Arab J Sci Eng (2017) 42:3621–3633
DOI 10.1007/s13369-017-2672-2

RESEARCH ARTICLE - COMPUTER ENGINEERING AND COMPUTER SCIENCE

Foreground Detection via Background Subtraction and Improved Three-Frame Differencing

Sandeep Singh Sengar[1] · Susanta Mukhopadhyay[1]

Received: 1 February 2016 / Accepted: 8 June 2017 / Published online: 19 June 2017

Abstract Moving object detection is a widely used and important research topic in computer vision and video processing. Foreground aperture, ghosting and sudden illumination changes are the main problems in moving object detection. To consider the above problems, this work proposes two approaches: (i) improved three-frame difference method and (ii) combining background subtraction and improved three-frame difference method for the detection of multiple moving objects from indoor and outdoor real video dataset. This work accurately detects the moving objects with varying object size and number in different complex environments. We compute the detection error and processing time of two proposed as well as previously existing approaches. Experimental results and error rate analysis show that our methods detect the moving targets efficiently and effectively as compared to the traditional approaches.

Keywords Moving object detection · Background subtraction · Frame differencing · Threshold · Morphology

1 Introduction

Foreground detection is an extensively used task in computer vision and artificial intelligence applications such as activity recognition [1], video surveillance [2], robotics [3], video retrieval [4]. It concerns how to take out moving objects from video frames and remove the background region and noise. The accurate detection of moving objects not only gives perfect result, but also avoids unnecessary computation for the wrong foreground objects. Moving objects are generally detected with the help of motion, color or shape features. However, it should be appropriate for complex circumstances, like fake motion, illumination variation, background clutter, night detection and Gaussian noise. All the above-mentioned problems lead to the detection of wrong area of the foreground objects. To resolve these problems, numerous methods have been presented, like optical flow [5], frame differencing [6] and background subtraction [7]. The optical flow method is based on the properties of flow vector of the object over time to detect moving object regions. However, this method is computationally complex. Usually, the detection of foreground objects using optical flow method is susceptible to noise, fake motion and illumination variation. While frame differencing is state forward and widely used method for moving object detection and segmentation, it utilizes the difference between consecutive video frames to detect moving objects. This approach is computationally less complex and appropriate for dynamically changing environments, but sometimes cannot detect all of the relevant foreground objects due to ghosting and foreground aperture problems. On the other hand, the background subtraction approach is a widely used method for this purpose. It comprises of two steps: (i) background modeling and (ii) computation of difference between the current background model and the current video frame. This method is susceptible to problem arising out of lighting and inappropriate events. Therefore, background image has to be updated regularly. It is direct and easy method to detect moving objects and extensively used due to its stability in dynamic environments and good real-time performance. However, background subtraction approach is susceptible to camera jitter and illumination change.

✉ Sandeep Singh Sengar
sandeep.iitdhanbad@gmail.com

[1] Department of Computer Science and Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad, Jharkhand 826 004, India

The three-frame differencing approach is proposed by Kameda et al. [8] to reduce the effect of ghosting and foreground aperture in two-frame differencing. This method can be implemented using the difference operation between frame at time t and $t-1$ as well as difference operation between frame at time t and $t+1$. However, it does not give accurate results with slowly moving objects and uniformly distributed pixels. In order to improve foreground detection result and reduce the false positives, we propose the improved three-frame differencing method and also present another approach with the help of our improved three-frame difference method and background subtraction.

The main novelties of our proposed work with respect to [9] and previously existing methods are as follows:

1. To detect the foreground objects, three-frame differencing approach generally combines two consecutive difference frames. In the existing techniques, traditional three-frame difference approach [8] uses logical 'AND' operation and the approach proposed by Fei et al. [9] performs pixel-wise sum to combine the difference frames. However, these techniques cause to loss in foreground information or generation of the false positives. To overcome the aforementioned problems, we have estimated the pixel-wise maximum operations to combine the difference frames (see Sect. 3.1).
2. To obtain the maximum amount of foreground object information with less noise and to reduce the computational complexity, we employ the 'MAX' operation directly between the difference frames instead of binary frames (see Sect. 3.1).
3. To estimate the accurate initial background frame for the background subtraction technique, we use pixel-wise median operation among frames. However, the method proposed by Fei et al. [9] takes the first frame as an initial background frame (see Sect. 3.2.1). Furthermore, the approach in [8] has not used the concept of background subtraction.

The structure of remaining part of this article is as follows. Section 2 presents the survey on related methodologies. In Sect. 3, we describe the proposed *improved three-frame differencing* as well as *combining background subtraction and improved three-frame differencing* techniques for foreground object detection. Experimental results with the qualitative and quantitative analysis are provided in Sect. 4. Finally, Sect. 5 gives the conclusion of our work.

2 Related Methodology

Lots of works have been done to deal with several complex scenarios in moving object detection, namely out-of-plane rotation, ghosting, object pose variation, occlusion, lighting changes and background distraction. To enhance the quality of detection of moving objects, numerous techniques with the help of nonparametric kernel density estimation have been built up [10]. Christian et al. [11] proposed a foreground detection approach to improve the results with non-static background. The method [12] develops both models (foreground and background) with the help of spatiotemporal reference data. The work [13] improves the object detection results using nonparametric techniques. However, this method is computationally complex and requires huge memory. Li et al. [14] have used the spatial features such as gradient and color to improve the quality of the background model for the non-moving camera. With the help of special stereo input, stereo-based segmentation [15] acquired the good results by fusing color, contrast and stereo matching. Olivier et al. [16,17] presented some techniques for target detection by comparing the background model with the current frame. This technique adopts neighboring pixels to build the background model. However, if the foreground and background have similar color (i.e., camouflage problem), then it will not detect the foreground accurately and will also be slow in removing ghost regions. Huang et al. [18] proposed a Bayesian method for moving object detection. This method uses spatial, temporal and spectral features to characterize the background appearance. Object motion is computed and utilized for moving object detection [19–21]. A foreground detection method for non-stationary background is proposed by Hu et al. [22]. This technique has used the Harris corner detector and the optical flow for feature detection and matching. Next, on the basis of multiple view geometry, these feature points are classified as either background or foreground. This method is good for real-time applications, but it does not give effective results for fast-moving objects. The region-matching-based motion estimation technique is presented by Ghosh et al. [23] to get the moving object from non-stationary background video. This method retained the object boundary for segmentation with the help of fuzzy edge strength of each pixel. Wang [24] proposed a technique to detect moving vehicles using a Markov random field model in different weather conditions. However, this technique works well only for grayscale videos. Yang et al. [25] presented an effective moving object detection approach by combining Gaussian mixture model, three-frame difference and cropped frame techniques. It can detect moving objects with an initialization method based on dynamic grid and density estimation.

3 Foreground Detection

In this paper, first we propose an improved three-frame differencing method. Next, we combine our improved method with the background subtraction approach to accurately detect the target.

3.1 Improved Three-Frame Differencing Method

This method is an improvement over the traditional three-frame differencing [8] as well as the one proposed by Fei et al. [9]. The traditional three-frame differencing method uses a double difference frame to detect the foreground object. The double difference frame is obtained by logical ‘AND’ operation between two successive threshold-based difference frame. The threshold-based difference frame can be computed by applying the difference operation between frames at time t and $t-1$ as well as difference operation between frames at time t and $t+1$ and, subsequently, performing the threshold operation on both the difference frames. The schematic diagram of traditional three-frame differencing is shown in Fig. 1.

If there is no significant texture content in moving objects, then this traditional approach does not correctly detect the foreground. Furthermore, if the speed of the moving objects is slow, then logically ‘AND’ operation in traditional three-frame difference method leads to loss of information in final foreground objects. Fei et al. [9] also proposed a three-frame differencing approach using ‘SUM’operation instead of ‘AND’ operation. It is an improvement over the traditional three-frame differencing method to combine the important information from the multiple frames of the same scene. Furthermore, it can also reduce the noise in the low-texture environments. However, this method has the deficiency due to the thresholding operation on the difference frame directly before applying the ‘SUM’ operation. So here if we choose the high threshold value, then the important information will be lost from each difference frames, and for the low threshold value this method will generate the noisy binary frames. Next, the binary frames with either lost information or noisy information will be added using the ‘SUM’ operation in the subsequent step and will produce inaccurate results. For that if we combine both the difference frames using ‘MAX’ operation before the thresholding operation, then we can get all the important information with less or negligible amount of noise. In the further step, we can apply the thresholding operation to detect the foreground. To keep this concept in mind, we incorporate an improved three-frame differencing approach. This proposed approach has two improvements over the previously mentioned methods as we have made: (i) use of ‘MAX’ operation in place of logical ‘AND’ and ‘SUM’ operation; ‘MAX’ operation can find the pixel-wise maximum value between two difference frames, and there is no loss of information; (ii) application of ‘MAX’ operation directly between the ‘difference frames’ instead of the ‘binary frames’ so as to detect the maximum amount of foreground object information.

Our improved three-frame differencing algorithm can be summarized in following steps:

1. Application of preprocessing operation to remove random noise from video sequences using Gaussian filter.
2. Estimation of difference frame $D_{t,t-1}(x,y)$ between frames F_{t-1}(x, y) and $F_t(x,y)$ and second difference frame $D_{t,t+1}(x,y)$ between frames $F_t(x,y)$ and $F_{t+1}(x,y)$. Here $F_{t-1}(x,y)$, $F_t(x,y)$ and $F_{(t+1)}(x,y)$ are three successive video frames.
3. Computation of pixel-wise maximum intensity value between two difference frames $D_{t,t-1}(x,y)$ and $D_{t,t+1}(x,y)$ to obtain the moving object frame *I(x,y)*.
4. Determination of binary image frame *BW(x,y)* from the moving object frame *I(x,y)* using thresholding method.
5. Accurate estimation of moving objects by post-processing with the help of morphological operation.

The schematic diagram of our improved three-frame differencing approach is given in Fig. 2.

3.2 Combined Approach

To detect the foreground objects in the current frame, we present a hybrid technique that uses both improved three-frame differencing and background subtraction. Background subtraction is a widely used simple technique for foreground detection in real-time video processing. The background subtraction approach uses the background frame that does not contain any moving object. We construct the current

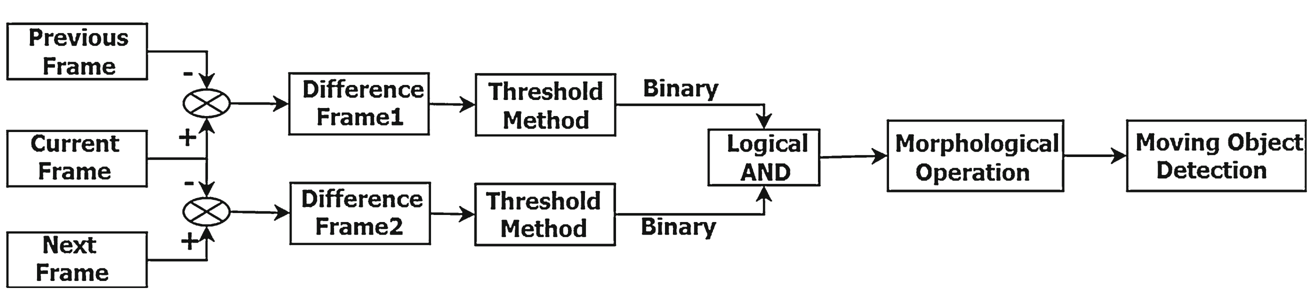


Fig. 1 Schematic diagram of traditional three-frame differencing approach

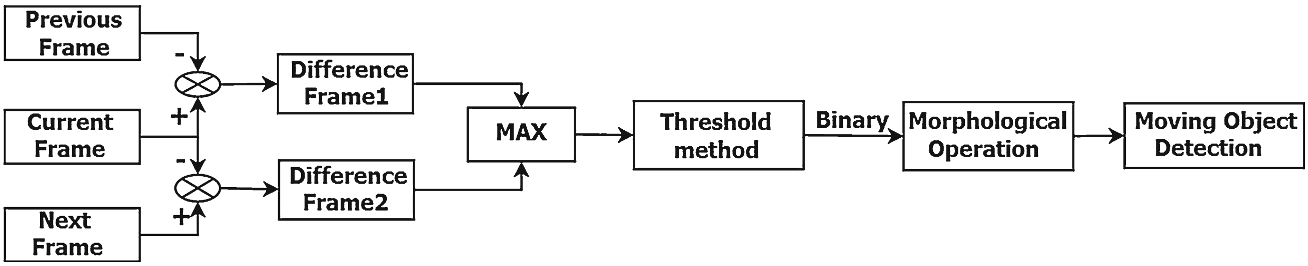


Fig. 2 Schematic diagram of our improved three-frame differencing approach

background frame using the background modeling technique (see Sect. 3.2.1). For implementing the combined approach, first we find the difference frame between the current frame and the previous frame. We also find the difference frame between the current frame and the background frame. Subsequently, we detect the foreground objects by finding the pixel-wise maximum intensity value between both the difference frames constructed in previous steps. Steps for combining background subtraction and three-frame differencing are summarized in Sect. 3.2.2. This approach is simple for foreground detection and has an improvement over the method proposed by Fei et al. [9].

3.2.1 Background Modeling

Background subtraction is the general approach to segment moving objects, which involves background modeling and updating. Background modeling must be able to recognize the correct background frame. Fei et al. [9] have used the first frame of a video sequence as an initial background for background subtraction approach. However, as shown in Fig. 5, the first frame may also contain moving objects; thus, we cannot get a correct background using this concept and detected foreground object will not be accurate.

To tackle the above-mentioned problem, we calculate the initial background model BG_0 using pixel-wise median operation as given in Eq. 1.

$$\mathrm{BG}_0(p) = median(F_i(p)) \tag{1}$$

Here $i = 1, 2, 3 \ldots, N$ and N is the number of video frames used to construct the first reference background frame. $F_i(p)$ is the intensity of pixel p of the ith frame.

In dynamic scenes, the background models need to be updated time to time to adapt the changes in the environment and correctly reproduce the current state of background. We use the following simple steps to update the background frame:

1. Computation of difference frame between current frame and current background frame.
2. Construction of binary image frame from difference frame using thresholding operation.
3. Filtering of binary frame using morphological operations.
4. Updating the background frame according to current background, if resultant binary image frame contains pixel value zero at particular position; otherwise, the background frame is updated using the current frame.

Representative frames of the tested video sequences and their corresponding background frames detected by our approach are displayed in Figs. 3, 4, respectively.

3.2.2 Steps for Combining Background Subtraction and Improved Three-Frame Differencing Approach

We use the improved three-frame differencing approach and updated background image frame to detect the significant prominent foreground objects. The procedure is summarized as follows:

1. Frame smoothing using Gaussian filter.
2. Estimation of difference frame $\mathrm{Diff}_{t,t-1}(x, y)$ between previous frame $F_{t-1}(x, y)$ and current frame $F_t(x, y)$ to extract the moving objects.
3. Estimation of second difference frame $\mathrm{Diff}_b(x, y)$ between current frame $F_t(x, y)$ and current background frame BG_t to find the variations of foreground objects.
4. Computation of foreground object frame $F_f(x, y)$ by taking the maximum pixel intensity value between $\mathrm{Diff}_{t,t-1}(x, y)$ and $\mathrm{Diff}_b(x, y)$.
5. Determination of binary frame *B(x,y)* from $F_f(x, y)$ using thresholding operation.
6. Accurate estimation of moving objects by post-processing using morphological filtering.

The schematic diagram of this approach is given in Fig. 6.

4 Experimental Results and Analysis

The methods proposed in this work are exclusively focused to get better quality, the speed and the usability of the back-

Fig. 3 Representative frames of tested video sequences. **a** Video 1, **b** Video 2, **c** Video 3, **d** Video 4, **e** Video 5, **f** Video 6

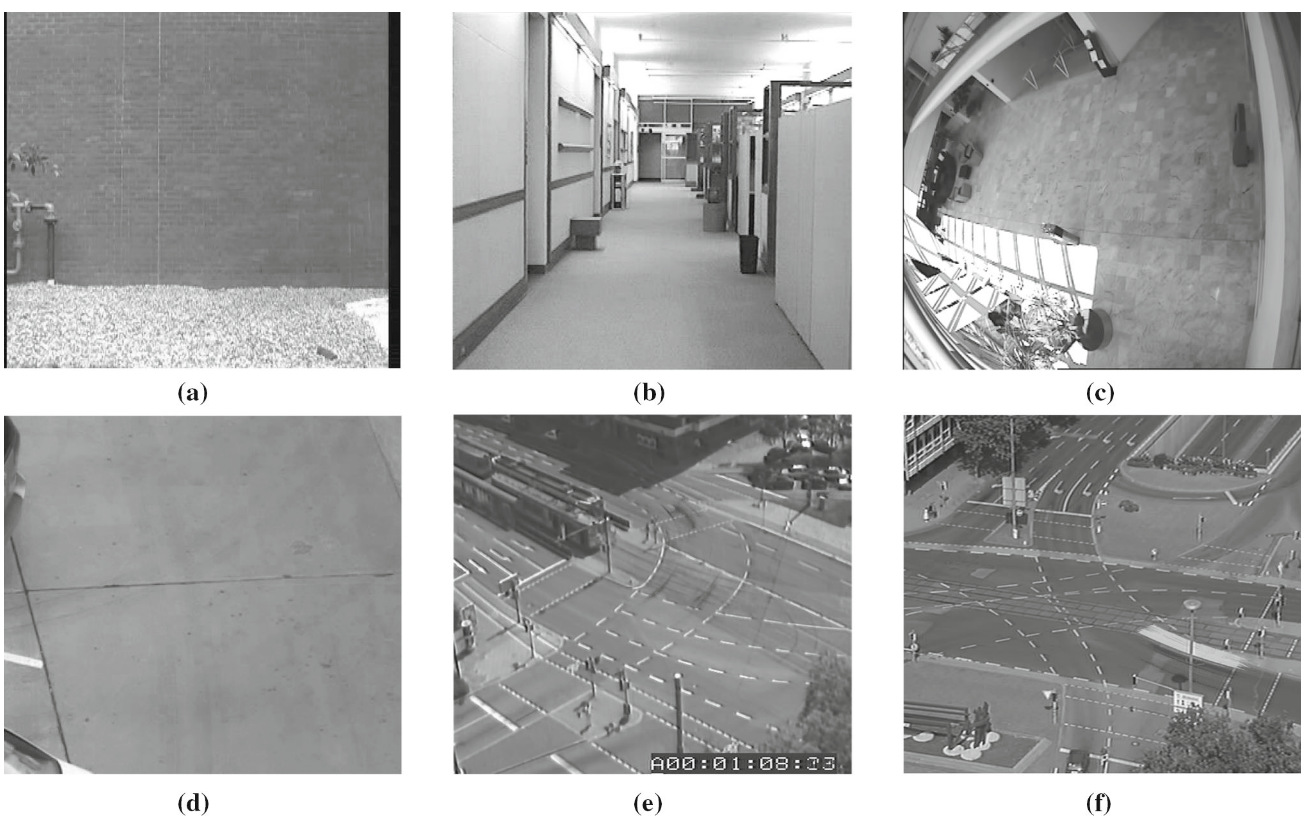


Fig. 4 Background frames obtained using background modeling. **a** Video 1, **b** Video 2, **c** Video 3, **d** Video 4, **e** Video 5, **f** Video 6

(a) (b) (c)

Fig. 5 Initial background frames used in Fei et al. [9]. **a** Video 1, **b** Video 5, **c** Video 6

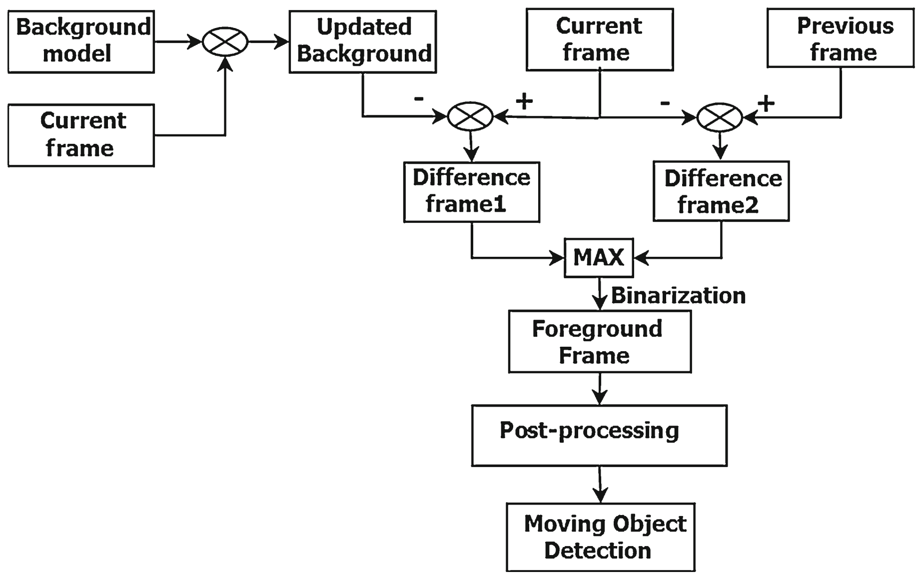


Fig. 6 Schematic diagram of proposed combining three-frame differencing and background subtraction approach

ground modeling in target moving object detection strategies. We have applied our methods on several complex indoor and outdoor standard benchmark video sequences containing critical situations such as illumination variation, background clutter, slow or fast motion and varying number as well as size of moving objects. The description of each tested video sequences is as follows:

– Video 1[1] shows the slowly moving person from one side to another. The frame size of this video is 240 pixels × 368 pixels. In this video, moving object has bright color than background color.
– Video 2[2] shows the movement of two persons from one place to another in opposite directions. One person picks up the box from table, simultaneously other person puts the briefcase on the other table. The frame size of the used sequence is 240 pixels × 352 pixels. This indoor video has a certain degree of illumination variations, and the color of the moving objects and background is very similar.
– In the video 3[3] two persons come from opposite directions and shake hands to each other and then go together. This video has complex background, and it suffers from illumination variation. The frame size of the used sequence is 384 pixels × 288 pixels.
– Video 4[4] represents two persons in the outdoor scene, they fight, shake hands and meet each other. The color of the target objects is very similar that of the background. The frame size of the video sequence is 352 pixels × 288 pixels.
– Video 5[5] shows the road traffic with varying number of moving cars, and people. These videos represent very complex background with dark and light regions, and

[1] https://vid.me/videodata.

[2] http://see.xidian.edu.cn/vipsl/database_Video.html.

[3] http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/.

[4] http://cvrc.ece.utexas.edu/SDHA2010/.

[5] http://clickdamage.com/sourcecode/cv_datasets.php.

color of some moving objects is similar to that of the background. The frame size of this video sequence is 512 pixels × 512 pixels.
- Video 6[6,7] also represents the complex road traffic with people along with moving vehicles with different size. Some vehicles are moving with high speed, while some vehicles have color similar to the background color. The frame size of the used sequence is 512 pixels × 512 pixels.

To establish the effectiveness of our approaches, the experimental results are compared with some of the existing techniques using both the qualitative (Sect. 4.1) and quantitative (Sect. 4.2) analysis.

4.1 Qualitative Analysis

We have selected some of the representative video frames for implementation purpose, and the experimental results are provided in Figs. 7, 8 and 9, which display the original video frames and foreground images obtained from five methods of Video 1 (the 15th, 72th, 126th frame), Video 2 (the 74th, 155th, 216th frame), Video 3 (the 316th, 334th, 433th frame), Video 4 (the 841th, 974th, 1220th frame), Video 5 (the 3th, 21th, 32th frame) and Video 6 (the 3th, 31th, 48th frame), respectively. The first row of Figs. 7, 8 and 9 represents the original frames, and next four rows (from top to bottom) of these figures show the results by inter-frame differencing [9], traditional three-frame differencing [8], our improved three-frame differencing and the one proposed by Fei et al. [9], respectively. The last row of the above-mentioned figures presents the results using the method of *combining background subtraction and improved three-frame differencing* proposed by us. In the following, we discuss the detection results of each tested video depending on the different challenging attributes.

4.1.1 Illumination Variations

Due to the significant illumination changes in the Video 2 and Video 3, the inter-frame differencing and traditional three-frame differencing approaches provide the inaccurate result with lots of false negative and high amount of noise (see columns d–f of 2nd and 3rd rows in Figs. 7, 8). Our improved three-frame differencing approach outperforms these two methods (see columns d–f of 4th row in Figs. 7, 8), but the results produced by Fei et al. [9] are more accurate than our improved approach due to less amount of noise (see 5th row of Video 2 and Video 3 in Figs. 7, 8). However, there are false positives (see column d of 5th row in Fig. 7 and columns d–f of 5th row in Fig. 8) and the upper portion of one person in the frame 155 of Video 2 is not detected (see column e of 5th row in Fig. 7). As shown in columns d–f of 6th row in Fig. 7 and columns a–c of 6th row in Fig. 8, our proposed combining approach reduces the shortcoming of existing techniques for these sequences and produces more efficient and accurate results with less false alarm.

4.1.2 Background Clutter

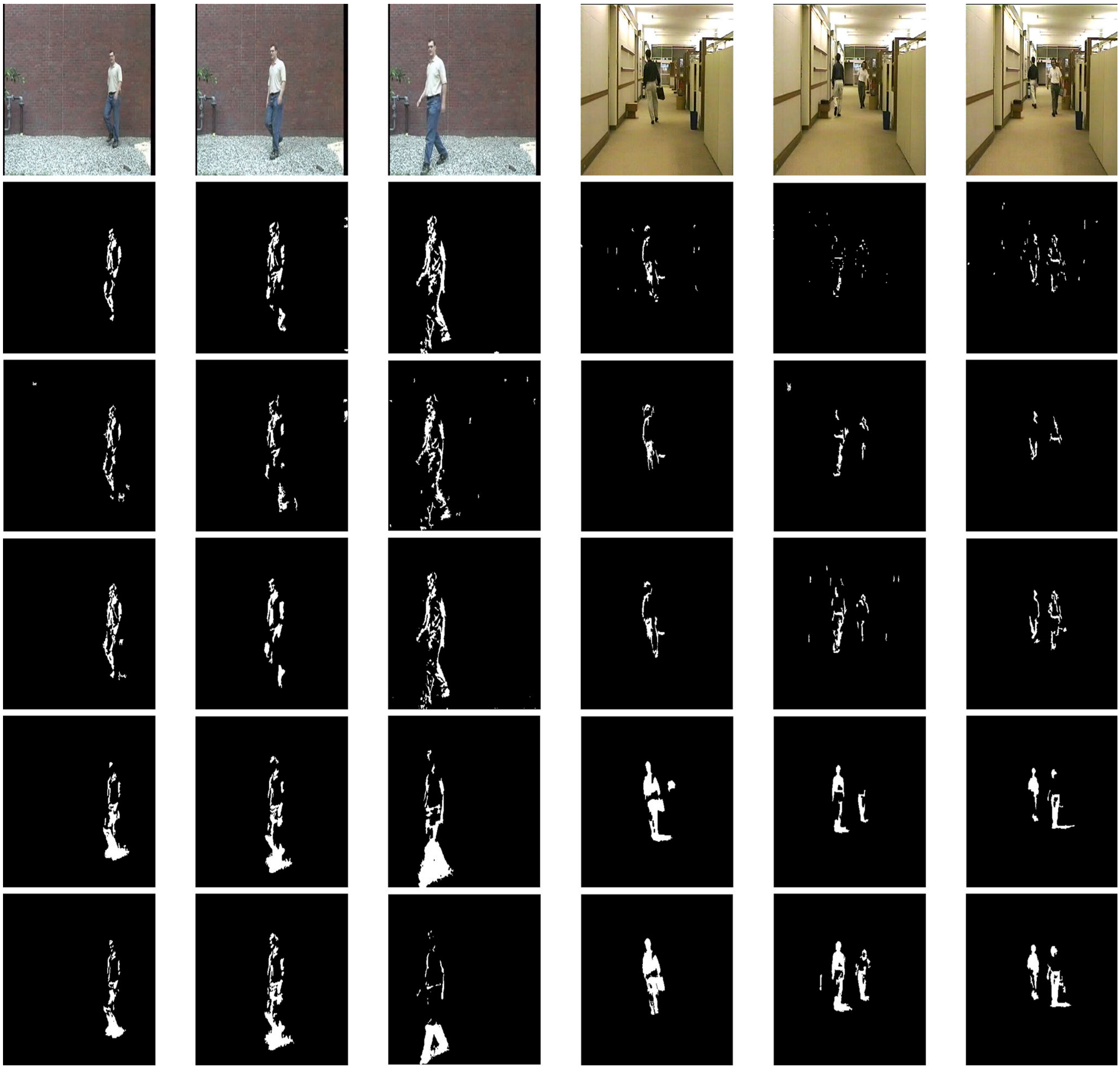
The color or texture information of the foreground object in the Video 2, Video 4, Video 5 and Video 6 is very similar to the background (see Fig. 3). As shown in columns d–f of 2nd, 3rd rows in Fig. 7, columns d–f of 2nd, 3rd rows in Fig. 8 and 2nd, 3rd rows in Fig. 9, inter-frame differencing and three-frame differencing methods face the problem of foreground aperture and generate the results with lots of false negatives. However, our method in the 4th row of aforementioned figures improves the performance, but with some amount of noise. The method proposed by Fei et al. [9] reduces the false negatives and detected results are more accurate, but with false positives. However, the number of false negatives in the Video 2 and Video 6 is significantly higher in this approach (see columns e, f of 4th row in Fig. 7 and column f of 4th row in Fig. 9). Furthermore, our proposed combining approach also has false negatives, but results of our approach are considerably more accurate than the [9].

4.1.3 Slow or Fast Motion

The inter-frame differencing and traditional three-frame differencing methods (i) suffer from the ghosting problem due to slow speed of person in Video 1 (see columns a–c of 2nd and 3rd rows in Fig. 7), and the fast speed of moving vehicles in Video 5 and Video 6 (see 2nd and 3rd rows in Fig. 9) (ii) also generates the false positives. Furthermore, the results produced by Fei et al. [9] in the Video 1, Video 5 and Video 6 are not much accurate and generate large false positives as well as false negatives (see columns b, c of 5th row in Fig. 7 and column f of 5th row in Fig. 9). However, the proposed combining approach reduces the false positives in all the aforementioned video sequences (see columns a–c of 6th row in Fig. 7 and 6th row in Fig. 9), but this approach produces the false negatives in Video 1. Thus, our combining approach does not perform well for Video 1. However, our improved three-frame differencing method (see columns a–c of 4th row in Fig. 7) outperforms all the tested approaches and produces considerably accurate results with low false positives as well as false negatives for Video 1.

[6] http://www.cvpapers.com/datasets.html.

[7] In this section and throughout the paper, the aforementioned resources are used.

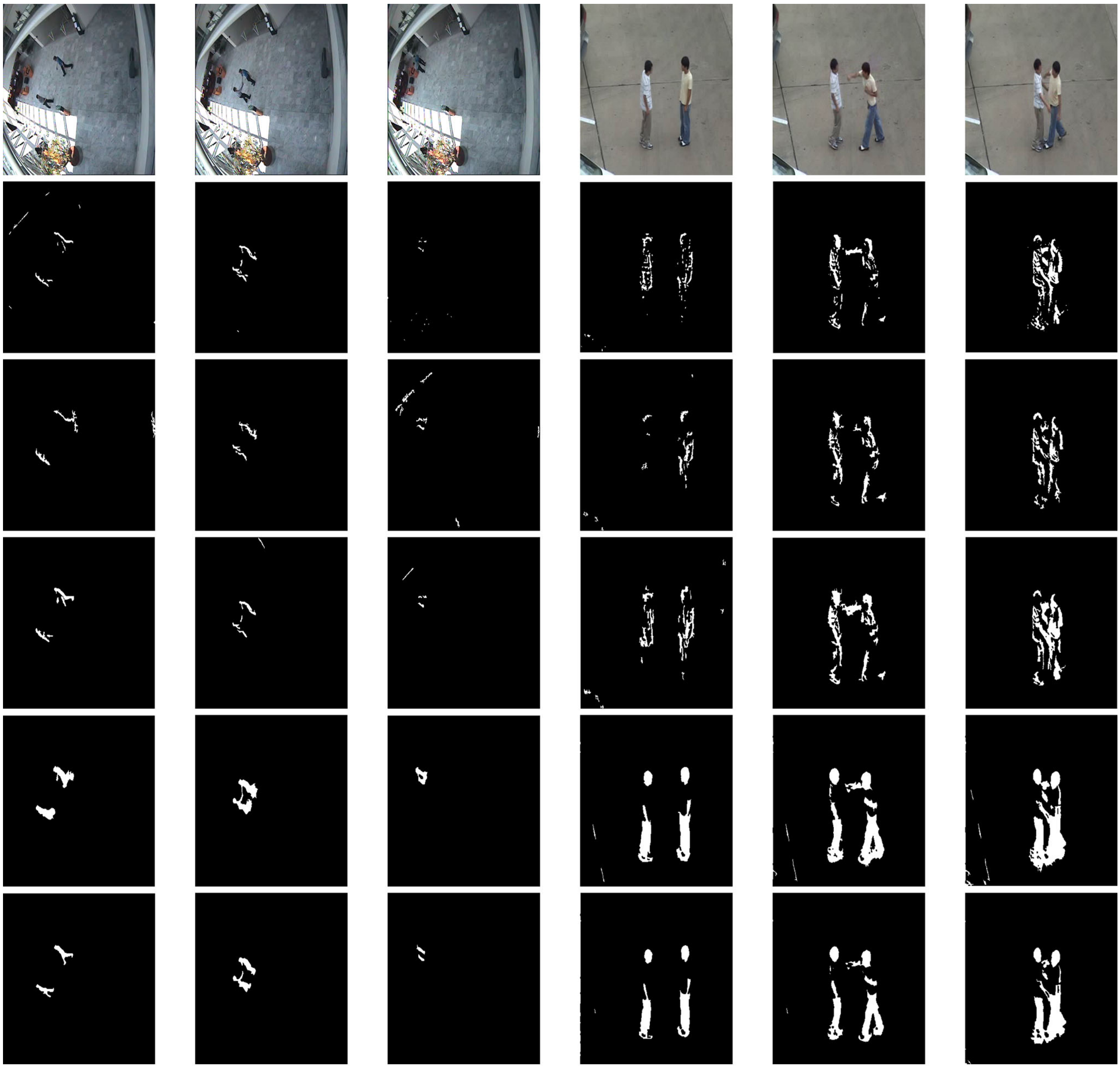


(a) (b) (c) (d) (e) (f)

Fig. 7 Comparison of proposed moving object detection methods with other existing approaches using Video 1 and Video 2, where **a** 15th frame of Video 1, **b** 72th frame of Video 1, **c** 126th frame of Video 1, **d** 74th frame of Video 2, **e** 155th frame of Video 2, **f** 216th frame of Video 2; *row wise*, *top to bottom*: original frames, inter-frame differencing, traditional three-frame differencing [8], our improved three-frame differencing, approach proposed by Fei et al. [9], our proposed combining approach

4.1.4 Varying Number and Size of Objects

The size of moving objects in the Video 3, Video 5 and Video 6 is small, and they are varying with the object movement. There are varying number of moving vehicles in the Video 5 and Video 6 (see Fig. 3), in which Video 5 also contains person with slow motion. In the aforesaid video sequences, both the inter-frame differencing and three-frame differencing approaches do not detect the moving objects accurately and generate the noise as well (see columns a–c of 2nd, 3rd rows in Fig. 8 and 2nd, 3rd rows in Fig. 9). Furthermore, our improved three-frame differencing methods outperform these techniques and detect the objects more accurately, but also have certain amount of noise. The approach proposed by Fei et al. [9] reduces the noise in Video 3. However, it also generates some false pos-



(a) (b) (c) (d) (e) (f)

Fig. 8 Comparison of proposed moving object detection methods with other existing approaches using Video 3 and Video 4, where **a** 316th frame of Video 3, **b** 334th frame of Video 3, **c** 433th frame of Video 3, **d** 841th frame of Video 4, **e** 974th frame of Video 4, **f** 1220th frame of Video 4; *row wise*, *top to bottom*: original frames, inter-frame differencing, traditional three-frame differencing [8], our improved three-frame differencing, approach proposed by Fei et al. [9], our proposed combining approach

itives; thus, the detected objects are not much accurate (see columns a–c of 5th row in Fig. 8 and 5th row in Fig. 9). Next, our proposed combining method outperforms all the tested techniques and detects the objects more accurately for these datasets (see 6th row for above-mentioned figures).

It has been shown from the qualitative analysis that our improved three-frame differencing method has attained better integrity of moving objects than inter-frame differencing as well as traditional three-frame differencing in all six videos. The second method proposed by us is also superior to the approach proposed by Fei et al. [9] for all six complex video sequences because our proposed method is susceptible to object movement: Even small objects with small movement can be detected in the different challenging situations.

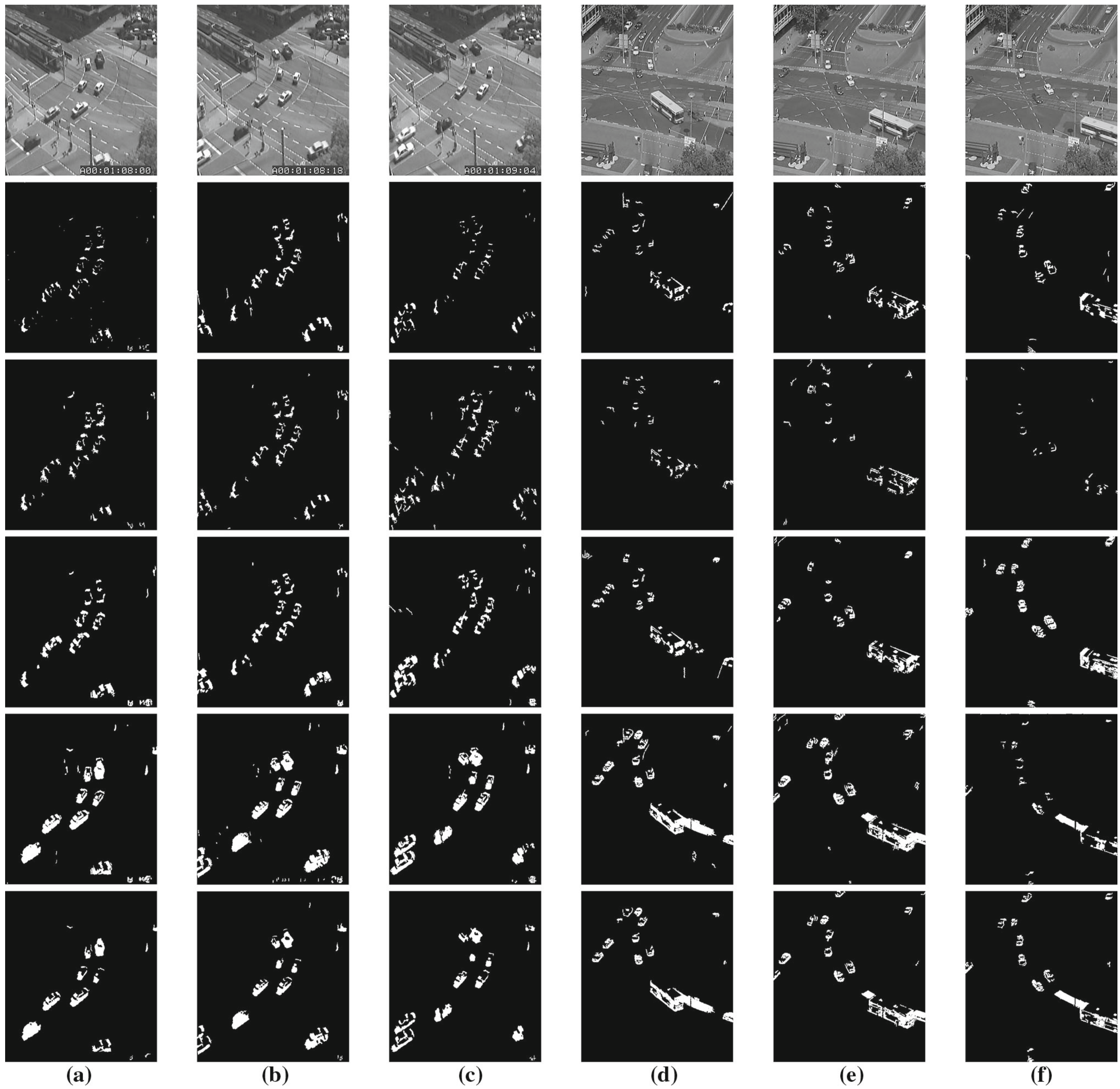


Fig. 9 Comparison of proposed moving object detection methods with other existing approaches using Video 5 and Video 6, where **a** 3th frame of Video 5, **b** 21th frame of Video 5, **c** 32th frame of Video 5, **d** 3th frame of Video 6, **e** 31th frame of Video 6, **f** 48th frame of Video 6; *row wise*, *top to bottom*: original frames, inter-frame differencing, traditional three-frame differencing [8], our improved three-frame differencing, approach proposed by Fei et al. [9], our proposed combining approach

4.2 Quantitative Analysis

To effectively compare the performance of our proposed methods with previous approaches, first we define the detection error rate (E^i) of each frame as well as average detection error rate ($\bar{E}$) of complete video sequences using the following equations.

$$E^i = \left| N_f^i - N_o^i \right| \Big/ N_o^i \tag{2}$$

$$\bar{E} = \sum_{i=1}^{N} E^i \Big/ N \tag{3}$$

Here N_f is the number of foreground objects detected by the moving object detection approach, N_o is the total number of

Table 1 Comparison of detection error rates of moving object detection methods in percentage

Approaches	Average detection error rate	
	Video 5	Video 6
Inter-frame differencing [9]	26.09	30.03
Traditional three-frame differencing [8]	15.12	37.86
Combining method proposed by Fei et al. [9]	10.53	18.27
Our improved three-frame differencing	15.10	22.54
Our proposed combining method	**6.78**	**11.48**

Detection error rate is lowest (best) for our combining method (in bold)

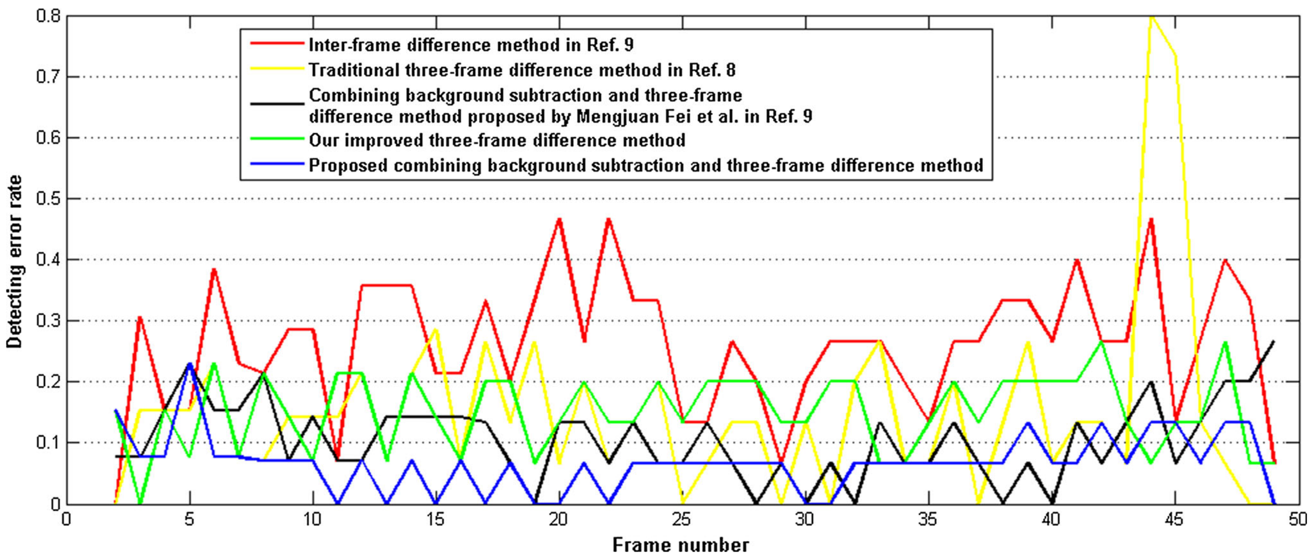


Fig. 10 Comparison of detection error rates of moving object detection methods with Video 5

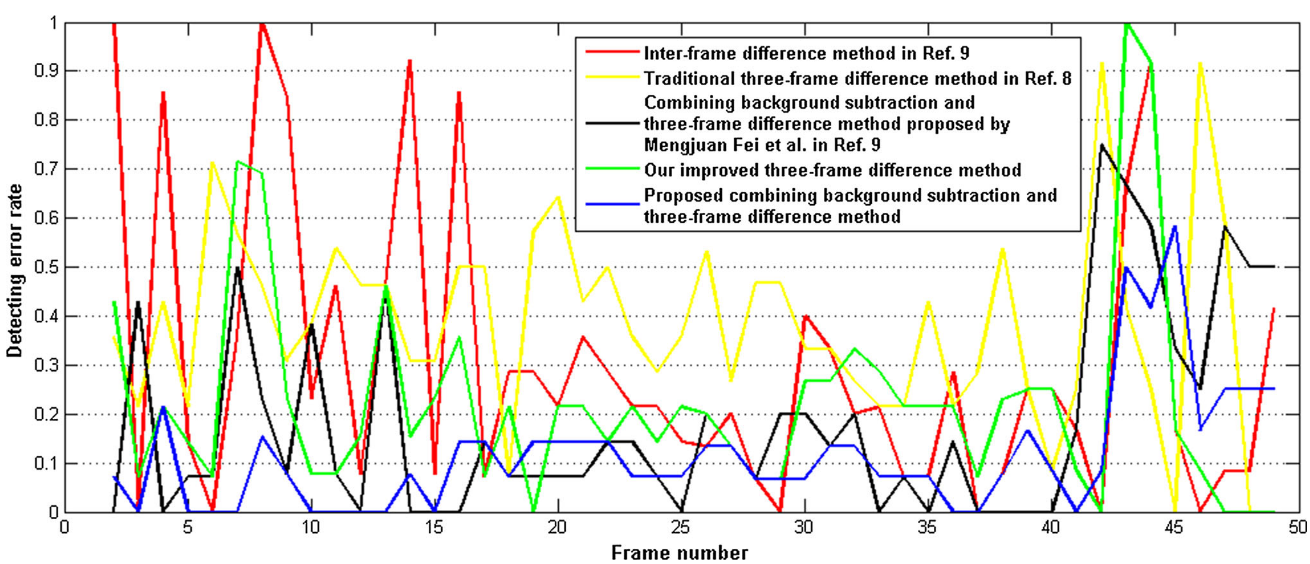


Fig. 11 Comparison of detection error rates of moving object detection methods with Video 6

truly moving objects in the original video frame, and N is the total number of frames in video.

We have selected the Video 5 and Video 6 for the quantitative analysis because these are the most complex among all six videos and contain multiple targets with varying size and numbers. Most of the target objects in these videos have almost same color as that of background. Table 1 gives the average detection error rate of each algorithm for Video 5 and Video 6.

Table 2 Comparison of processing time per frame of moving object detection methods in sec

Approaches	Average processing time per frame	
	Video 5	Video 6
Inter-frame differencing [9]	0.2418	0.2326
Traditional three-frame differencing [8]	0.3257	0.3192
Combining method proposed by Fei et al. [9]	0.3022	0.2888
Our improved three-frame differencing	**0.2386**	**0.2149**
Our proposed combining method	0.2502	0.2335

Average processing time of our another proposed improved three-frame differencing method is lowest (best) (in bold)

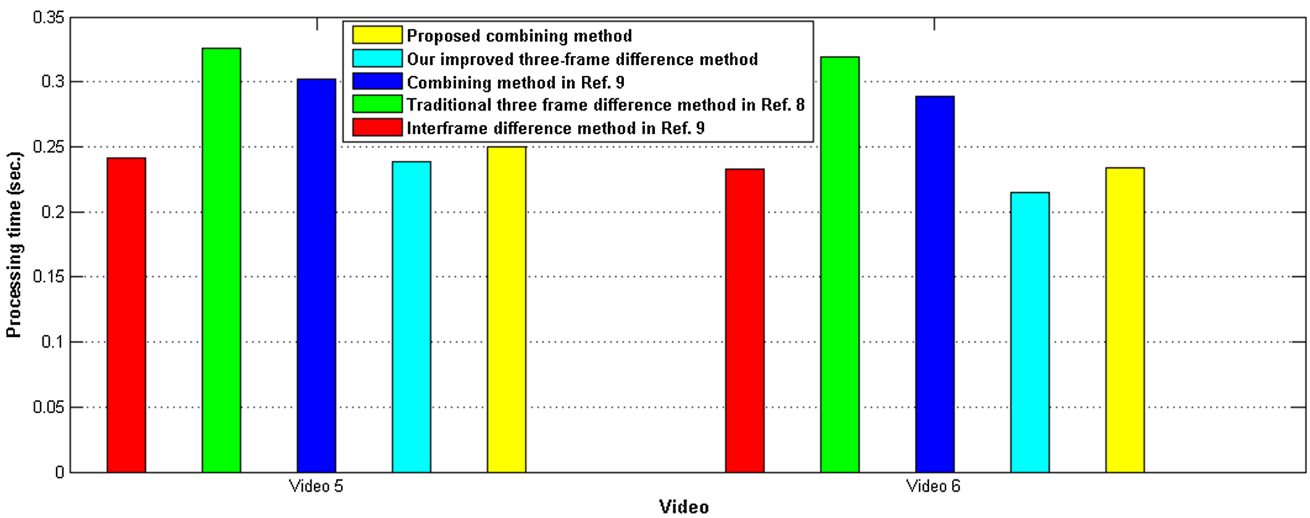


Fig. 12 Comparison of different methods in terms of processing time

Here the numeric values in bold indicate the best results. Figures 10 and 11 display the frame-wise detection error rate of each approach for Video 5 and Video 6, respectively.

It is observed that the improved three-frame differencing has low detection error rate than the inter-frame differencing as well as the traditional three-frame differencing method. Furthermore, the detection error rate of our proposed combining approach is less than those produced by the tested methods.

We also do the comparison in terms of processing time. Table 2 shows the average processing time per frame for the selected video sequences. The comparison in terms of processing time is also shown with the help of a bar graph in Fig. 12.

The performance analysis agrees that our approaches yield acceptable results as compared to previously existing methods; accuracy in terms of the detection error rate of the proposed methods is significantly higher; and the average processing time per frame is considerably lower than those of existing approaches. Furthermore, the processing time of inter-frame differencing is less than that required by our combining method. However, accuracy of our approach is relatively higher. Additionally, the methods proved to be robust under different critical situations.

5 Conclusions

In this work, we deal with the complicated task of moving object detection in complex scenes. Using the proposed approaches, we can detect almost all the foreground objects with fairly high accuracy. These methods have several enhancements than existing methods, viz. increased accuracy, less complexity and lesser computation time. In order to display its effectiveness, we have used a set of six indoor and outdoor real video sequences with different complex environments. Experimental results prove that our methods can handle slowly and fast-moving objects, ghosting, illumination changes, background clutter problems and variation of the object size as well as numbers with high accuracy. In terms of error rate detection and processing time, we have compared our proposed methods with three previously existing approaches. Qualitative as well as quantitative analysis show that our methods outperform the previously existing techniques. In future, we wish to upgrade our approach with non-stationary camera for object detection and tracking.

References

1. Zhang, Y.; Liu, X.; Chang, M.C.; Ge, W.; Chen, T.: Spatio-temporal phrases for activity recognition. In: Computer Vision-ECCV, vol. 7574, pp. 707–721. Springer, Berlin, Heidelberg (2012)
2. Sengar, S.S.; Mukhopadhyay, S.: Moving object tracking using Laplacian-dct based perceptual hash. In: International Conference on Wireless Communications, pp. 2345–2349. Signal Processing and Networking, IEEE (2016b)
3. Fu, K.S.; Gonzalez, R.; Lee, C.G.: Robotics: Control, Sensing, vision, and Intelligence. Tata McGraw-Hill Edition, Delhi (2008)
4. Aslandogan, Y.A.; Yu, C.T.: Techniques and systems for image and video retrieval. IEEE Trans. Knowl. Data Eng. **11**(1), 56–63 (1999)
5. Sengar, S.S.; Mukhopadhyay, S.: Moving object area detection using normalized self adaptive optical flow. Optik-Int. J. Light Electron Optics **127**(16), 6258–6267 (2016a)
6. Sengar, S.S.; Mukhopadhyay, S.: A novel method for moving object detection based on block based frame differencing. In: 3rd International Conference on Recent Advances in Information Technology, IEEE, pp. 462–472 (2016c)
7. Sobral, A.; Vacavant, A.: A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. Comput. Vis. Image Underst. **122**, 4–21 (2014)
8. Kameda, Y.; Minoh, M.: A human motion estimation method using three-successive video frames. In: ICVSM, pp. 135–140 (1996)
9. Fei, M.; Li, J.; Liu, H.: Visual tracking based on improved foreground detection and perceptual hashing. Neurocomputing **152**, 413–428 (2015)
10. Benezeth, Y.; Jodoin, P.; Emile, B.; Laurent, H; Rosenberger, C.: Review and evaluation of commonly-implemented background subtraction algorithms. In: Pattern Recognition, IEEE Int'l Conference, pp 1–4 (2008)
11. Cristani, M.; Farenzena, M.; Bloisi, D.; Murino, V.: Background subtraction for automated multisensor surveillance: a comprehensive review. EURASIP J. Adv. Signal Process. **2010**, 1–24 (2010)
12. Zhang, X.; Yang, J.: Foreground segmentation based on selective foreground model. Electron. Lett. IEEE **44**(14), 851–852 (2008)
13. Ko, T.; Soatto, S.; Estrin, D.: Background subtraction on distributions. In: Computer Vision-ECCV, Lecture Notes in Computer Science **5304**, pp. 276–289 (2008)
14. Li, L.; Huang, W.; Gu, I.; Tian, Q.: Foreground object detection from videos containing complex background. In: Multimedia 11th ACM International Conference, Berkeley, USA, pp. 2–10 (2003)
15. Kolmogorov, V.; Criminisi, A.; Blake, A.; Cross, G.; Rother, C.: Probabilistic fusion of stereo with color and contrast for bi-layer segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **28**(9), 1480–1492 (2006)
16. Barnich, O.; Droogenbroeck, MV.: Vibe: a powerful random technique to estimate the background in video sequences. In: Acoustics, Speech, and Signal Processing, International Conference, pp. 945–948 (2009)
17. Barnich, O.; Droogenbroeck, M.V.: Vibe: a universal background subtraction algorithm for video sequences. IEEE Trans. Image Process. **20**(6), 1709–1724 (2011)
18. Li, L.; Huang, W.M.; Gu, I.H.; Tian, Q.: Statistical modeling of complex background for foreground object detection. IEEE Trans. Image Process. **13**(11), 1459–1472 (2004)
19. Hosaka, T.; Kobayashi, T.; Otsu, N.: Object detection using background subtraction and foreground motion estimation. IPSJ Trans. Comput. Vision Appl. **3**, 9–20 (2011)
20. Wolf, C.; Jolion, JM.: Integrating a discrete motion model into gmm based background subtraction. In: Pattern Recognition, 20th IEEE International Conference, pp. 9–12 (2010)
21. Zhang, W.; Fang, X.Z.; Yang, X.: Moving vehicles segmentation based on bayesian framework for gaussian motion model. Pattern Recognit Lett. **27**(9), 956–967 (2006)
22. Hu, W.C.; Chen, C.H.; Chen, C.M.; Chen, T.Y.: Effective moving object detection from videos captured by a moving camera. In: Intelligent Data Analysis and Applications, Euro-China Conference **1**, pp. 343–353 (2014)
23. Ghosh, A.; Subudhi, B.N.; Ghosh, S.: Object detection from videos captured by moving camera by fuzzy edge incorporated markov random field and local histogram matching. IEEE Trans. Circuits Syst. Video Technol. **22**(8), 1127–1135 (2012)
24. Wang, Y.: Joint random field model for all-weather moving vehicle detection. IEEE Trans. Image Process. **19**(9), 2491–2501 (2010)
25. Yang, J.; Yang, W.; Li, M.: An efficient moving object detection algorithm based on improved GMM and cropped frame technique. In: Mechatronics and Automation, IEEE International Conference on, pp. 658–663 (2012)