It is further written in matrix form:

$$y_i = Ax_i = [A_i, I] \begin{bmatrix} x \\ e \end{bmatrix} \tag{10}$$

wherein,  $A_i$  is the background template; I is an  $(N \times N) \times (N \times N)$  two-dimensional matrix, it is called trivial template (Zhao *et al.*, 2011); x represents the background template coefficients; e represents trivial coefficient template or noise coefficients (Gao *et al.*, 2012).

According to the idea of sparse representation theory, the following conclusions can be obtained:

- If the image block  $y_i$  is belonging to the background input block matrices  $A_i$ , the value of the coefficient vector  $x_i$  should be mainly the background template coefficient  $x$  in the Eq. 10 and the  $x$  element value is not greater than 1
- If the input image block  $y_i$  does not belong to the background block matrices  $A_i$ , the obtained coefficient vector  $x_i$  value should be mainly in the trivial template coefficient  $e$  and because trivial template is a unit matrix, the element value of the  $e$  is the pixel value of each pixel in the input block, it is generally greater than 1

Thus, the background template coefficient of the coefficient vector  $x_i$ , the range and distribution of trivial template coefficient values are analyzed, we can know that the input image block is part of the background or foreground.

For the input image block process, the background image is determined by the above method, the image block pixels are directly set to 0, this clears the background. If the input block does not belong to the background image, the coefficient value of the trivial background template minus multiplication of background template coefficient and background matrix and then it is reverted to the block, this realizes the extracted foreground.

**Algorithm steps:** The main steps to build a background set are algorithms based on RPCA:

- Step 1:**  $n$  frames of images with  $r \times l$  are read in turn, each frame image will be expanded as a column vector and in order of priority, the observation matrix  $E$  is constituted
- Step 2:** ALM algorithm is used to get the background matrix  $B$
- Step 3:** To restore the background matrix  $B$  operation, they will be restored to  $n$  frame  $r \times l$  images
- Step 4:** Each of the background image set is divided into  $N \times N$  non-overlapping blocks
- Step 5:** The image block of different images at the same position is expanded as the column vectors and they are added in the trivial cascade template to form a matrix  $A_i$

The main steps of block analysis method with sparse representation are as follows (Fig. 2):

- Step 1:** The input image is also divided into  $N \times N$  non-overlapping blocks
- Step 2:** An image block of the input image is expanded into a column vector  $y_i$  in the raster scan manner
- Step 3:** The corresponding matrix  $A_i$  is input in the background set
- Step 4:**  $l_1$  norm minimization method is used for solving equations  $y_i = A_i x$ , the solution size and distribution of the value  $x$  are analyzed, it is obtained whether the input image block belongs to the background set
- Step 5:** If the input image block belongs to the background set, its pixel value will be removed, if the input image block does not belong to the background set, the corresponding differential treatment is continued
- Step 6:** Restore to block, appends the following to get final test results

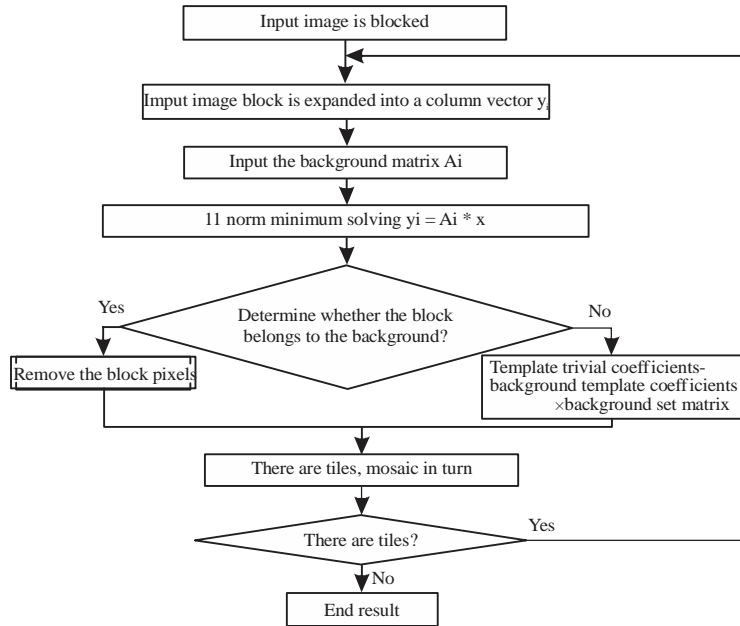


Fig. 2: Implementation flow analysis based on block sparse representation method

## RESULTS AND DISCUSSION

**Simulation of analysis algorithm based on block sparse representation:** To verify this study, the target detection method and sparse representation are used, it can be accurately distinguished in the input image whether the block belongs to the background set, the following simulation is carried out.

The OTCBVS benchmark database video is selected to do simulation experiments, the size of the video sequence is  $240 \times 320$  in the simulation experiment, the block side length is  $N = 8$ , the training video sequence frames is  $n = 50$ . Therefore,  $A$  is a two-dimensional matrix of  $64 \times 114$ , the calculated coefficient vector  $x_i$  is a 114-dimensional vector, which 50 values in front are the background template coefficients, the 64 values behind are trivial template coefficients. The simulation results are shown in Fig. 3: the input image (a) is a grayscale image, the upper left corner of the figure labeled building block represents the  $8 \times 8$  pixel block foreground, the characters which is marked on the lower right corner represent the  $8 \times 8$  pixel block of the background; when the input vector  $y_i$  in Fig. 3b is equal to the background block in Fig. 3a, a histogram values of the coefficient vector  $x_i$  is obtained by solving Eq. 10; Fig. 3c is the foreground block when the input vector  $y_i$  is equal to Fig. 3a, a histogram of the coefficient vector  $x_i$  values is obtained by solving the Eq. 10.

In Fig. 3b and c, the abscissa represents the coefficient vector index, the ordinate represents the size of the element corresponding to subscript. The 0-49 on the horizontal axis is the background template range, 50-114 range is trivial template range.

**Target detection experiment simulation:** In order to verify the test effectiveness of the proposed method, two simulations are made by our algorithm, Gaussian mixture algorithms, codebook algorithm. Experiment consists of a set of outdoor illumination equalization surveillance video and a set of outdoor affected video due to clouds. In the surveillance video experiments of



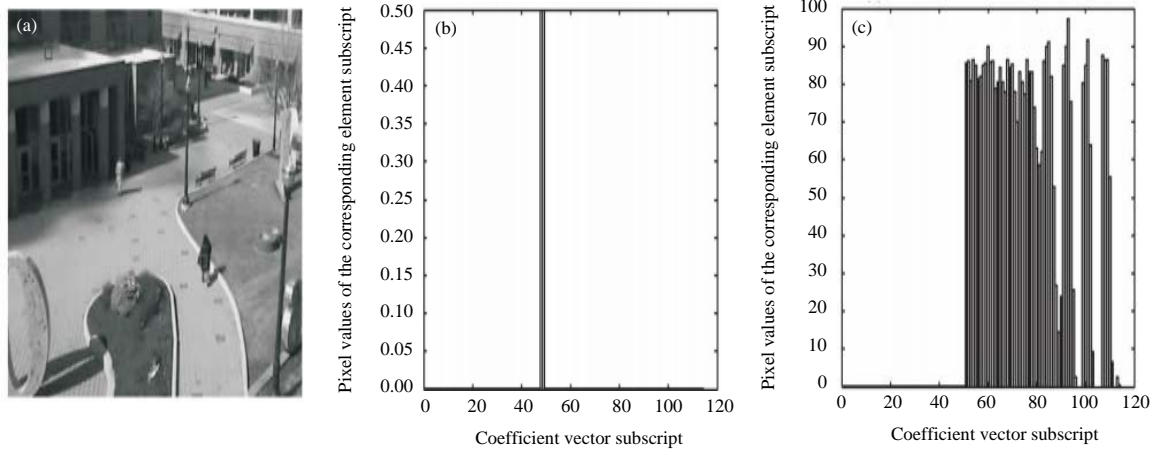


Fig. 3(a-c): Block analysis method simulation based on sparse representation, (a) Gray-scale image of the input frame, (b) Coefficient of solutions when the input block is the background block and (c) Coefficient of solutions when the input block is the foreground block

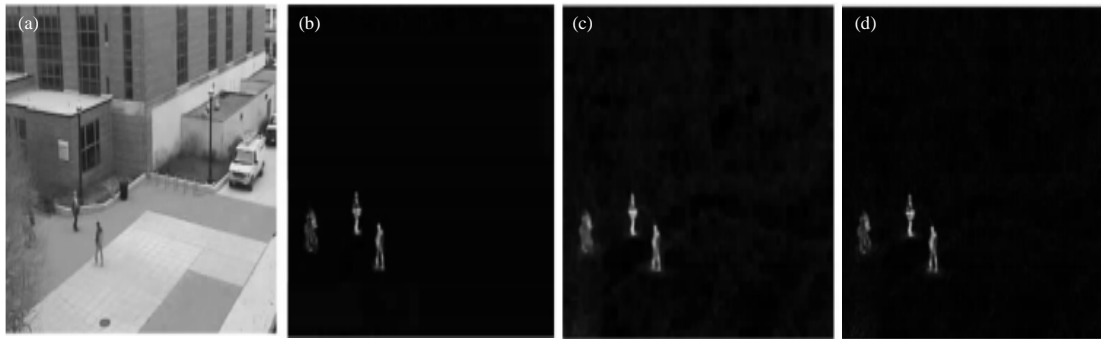


Fig. 4(a-d): Detection effect contrast in outdoor lighting balanced case, (a) Input image, (b) Our algorithm test results, (c) Gaussian mixture algorithm test results and (d) Codebook algorithm test results

strong illumination changes, the number of target pixels and the pixel number of false detection point are statisticised for each algorithm, the detection rate and false detection rate are calculated (Xu *et al.*, 2011), which is used as an indicator for the three algorithms:

$$\text{Hit rate} = \frac{\text{Detected target pixel points}}{\text{Actual target pixel points}}$$

$$\text{False detection rate} = \frac{\text{False detected pixel points}}{\text{All detected pixel points}}$$

Figure 4 is the detection effect visual comparison chart of outdoor lighting balanced situation in three target algorithm. Meanwhile, a random sample of 30 images are extracted in

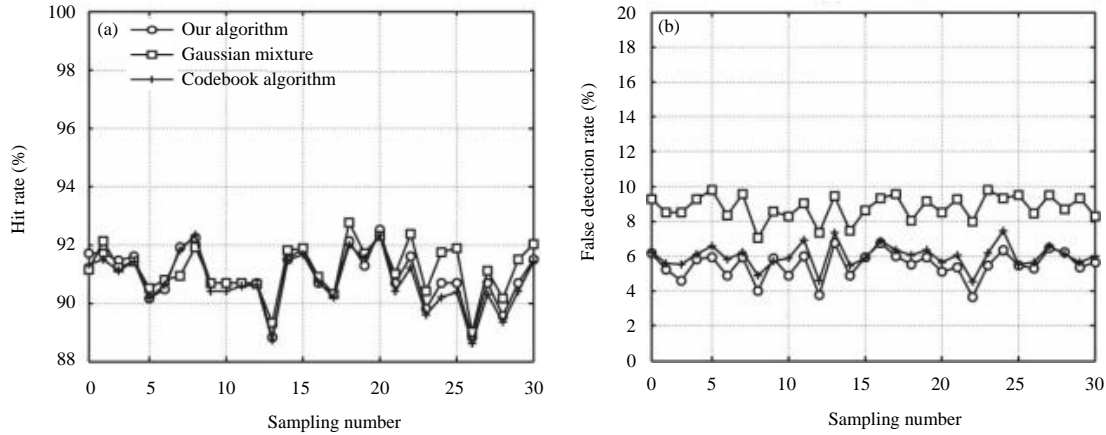


Fig. 5(a-b): (a) Hit rate and (b) False detection rate comparison of three algorithms in outdoor lighting balanced case

this set of experiments and the hit rate and false detection rate are calculated with three target detection algorithm, the statistical results are shown in Fig. 5.

As can be seen from the above tests and in the light of a balanced situation, the proposed algorithm has the same ideals of detection with other two algorithms.

Figure 6 is video frames under illumination change with influence of moving clouds strongly. Figure 6a, c, e is respectively 15th frame, 20th frame, 25th frame grayscale image in the test video, it can be seen from these figures that due to the influence of clouds, there is ambient light mutant situation in a video sequence. Figure 6b, d and f is, respectively three-dimensional histogram for 15th frame, 20th frame, 25th frame gray image pixel values, wherein, x-axis represents the image column, y-axis represents the image lines, z-axis represents the pixel value of the image pixel, it can be seen from the figure that in the region of the corresponding light mutation, columnar distribution of pixel values has changed significantly.

Figure 7 is intuitive comparison chart of object detection effects with the three algorithms and in outdoor lighting changes. Similarly, a sample of 30 images are extracted randomly in this set of experiments and the hit rate and false detection rate were calculated by three target detection algorithms, the statistical results are shown in Fig. 8.

## DISCUSSION

As is seen in Fig. 3b, with respect to the background block, the value of the coefficient vector are mainly distributed in front of the background template coefficients and the size value of the vertical axis is less than 1 and it is in Fig. 3c, for values of the foreground block, the coefficient vector is mainly distributed in the back of the trivial template coefficients, the value of the vertical axis represents the majority of the size, which is greater than 1. The simulation results of Fig. 3 illustrates that the visual analysis method of block sparse representation can determine that the input block is part of the background or foreground.

Intuition effect of Fig. 7 is conjunction with statistical result of Fig. 8, the target of three algorithms are more accurate determination in strong outdoor lighting changes, the missed case is not exist, so the hit rate is higher. However, the Gaussian mixture method is poor adaptability for rapidly changing scene and in false positive rate and therefore there is a high false alarm rate;

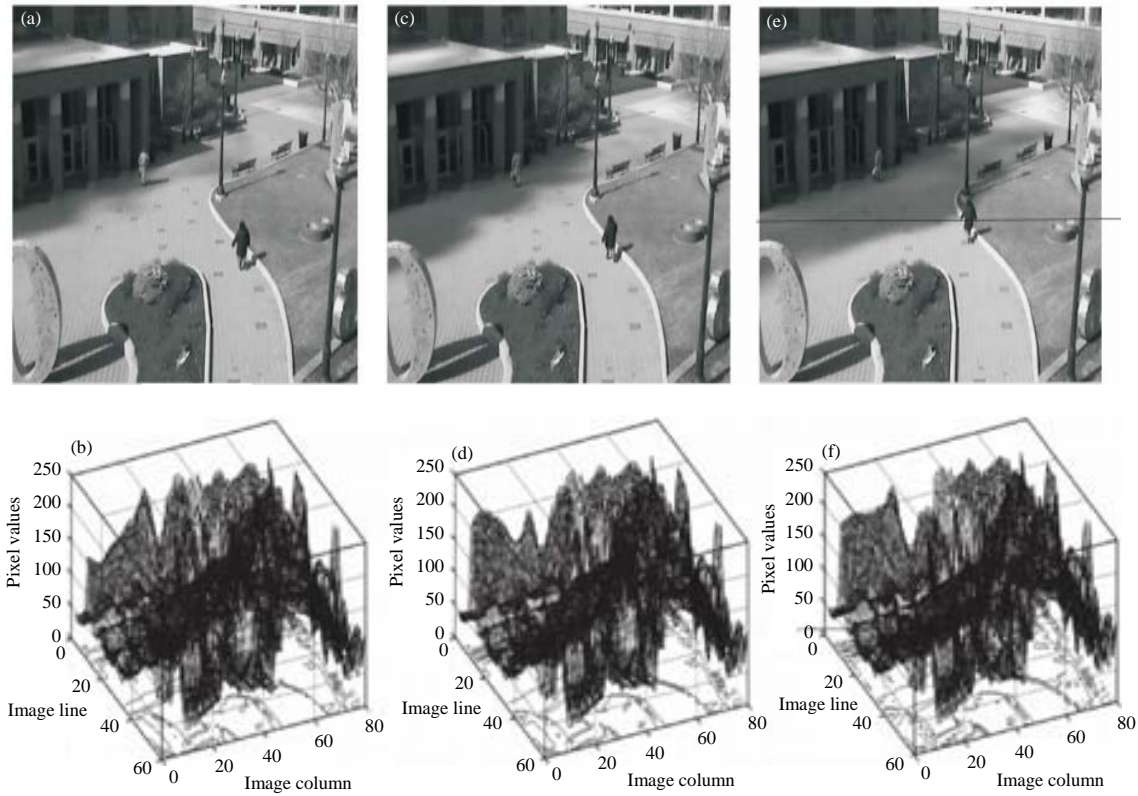


Fig. 6(a-f): Light mutation test video, (a) 15th frame grayscale image, (b) Three dimensional histogram of 15th frame image, (c) 20th frame grayscale image, (d) Three dimensional histogram of 20th frame image, (e) 25th frame grayscale image and (f) Three dimensional histogram of 25th frame image



Fig. 7(a-d): Contrast of detection effect in case of outdoor lighting changes, (a) Input image, (b) Our algorithm test results, (c) Gaussian mixture algorithm test results and (d) Codebook algorithm test results

codebook modeling algorithm is not capable of accurate information with codebook and for the case of the illumination mutant. The false detection rate is also higher and a set is established for the background image if the proposed algorithm is used. If the blocks of collection are used to describe

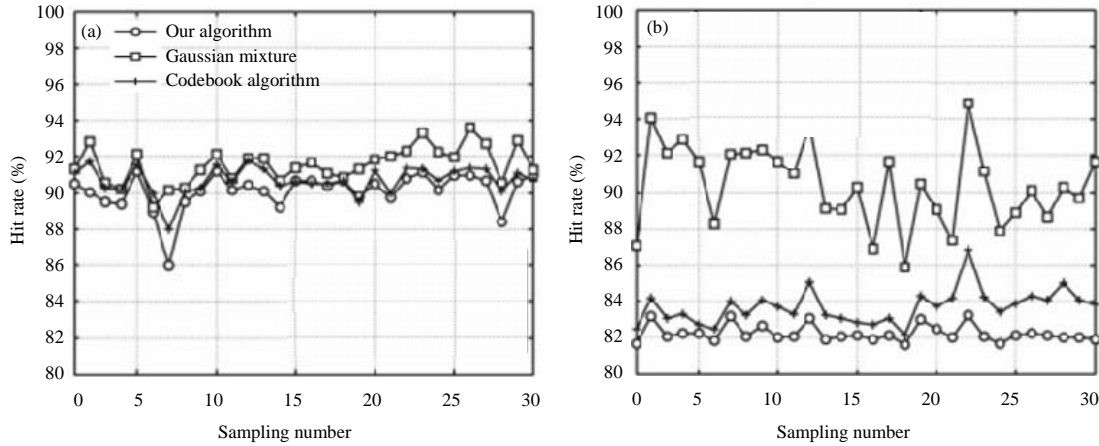


Fig. 8(a-b): Three algorithms, (a) Hit rate and (b) False detection rate comparison in outdoor lighting changes

Table 1: Processing time analysis of three algorithms

Algorithm	Processing frames	Consuming (s)	Average consuming per frame (s)
Gaussian mixture algorithm	60	225.18	3.75
Codebook algorithm	60	303.96	5.07
Our algorithm	60	695.58	11.59

the background, it avoids light sensitive issue of a single pixel, the impact of the illumination changes are effectively eliminated on target detection, false detection rate is much lower than the other two algorithms, the better detection results are obtained.

Meanwhile, the time complexity of the three algorithm is analyzed and compared in this study, the image sequence of the same period video are taken in OTCBVS benchmark database, three algorithms were used for testing, whose image size is 240×320, the three algorithm processing times are, respectively statistised, The results are shown in Table 1.

As can be seen from Table 1, the real-time is better in Gaussian mixture algorithm and codebook algorithm, but this algorithm relates to a large matrix, the operation takes a long time. Simulation results show that the method of the present study can better handle background ambient lighting mutation, the ambient noise was effectively eliminated on target detection, however, the algorithm is related to the operation of a large matrix, it takes a long time, it will be the next focus of the study how to ensure high accuracy at the same time have a good real-time performance.

## CONCLUSION

Moving object detection means that interest moving objects are extracted from the video images, the location of the moving target image is determined. This process can also be seen as to separate the target pixel of motion video image and background pixels. Due to various disturbances of scene change, uncertainty exists two main categories in the target extraction: uncertainty spatial location and undefined pixel values. Spatial position uncertainty is mainly because of camera shake in mobile scenarios, the fluctuating background scene is caused by the lens shift; undefined pixel values mainly refers to fixed scenes. Due to various disturbances, the pixel value change of the background image portion is uncertain situation. Two type uncertainties will lead to a significant decline in target detection accuracy. At present, background subtraction has been widely used in

target detection, it exists a typical pixel values of uncertainties, which is mainly to remove the background pixels through the image sequence frame of the current frame and the background subtraction, separation is achieved between foreground pixels and background pixels. When there is interference in the scene, if the foreground contains a false background pixels in the difference results, the target detection error rate is higher.

When the light changes occurred against the background, the background modeling methods based on the existing single-pixel can not effectively detect the moving objects, the present study was designed and a moving target detection method was implemented based on background image collection and analysis of the sparse block method, RPCA build background image collection was used, the separation of foreground blocks and background blocks was achieved by block analysis method of sparse representation and the moving object extraction was ultimately realized.

## REFERENCES

- Candes, E., X. Li, Y. Ma and J. Wright, 2010. Robust principal component analysis?: Recovering low-rank matrices from sparse errors. Proceedings of 2010 IEEE Sensor Array and Multichannel Signal Processing Workshop, October 4-7, 2010, Jerusalem, pp: 201-204.
- Candes, E.J. and T. Tao, 2006. Near-optimal signal recovery from random projections: Universal encoding strategies? IEEE Trans. Inform. Theory, 52: 5406-5425.
- Chen, S.S., D.L. Donoho and M.A. Saunders, 1998. Atomic decomposition by basis pursuit. SIAM J. Sci. Comput., 20: 33-61.
- Gao, C., T. Zhang and Q. Li, 2012. Small infrared target detection using sparse ring representation. IEEE Aerospace Electron. Syst. Magaz., 27: 21-30.
- Kim, K., T.H. Chalidabhongse, D. Harwood and L. Davis, 2004. Background modeling and subtraction by codebook construction. Proceedings of the International Conference on Image Processing, Volume 5, October 24-27, 2004, Singapore, pp: 3061-3064.
- Lin, Z., M. Chen and Y. Ma, 2009. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. Math. Programm., 22: 55-78.
- Mallat, S.G. and Z. Zhang, 1993. Matching pursuits with time-frequency dictionaries. IEEE Trans. Signal Process., 41: 3397-3415.
- Stauffer, C. and W.E.L. Grimson, 2000. Learning patterns of activity using real-time tracking. IEEE Trans. Pattern Anal. Mach. Intell., 22: 747-757.
- Xu, X.M., Q. Mo, L. Ni, Q.Y. Guo and A. Li, 2011. Moving object detection based on local updated layered codebook. J. Comput. Applic., 31: 3399-3402.
- Zhao, J.J., Z.Y. Tang, J. Yang, E.Q. Liu and Y. Zhou, 2011. Infrared small target detection based on image sparse representation. J. Infrared Millimeter Waves, 30: 156-161.