STATISTICAL BACKGROUND MODELING BASED ON VELOCITY AND ORIENTATION OF MOVING OBJECTS

S. Sharifzadeh*, G. Ardeshir and H. Miar Naimi

Department of Computer and Electrical Engineering, Noushirvani Institute of Technology
University of Mazandaran, P.O. Box 47144, Babol, Iran
ssharifzadeh@stu.nit.ac.ir – g.ardeshir@nit.ac.ir - h_miare@nit.ac.ir

*Corresponding Author

(**Received: June 28, 2007 - Accepted in Revised Form: November 22, 2007)**

Abstract Background modeling is an important step in moving object detection and tracking. In this paper, we propose a new statistical approach in which, a sequence of frames are selected according to velocity and direction of some moving objects and then an initial background is modeled, based on the detection of gray pixel's value changes. To have used this sequence of frames, no estimator or distribution is required. In this work, smoothly changing pixels are modeled by averaging and other statistical calculations are done for more varying pixels. The proposed algorithm works well, even if the moving objects are present in all frames. We have evaluated this novel method successfully on highly textured scenes, with challenging phenomena such as dynamic background, area of high foreground traffic and moving objects with different speeds and sizes.

Keywords Background Modeling, Orientation, Velocity, Modified Set of Frames

چکیده در آشکارسازی و ردیابی اشیای متحرک، مدل سازی زمینه تصویر مرحله مهمی محسوب می‌شود. در این مقاله رویکرد آماری جدیدی در این زمینه ارائه می‌گردد. در روش پیشنهادی، رشته ای از فریم ها بر اساس سرعت و جهت اشیای متحرک انتخاب می‌شوند و سپس زمینه اولیه براساس آشکارسازی تغییرات مقادیر سطح خاکستری پیکسل ها مدل می‌شود. با بکارگیری چنین رشته فریم هایی، نیازی به استفاده از تخمین زننده ها و توزیع های آماری وجود ندارد. در این روش پیکسل هایی که به آرامی تغییر می‌کنند، از طریق میانگین گیری ساده مدل می‌شوند و سایر محاسبات آماری برای پیکسل هایی که تغییرات شدیدتری دارند، انجام می‌شود. الگوریتم پیشنهادی، حتی در شرایطی که اشیای متحرک در همه فریم ها حضور داشته باشند، به خوبی کار می‌کند. این رویکرد نوین را در صحنه هایی با بافت های متنوع تحت شرایط بحث بر انگیزی همچون زمینه هایی با ماهیت دینامیک، صحنه های شلوغ و اشیای متحرکی با ابعاد و سرعت های مختلف با موفقیت آزموده ایم.

1. INTRODUCTION

Background is the static part of a scene that does not have obvious changes along the time. For detection of moving objects, it seems easy to subtract a reference frame of the stationary background from the observed frame and then threshold the resulting image, and finally foreground objects could be extracted. However, since several factors such as lighting changes, camera noise and scenery movements are the elements effecting the changes in the background; this method does not work properly. One solution for this problem is to update background model in order to overcome these difficulties and effectively detect foreground objects.

The most common approach in background modeling was adaptive filters like Kalman and Wiener which predicted the intensity of pixels. These filters could not handle the problems of removed or introduced objects [1-3].

Other significant methods are statistical approaches in which a probability, according to a statistical model for each pixel, is obtained for a new pixel value, and then it is assigned to a background or foreground. Gaussian model has been widely used in this field [4-7].

Single Gaussian model (SGM) or mixed Gaussian models (MGM) for gray levels of images along the background updating and using on-line Expectation Maximization (EM) algorithm have been applied for surveillance applications, using color information of pixels. A new faster and more accurate method by improving the updating equations was proposed in [8]. A similar work in this field has been done in [9]. Each pixel is defined as layers of 3D multivariate Gaussian due to different appearance of each. A Bayesian approach was used to estimate probability distribution of mean and variance, rather than directly estimating these parameters for each layer.

Another statistical approach is nonparametric modeling of background. These algorithms reduce the dependency of background model to the distribution parameters and adapt the background model faster to the scene changes. Several samples of intensity values for each pixel are collected to estimate the density function [10,11]. Kernel density estimator computes the likelihood of new samples for each pixel. New samples are replaced with old ones in order to update the model.

Other works in this field deal with dynamic scenes in which, during learning stage the behavior of background is quantized and dynamic regions are modeled by aggregating the observations from the whole region [12,13].

Hidden Markov models are also used for background modeling. HMMs with fixed topology and parameter updating stage or with dynamic and time varying topology (Topology free HMM) were implemented [14,15].

In some recent works, on-line Adaboost (on-line version of boosting) was used for feature selection, which combines weaker classifiers to make a stronger one [16]. Classifiers are made in initializing learning stage assuming that all input fames contain (even changing) background. Then, the modeled background will be updated for new frames.

In a frame stream, the average of gray values for each pixel would belong to a background if it was occluded in a few frames. In other words, the quality and quantity of frames will affect the correct estimation of background pixel values. Variety of different scenes, often do not satisfy these conditions. Therefore Statistical models have been used to assign every new sample to foreground or background according to its probability. In our work, we chose a number of particular frames among input stream using movement information of the scene and therefore background modeling based on density of each gray value is improved. In this algorithm velocity and direction of moving objects have been used as a criterion to select a proper number of appropriate sequences of frames and then background is modeled by analyzing each pixel gray level in these frames. There is also an updating stage to handle scene changes. Due to selection of a particular sequence, more complicated calculations are done for significantly varying pixels while the smoother changing ones are modeled by a simple averaging.

This paper is organized as follows: The framework is introduced in Section 2. Estimation of direction, velocity of moving objects and the proposed background modeling algorithm are explained in this section. The experimental results are shown in Section 3.

2. BACKGROUND MODELING

In a sequence of frames, while no moving object is located in the pixel view, frequent changes in illumination especially in outdoor scenes, will smoothly vary the gray level of each pixel. On the other hand, prominent changes would occur if any movement happens in the visual field of the pixel. Figure 1 illustrates a street where cars are moving in different directions. Two pixels in different locations are shown in Figure 2a. Smooth changes of the first pixel and sharp changes of the second one through 100 frames are shown in 2b and 2c respectively. No moving object crosses the view of the former, but variations of the latter are considerable due to moving cars. It is not difficult to estimate the intensity of pixels like the first one. But the algorithm should be more accurate for those with prominent changes, to model background correctly.

To handle such pixels, a special sequence of frames should be selected among input stream so that in this sequence, moving objects haven't frequently occluded the pixel view when moving

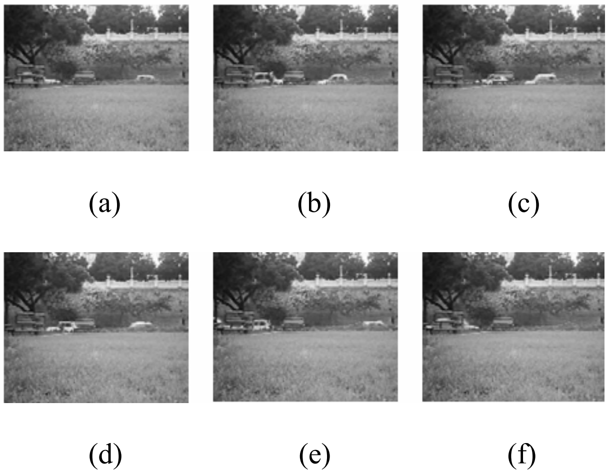


Figure 1. (a-f) show cars moving in a street in two directions.

in different directions. In other words, the pixel is not hidden in most frames. Consider the observation set of such pixel intensities in the mentioned sequence. Under above conditions, it is more probable that majority of gray values in the set belong to the background. Therefore, background modeling based on estimating each pixel gray value will be strongly improved.

However, there isn't such feature in every sequence of input stream, especially in scenes with several foreground objects, moving frequently in different directions. Therefore, it is necessary to find a way to provide a desired sequence of frames. Another important key is the number of frames so that slowest moving objects as a part of foreground would wholly move along their initial places. This requires information about their velocity.

2.1. Getting an Appropriate Sequence We shall prepare a sequence with suitable length and a particular direction of a major moving object. As discussed before, such sequence is the optimum choice, in order to model the whole background based on gray values of each pixel.

2.1.1. Orientation of moving objects In this section we explain how to distinguish the direction of moving objects in a frame sequence. The main idea is to consider variations in row or column numbers, of a moving object location in an image

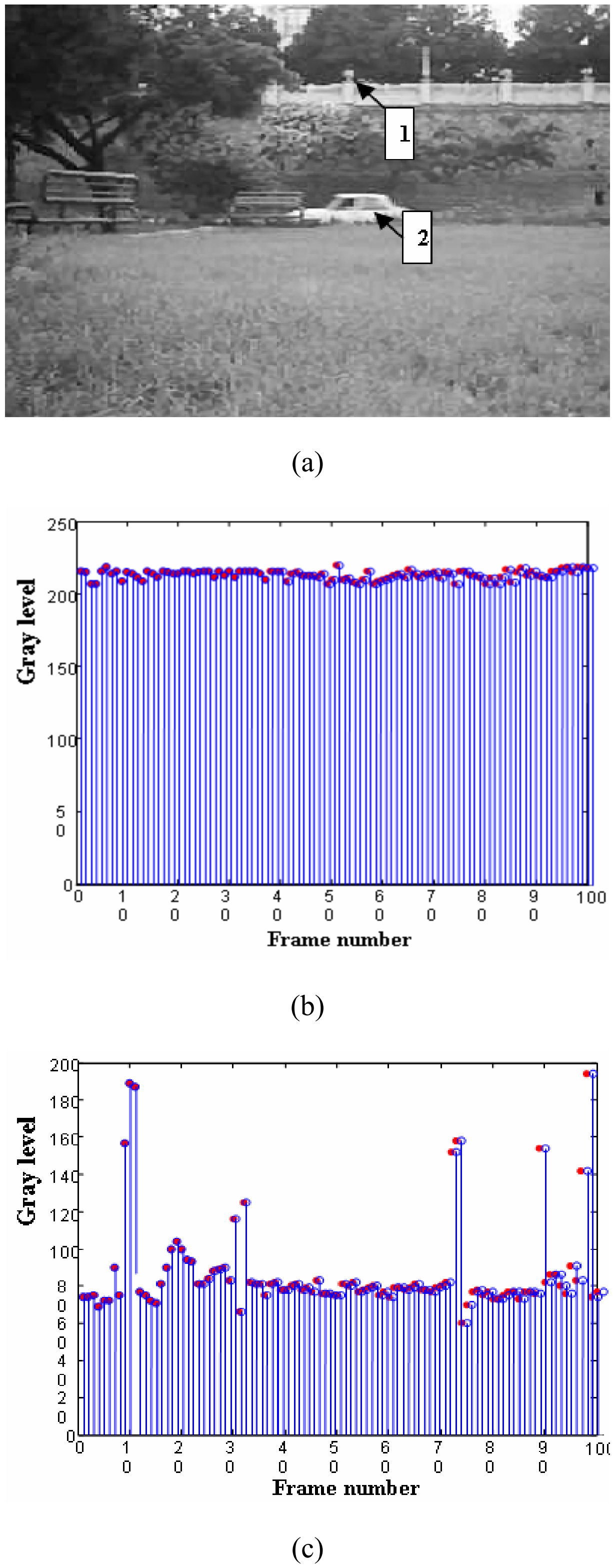


Figure 2. (a) 1 and 2 are pixels where the rate of gray level changes differs due to their location in the image (b) smoothly changes of the first pixel (c) sharp changes of the second pixel.

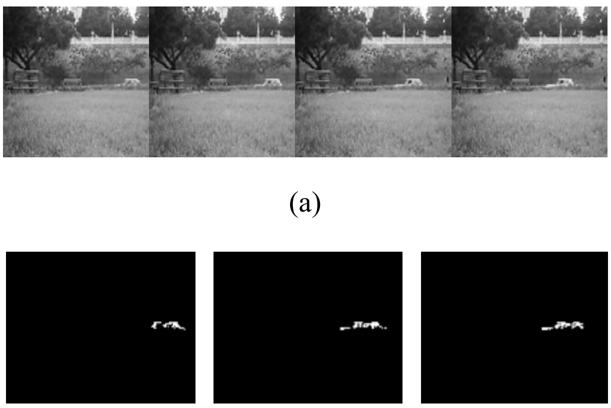
matrix. Horizontal and vertical movement changes the column and row numbers respectively. To detect these changes through a sequence of N frames, we have used threshold difference of consecutive frames to detect active regions, as well as removing scattered unwanted points. Figure 3 shows consecutive differences of four frames. Therefore, a set of differentiated frames is obtained:

$$F = \{f_1, f_2, \ldots f_N\} \quad (1)$$

$$I = \{I_{d1}, I_{d2}, \ldots, I_{d(N-1)}\} \quad (2)$$

Where F is the sequence of N frames and I is the set of their consecutive differences. At this stage connected component labeling is applied to these frames in order to find the most prominent changing region in each one, which deals with the most effective moving object of the scene. The mean values of rows and columns of these regions are calculated in the following sets. Gradient function have been handled in this step, noticing that ascending set of numbers have positive gradient values and descending ones have negative gradient values which can help us discover whether the number of rows and columns are increasing or decreasing.

$$Row = \{r_{m1}, r_{m2}, \ldots r_{m(N-1)}\} \quad (3)$$



(a)

(b)

Figure 3. (a) Consecutive frames of the scene (b) the differences of consecutive frames.

$$Col = \{c_{m1}, c_{m2}, \ldots c_{m(N-1)}\} \quad (4)$$

$$\nabla Row = \frac{\partial Row}{\partial f} = \{gr_{m1}, gr_{m2}, \ldots gr_{m(N-1)}\} \quad (5)$$

$$\nabla Col = \frac{\partial Col}{\partial f} = \{gc_{m1}, gc_{m2}, \ldots gc_{m(N-1)}\} \quad (6)$$

Where f is frame variable which is used to extract how many rows or columns are passed per frame. By calculating the number of positive and negative values in the above sets we have:

$$\nabla Row^{+} = \{gr_{mi} \mid gr_{mi} > 0\} \quad (7)$$

$$\nabla Row^{-} = \{gr_{mi} \mid gr_{mi} < 0\} \quad (8)$$

$$\nabla Col^{+} = \{gc_{mi} \mid gc_{mi} > 0\} \quad (9)$$

$$\nabla Col^{-} = \{gc_{mi} \mid gc_{mi} < 0\} \quad (10)$$

Combining these results with corresponding direction of ascending or descending sets of rows or columns, the orientation of major moving object in the scene is found. For example if the maximum number belongs to ∇Col^{+}, it means that the moving object moves horizontally from left to right side as the column numbers are increasing. Diagonal movements are also classified in horizontal or vertical classes based on angle value of the path.

In a sequence of frames, several moving objects may move in different directions and even change their directions. Initially, there isn't any information about moving objects, such as their size, direction and velocity. In order to minimize the probability of presence of several moving components in frames, it is necessary to limit the number of frames. On the other hand, it is important that the object moves along its covered area among these frames. A special analysis on effective parameters (velocity and size) is performed.

According to the experiments, a normal walking person's speed is about 3 Km/H (83 Cm/S) and for a slow walking person, it's about

0.7 Km/H (20 Cm/S). The mean width of a normal size person is about 50 Cm. Therefore, it takes about 2.5 Seconds for such person to cover the width slowly. Similar calculations have been done for a slow car with the normal width of 2 meters at the speed of 10 Km/H (2.7 m/S). It takes 0.72 Seconds to cover its width. To compromise our experiments and extreme circumstances, the required frames for a slow object at the speed of 3 Km/H and about 4 meters width is considered. According to the mentioned considerations, it takes 5 seconds for such a slow object to move along its covered area. Assuming to have 5 frames per Second, every 25 frames is orientated separately. Surely, these frames are also sufficient for objects at higher speed or smaller size, to move along their covered area. To increase the probability to have more background gray values than foreground, one or more groups of 25 frames are selected according to direction, velocity and frame dimension for background modeling. This will improve the statistical estimation of pixels gray values. More detailed explanations will be given in Section 2.1.2 and 2.1.3.

2.1.2. Velocity of moving objects We have explained that in our algorithm, velocity of moving objects plays a significant role. Many numbers of frames are required so that a slow moving object leaves its initial position while fewer numbers are needed for faster ones.

To distinguish the velocity of an object moving in a particular direction, we have used absolute values of corresponding gradients of rows and columns. This is based on this fact that there are bigger changes in gradient values of faster objects than slower ones.

In previous section we have used a number of positive and negative values, for gradient sets of rows and columns, to find the direction of moving object. The densest sign refers to the direction of the major moving object in the sequence. Whereas the absolute values corresponding to this sign offers the instantaneous velocity associated with moving object per frame. The average values of these velocities are the mean velocity of moving object in its directional path.

For example if the densest sign among the sets in (7-10), belongs to ∇Col^{+}, their velocity set and mean value are as follows:

$$V = \{\left|gc_{p1}\right|, \left|gc_{p2}\right|, \ldots \left|gc_{pk}\right|\} \tag{11}$$

$$V_m = \frac{1}{k}\sum_{i=1}^{k} gc_{pi} \tag{12}$$

Where k is the number of positive signs in ∇Col^{+}.

As mentioned before, we have divided the initial stream into groups of 25 frames, to distinguish the direction of moving objects. To find the mean velocity of moving objects in total stream (v_t), we have calculated the average mean velocities of all groups. This will help us determine the proper number of frames for estimating the intensity of each pixel. A special criterion is used for number of these groups. Logically, we should consider the maximum required length for initial stream. Since slower objects leave the area they have covered in more frames, the limiting factor of maximum length (L_{max}), is defined based upon movement characteristics of normally the slowest objects and also the dimension of our frames. Therefore, number of initial groups with the length of 25 is allocated according to the following formula:

$$N_g = \frac{L_{max}}{25} \tag{13}$$

$$v_t = \frac{1}{N_g}\sum_{i=1}^{N_g} v_{mi} \tag{14}$$

Where N_g is the number of these groups oriented in initial stage separately and v_{mi} is the mean velocity of i^{th} group.

2.1.3. Forming appropriate sequence of frames Considering the average velocity and directions, obtained for N_g number of initial groups, we are able to estimate an appropriate length for the desired set of frames to perform successful background modeling.

In this work, the width (column numbers) and length (row numbers) of each frame are 240 and 320 respectively. Consider an object, moving horizontally at an average velocity of V_t columns per frame. It will pass all the columns in approximately [320/V_t] unidirectional consecutive frames as well as [240/V_t] frames in vertical movements. Since it is essential to have enough number of frames, and considering various size and distance of different moving objects and also the fact that they may not move uniformly, this number have been doubled to increase the accuracy.

$$L_S = 2 \times \text{round}\left(\frac{230}{V_t}\right) \rightarrow \text{horizental}$$
$$L_S = 2 \times \text{round}\left(\frac{240}{V_t}\right) \rightarrow \text{vertical} \qquad (15)$$

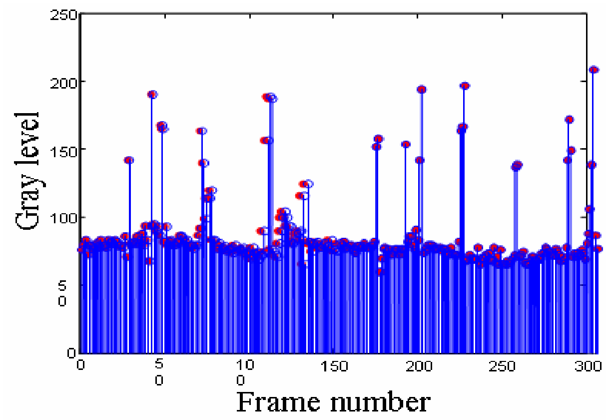
Where L_S is the length of the modified set of frames. At this stage, this value is used in preparing sufficient subsets of 25 frames with a particular direction. As N_g groups of initial stream have been oriented in primary stage. Maximum number of consecutive unidirectional groups is found and according to the estimated length for the modified set (L_S) a sequence of N_S subsets will be selected among them:

$$N_S = \text{round}\left(\frac{L_S}{25}\right) \qquad (16)$$

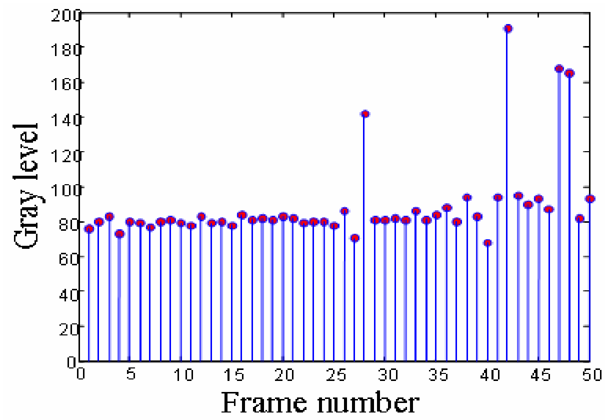
These unidirectional Consecutive subsets are merged together as a set which can be used to estimate pixel intensities statistically and its length is the multiple of 25. This sequence of frames includes objects moving in a certain direction and its length according to velocity observations, assures sufficient background gray values as compared to foreground for each pixel. In Figure 4 the intensity variations of a pixel in an initial stream and modified set have been shown. It points out better homogeneity of intensity values in the second sequence, which efficiently offers successful estimation of background gray level. Figure 5 shows flowchart of the whole algorithm including orientation and velocity extraction to estimate the length of modified set.

2.2. The Proposed Approach As discussed in previous section, by using movement information, modified sequence of frames are provided, in order to increase the probability of viewing background intensities in most frames. However, smooth changes of illumination in a scene affects background gray values and therefore, background may change continuously and no certain gray value may have the highest density to be assigned to a pixel which could be clearly understood in Figure 4b. Under this circumstance, we should make a compromise between these changing gray values of background to approximate pixel intensity.

According to the type of sequence we use, there is higher probability to have limited number of moving object gray values in this set with respect to the ones in the background. Consider such a modified observation set for one pixel. The distance of background members to each other is lower than foreground values because of their



(a)



(b)

Figure 4. (a) Variations of pixel intensities in initial stream (b) variations of intensities in modified set.

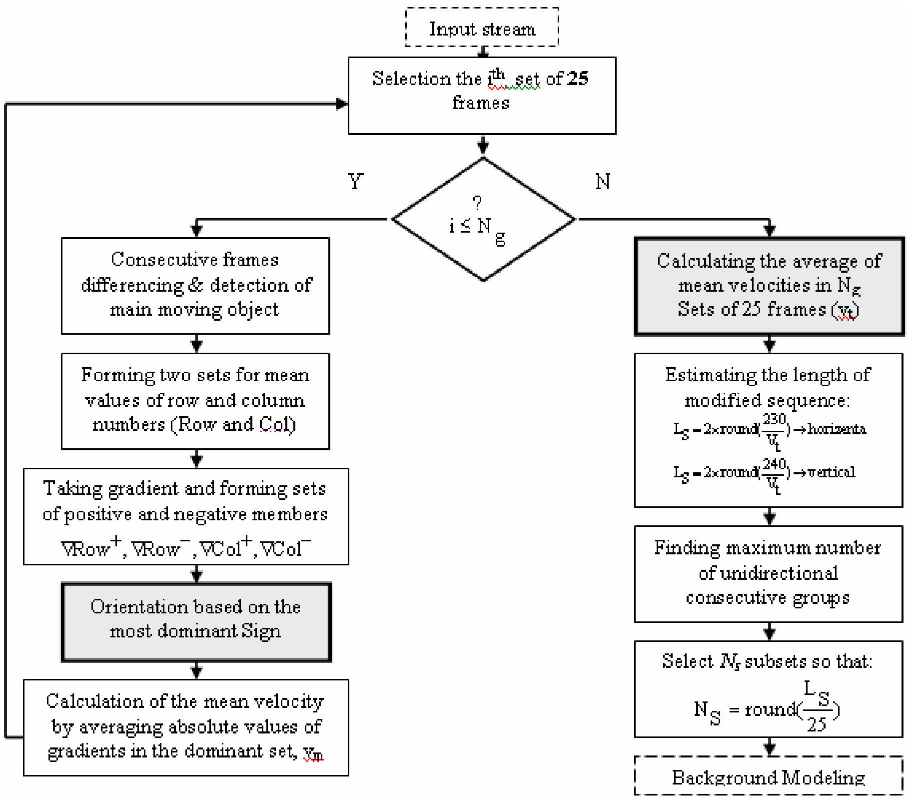


Figure 5. Flowchart representing extraction the direction and velocity of moving objects to estimate the length of modified sequence of frames.

smooth changes. However they have higher density. Therefore, if we calculate the distance between members and then, the sum of these distances, higher values will be referred to foreground members in comparison with background. We have used this feature to model background. In Figure 6, we have shown the modified set of the pixel displayed in Figure 4b again. It is clear that the sum of minimum values in 6c refers to majority of gray levels in 6a which have higher densities in 6b. At this stage, different procedures of background modeling algorithm will be described. The flow chart of this algorithm is shown in Figure 7.

At first, the past history of each pixel is considered through modified frames:

$$P = \{g_1, g_2, \dots g_{L_S}\} \tag{17}$$

Where L_S is the length of modified set (calculated in Section 2.1.3) and g_i is the gray level of the pixel in i^{th} frame and may be equal to or different from other samples. These samples are rearranged as values of X_i in an ascending order and not repeating order in the set P' so that the smallest one is X_1 and the largest one is X_m. The corresponding density of each value is also presented in the set D:

$$P' = \{X_1, X_2, \dots X_m\} \tag{18}$$

$$D = \{d_1, d_2, \dots d_m\} \tag{19}$$

At this stage, the distance of each gray level from others in the set P' is computed and sum of these distances is arranged in the set S:

$$S = \{S_1, S_2, \dots S_m\} \tag{20}$$

$$S_i = \sum_{j=1, j \neq i}^{m} (X_i - X_j) \tag{21}$$

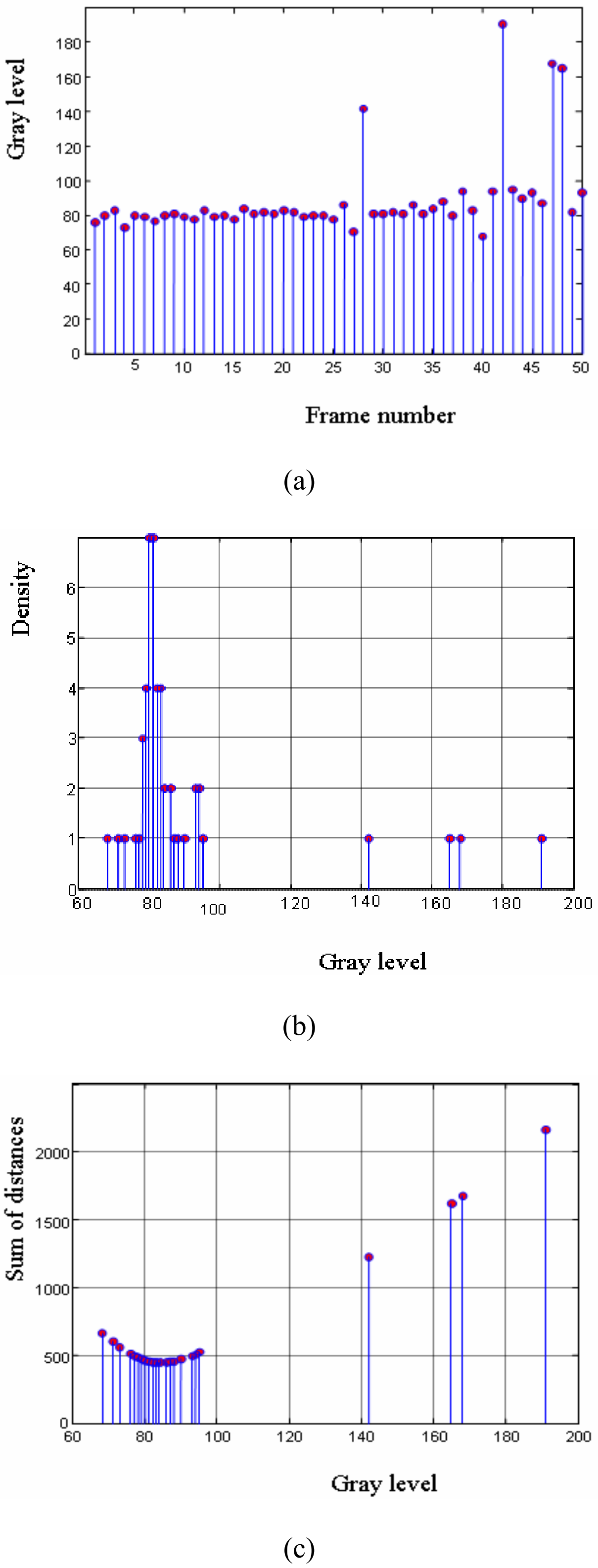


Figure 6. (a) Variation of pixel intensity in modified set (b) variation of density with respect to gray levels (c) variation of sum of distances with respect to gray levels.

The maximum and minimum values of S are calculated which refer to foreground and background respectively as demonstrated in Figure 6c. Then those members of S, nearer to its minimum member S_{mi} than maximum Member S_{mx} are collected in Se, also their corresponding gray levels and densities in the sets X and D are shown in Xe and De as follows. These are referred to background gray values with smooth changes to each other.

$$Se = \{Se_1, Se_2, \ldots Se_t\} \tag{22}$$

$$Se_i = \{s_i \| s_i - s_{mx}| > |s_i - s_{mi}|\} \tag{23}$$

$$Xe = \{Xe_1, Xe_2, \ldots Xe_t\} \tag{24}$$

$$De = \{de_1, de_2, \ldots de_t\} \tag{25}$$

Where t is the number of members of S satisfy (23). We will proceed with these two recent sets to estimate the gray value of the pixel. Those values of Xe associated with higher densities are surely the best choices in this homogeneous set. Though among members of Xe, those corresponding to the maximum value in De are found and their mean value is assigned to the pixel in background image.

$$U_M = \{u_{m1}, u_{m2}, \ldots u_{mx}\} \tag{26}$$

$$d_{m1} = d_{m2} = \ldots = d_{mx} = Max(D_e) \tag{27}$$

$$gr_f = \frac{1}{x} \times \sum_{i=1}^{x} u_{mi} \tag{28}$$

Where the set of members of Xe with maximum density is shown by U_M, the density of u_{mi} in De is shown by d_{mi} and the final intensity estimated for the pixel is shown by gr_f.

Although the whole background can be modeled with the described algorithm, performing these calculations for all pixels, especially in (21) and (23) is time consuming. As a solution, the primitive density of gray values in the set D can be used to estimate the intensity of some pixels directly. If it is possible to find dominant densities among members of D, the mean value of their

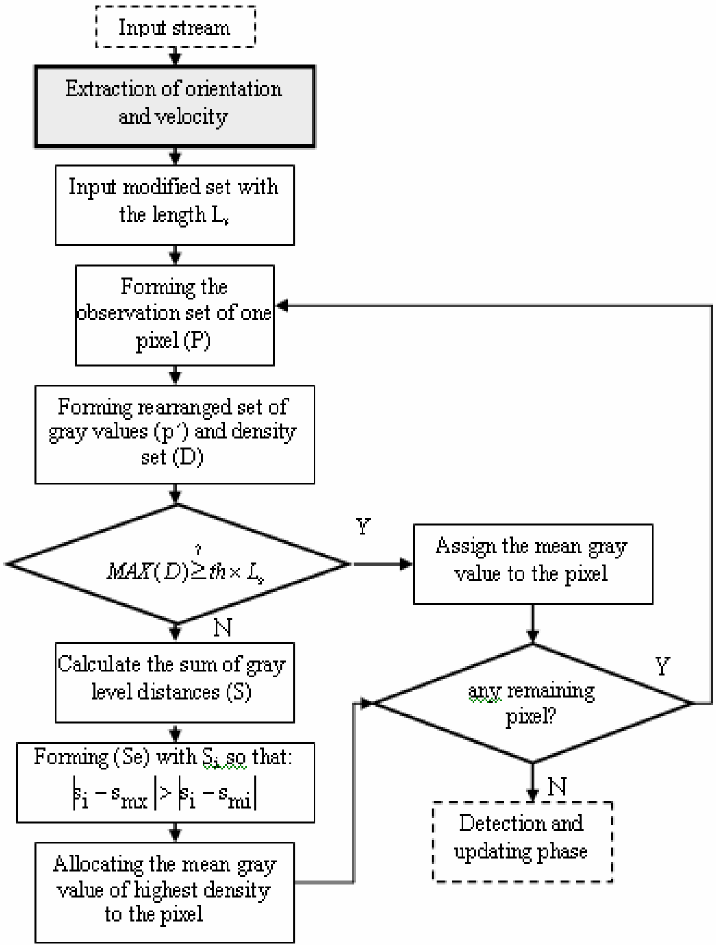


Figure 7. Flowchart representing background modeling algorithm.

corresponding gray levels can be assigned to the pixel. Because, according to the characteristic of our modified set of frames, repetition of one or more gray level(s) of a pixel dominantly, represents smooth environmental changes in the view of that pixel which refers to background than foreground like the one shown in Figure 2b. We have experimentally defined a threshold value based on the maximum time that a change takes place which is a coefficient of the length of modified set (L_S). Therefore estimating the intensity of the pixel is simplified as follows:

$$Y_M = \{y_{m1}, y_{m2}, \ldots y_{mn}\} \tag{29}$$

$$d_{m1} = d_{m2} = \ldots = d_{mn} = MAX(D) \tag{30}$$

$$MAX(D) \geq th \times L_S \tag{31}$$

$$gr_f = \frac{1}{n} \times \sum_{i=1}^{n} y_{mi} \tag{32}$$

Where Y_M is the set of gray values corresponding to dominant densities with the length n, d_{mi} is the density of y_{mi} and th is the defined threshold. The algorithm will be done faster, as the rate of calculations is reduced for those pixels that satisfy (31).

Initial background is updated after a detection period for changed regions of the last frame in similar way.

The whole background is also updated during distinct intervals and through the most recent sequence of frames with the length defined before.

3. EXPERIMENTAL RESULTS

In this section we present experiments in which the proposed algorithm is used to model background and detect moving objects. This work has been tested for moving objects in different sizes and speeds. In order to demonstrate the potential of our algorithm to handle with some challenging problems, we will show the results of our experiments in different highly textured backgrounds, involving environmental changes, or high foreground traffic and also different velocities. To evaluate the efficiency of foreground detection an error rate is defined as false detection rate (FD):

$$FD = \frac{1}{n}\sum_{i=1}^{n} \frac{\text{number of falsly detected pixels}}{\text{number of the whole pixels}} \qquad (33)$$

Where n is the number of frames. We have achieved a rate of 5 frames per second with the size of 240 × 320 pixel for our experiments.

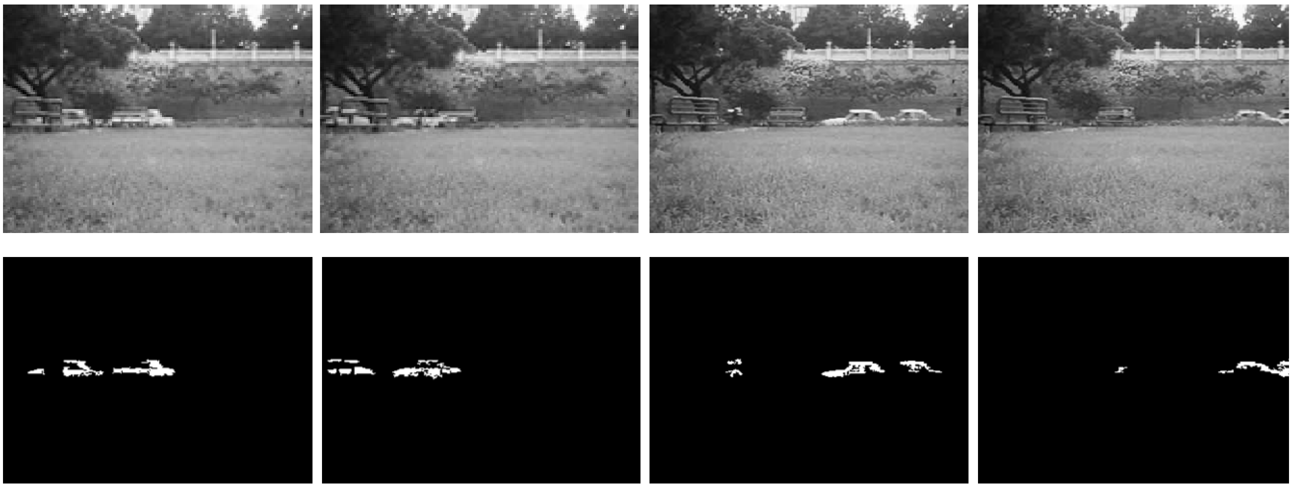
At first we have chosen a dynamic scene with moving objects at high speed. It is performed inside a park toward a street where cars are moving fast in two different directions. Movement of branches and bushes changes the background view continuously. Figure 8b and 8a illustrate the results of background modeling and moving object detection respectively. Figure 8c is the updated background and extracted information is listed in Table 1. According to the listed information, the algorithm has oriented 12 subsets of 25 frames. In subsets number (1-6), 9 and 12, cars were moving horizontally from right to left side as well as moving in the opposite direction in the remaining subsets from left to right. No vertical movement was reported as actually there was no object in the scene moving vertically. The mean velocity is estimated as 18.32 columns per frame and according to this velocity only on subset with 25 frames has been used in background modeling. The FD factor is reported as 5.21 percent.

The second experiment is done in a far away scene of a road in windy condition. The cars seemed smaller than the first experiment. In Figure 9b, we have shown the modeled background and in 9a the results is presented. Figure 9c is the updated background. The experimental results are listed in Table 2. The subset numbers, in which cars were moving horizontally from right to left or left to right, are shown in Table 2. According to the mean velocity of 16.43 column per frame, totally 2 unidirectional subsets with 50 frames were used to model background. The percentage of error is 6.38. It is more than the first experiment as the wind causes more dynamic changes in the scene.

The third experiment is done in the view of a building where a person was walking around. Figure 10b, shows the modeled background and in 10a the detection results is displayed. Figure 10c is the updated model of the background. The results of orientation, velocity and other information are listed in Table 3. We expect that the algorithm distinguishes lower velocity than previous experiments. It is truly estimated as 12.17 columns per frame. The error rate in this experiment has decreased because there is lower environmental change in the scene. Our final experiment presents the view of a window. In a narrow part of the scene, walking people in a crowded street could be seen. The most significant change is the illumination and it happens when the stores turn on their lights. Figure 11b and 11a illustrate the results of background modeling and moving object detection respectively. Figure 11c is the updated background. The results of experiment are listed in Table 4. The mean velocity of people in crowded street has been estimated lower than other experiments and 3 unidirectional subsets with 75 frames are used in background modeling. The error rate has increased due to artificial light changes.

The diagram of the length of modified sets with respect to extracted velocities has been displayed in Figure 12. This diagram shows that in lower velocities, more frames are used for background modeling. This demonstrates the fact that slower objects move along their covered area in more frames.

In order to compare the effect of different



(a)

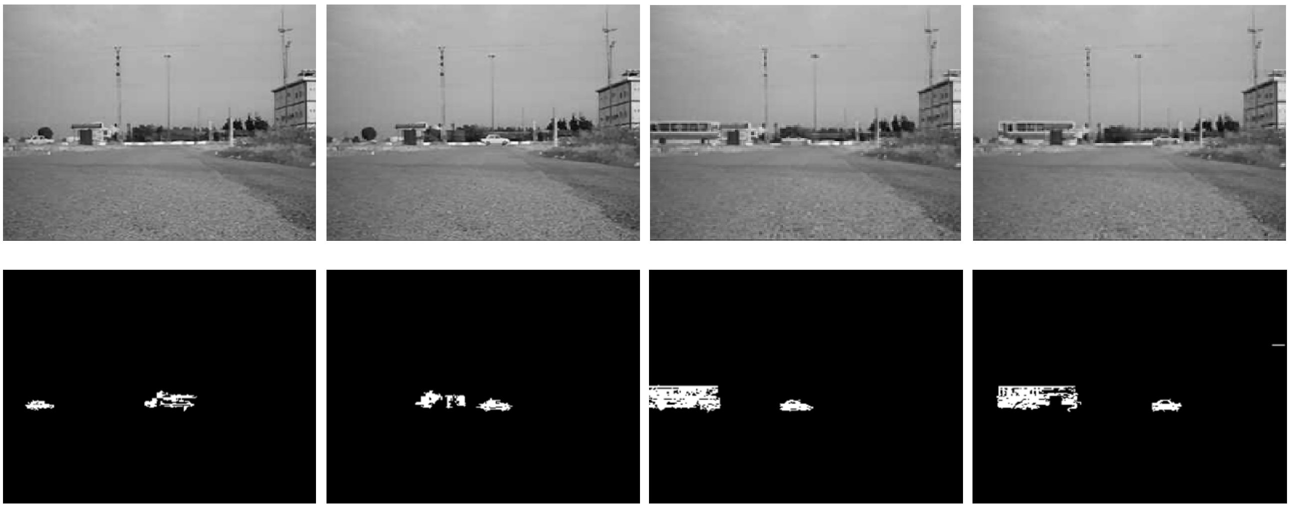
(b)

(c)

Figure 8. (a) The top row displays frames in which cars are moving in the street and the second row is the detection results using background model (b) the modeled background (c) the updated background.

TABLE 1. Extracted Movement Information in First Experiment.

Orientation:	Subset Number	Mean Velocity
Horizontal from right to left	(1-6), 9, 12	18.3226 (Column/f)
Horizontal from left to right	(7,8), (10,11)	Length of modified set (L_S)
		25 (frames)
Vertical from up to down	-	% FD
Vertical from down to up	-	5.21
Orientation:	Subset number	Mean velocity
Horizontal from right to left	(1-6),9,12	18.3226 (Column/f)
Horizontal from left to right	(7,8), (10,11)	Length of modified set (L_S)
		25 (frames)
Vertical from up to down	-	% FD
Vertical from down to up	-	5.21



(a)

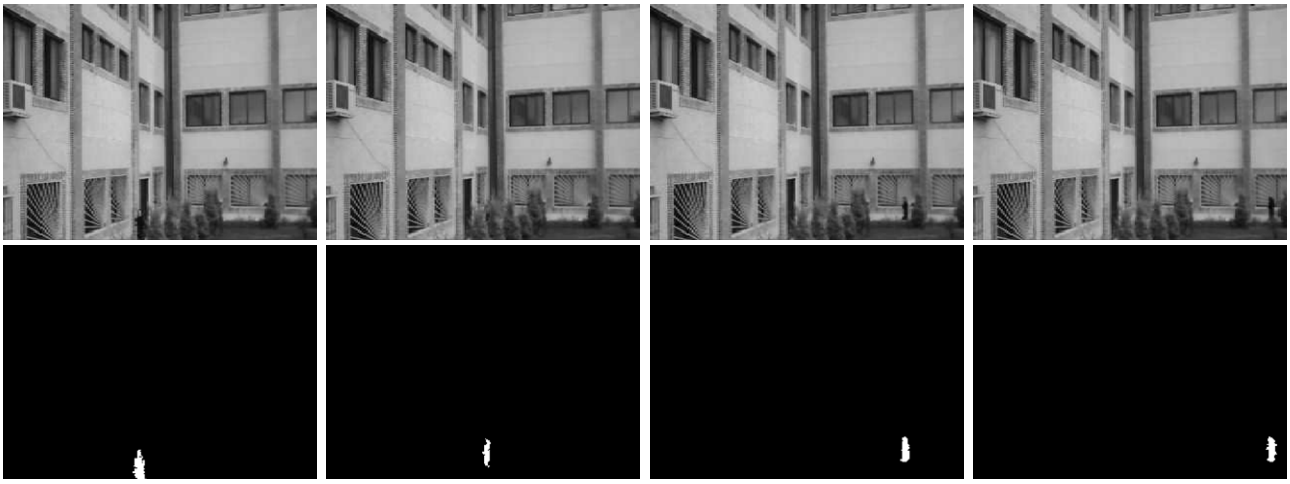
(b) (c)

Figure 9. (a) The top row displays frames in which cars are moving in a far away road and the second row is the detection results using background model (b) the modeled background (c) the updated background.

TABLE 2. Extracted Movement Information in Second Experiment.

Orientation:	Subset Number	Mean Velocity
Horizontal from right to left	(3,4), (7,8), 12	16.4342 (Column/f)
Horizontal from left to right	(1,2), (5,6), (9,10,11)	Length of modified set (L_S)
		50 (frames)
Vertical from up to down	-	% FD
Vertical from down to up	-	6.38

environmental factors that increase the error rate in our algorithm, it is necessary to have similar conditions of velocity and size for Moving objects in the scene. Typically, three experiments were done, in which cars were moving close to each other. These experiments were repeated three times



(a)

(b)

(c)

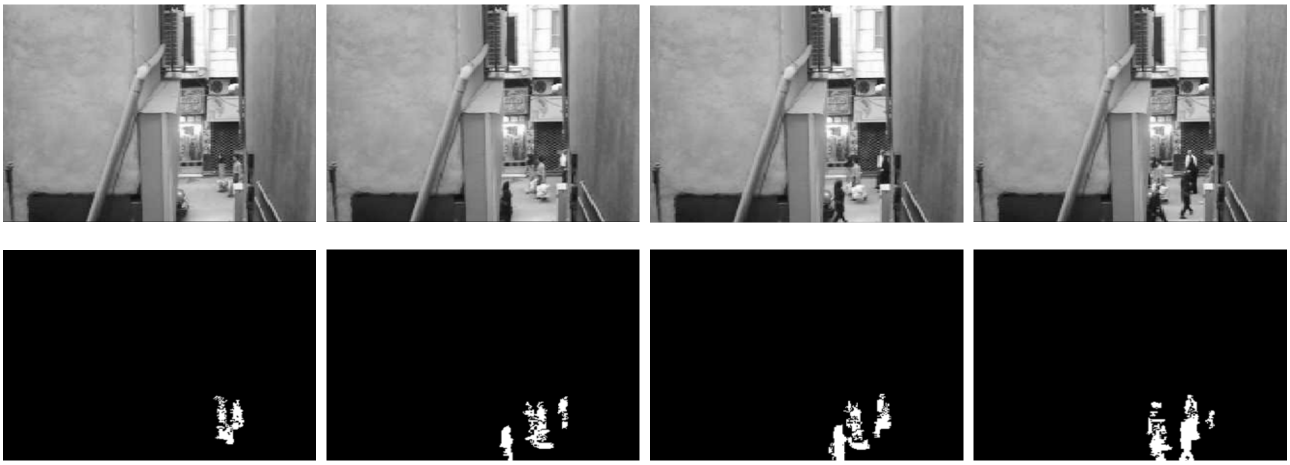
Figure 10. (a) The top row displays frames in which a person is walking and the second row is the detection results using background model (b) the modeled background (c) the updated background.

TABLE 3. Extracted Movement Information in the Third Experiment.

Orientation:	Subset Number	Mean Velocity
Horizontal from right to left	(7,8,9,10)	12.1784 (Column/f)
Horizontal from left to right	(1,2), (5,6), (11,12)	Length of modified set (L_S)
		50 (frames)
Vertical from up to down	-	% FD
Vertical from down to up	(3,4)	3.78

and in three different conditions; in Normal smooth changes of day light, artificial lights and in dark and windy condition.

Figure 13 displays the views of the scenes in three experiments and Figure 14 shows the various percentage of FD in a descending order with



(a)

(b) (c)

Figure 11. (a) The top row displays frames in which people are walking in a street and the second row is the detection results using background model (b) the modeled background (c) the updated background.

TABLE 4. Extracted Movement Information in the Fourth Experiment.

Orientation:	Subset Number	Mean Velocity
Horizontal from right to left	2, (7,8)	8.5234 (Column/f)
Horizontal from left to right	1, (3,4,5,6), (9,10,11,12)	
Vertical from up to down	-	Length of modified set (L_S)
		75 (frames)
Vertical from down to up	-	% FD
		10.54

respect to the most effective agents of error in each experiment. It can be understood from this diagram that windy condition and artificial light can cause more error than the normal illumination changes.

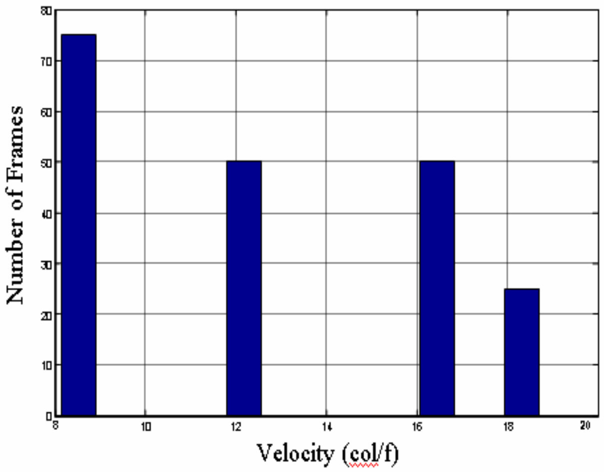


Figure 12. Length of modified sets in comparison with velocities in 4 experiments.

(a) (b) (c)

Figure 13. (a) Windy situation (b) artificial light in dark (c) normal smooth changes in illumination.

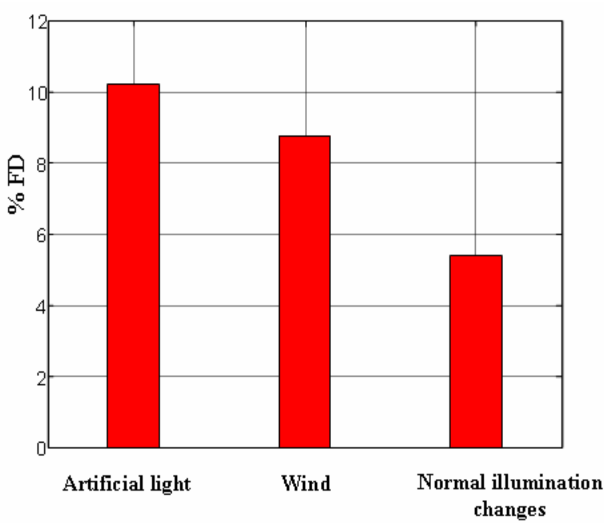


Figure 14. The percentage of error rate in 3 different conditions.

In comparison with other background modeling methods, this work is independent of any statistical distribution. Therefore there is no need to estimate and update any parameter of distribution that reduces their estimation error in different stages of algorithm. Also, unlike some nonparametric methods, there isn't any need to observe the initial background in a learning stage.

4. CONCLUSIONS

Anew statistical method is presented for background modeling while objects are frequently moving in the scene. This method is based on detection of changing pixel's gray level values in a modified sequence of frames. Modification of initial Sequence is performed by extracting the direction and velocity of moving objects. There is also an updating stage to handle changes in the scene. In comparison with other statistical methods, our algorithm is less complicated and independent of statistical distributions and their parameters. The results show successful detection of moving objects.

5. REFERENCES

1. Karmann, K. P., Brandt, A. V. and Gerl, R., "Moving Object Segmentation Based on Adaptive Reference Images", ***Signal Processing V: Theories and Application***, Elsevier Science Publishers B.V., (1990), 951-954.
2. Koller, D., Weber, J., Huang, T., Malik, J., Ogasawara, G., Rao, B. and Russell, S., "Towards Robust Automatic Traffic Scene Analysis in Real-Time", ***Proceedings of the International Conference on Pattern Recognition***, Vol. 4, (1994), 3776-3781.
3. Karmann, P. and Vonbrandt, A., "Moving Object Recognition using an Adaptive Background Memory", ***Time-Varying Image Processing and Moving Object Recognition***, Elsevier Science Publishers B.V., (1990), 289-296.
4. Gonzalez, R. C. and Woods, R. E., "Digital Image Processing", Prentice Hall, Upper Saddle River, New Jersy, U.S.A., (2004).
5. Wren, C., Azarbayejani, A., Darrell, T. and Pfinder, P., "Real-time Tracking of the Human Body", ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, (July 1997), 780-785.
6. Stauffer, C. and Grimson, W. E. L. "Adaptive

Background Mixture Models for Real-Time Tracking", ***Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition***, Vol. 2, (1999), 252-256.

7. Stauffer, C. and Grimson, W. E. L., "Learning Patterns of Activity using Real-Time Tracking", ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, (August 2000), 747-757.
8. KaewTraKulPong, P. and Bowden, R., "An Improved Adaptive Background Mixture Model for Real Time Tracking with Shadow Detection", ***Proceeding 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01***, Kingston, UK., (September 2001), 149-158.
9. Tuzel, O. Porikli, F. and Meer, P., "A Bayesian Approach to Background Modeling", ***IEEE Workshop on Machine Vision for Intelligent Vehicles (MVIV)***, Vol. 3, (June 2005), 58-63,
10. Elgammal, A., Harwood, D. and Davis, L., "Non-Parametric Model for Background Subtraction", ***Proceedings of the 6th European Conference on Computer Vision-Part II***, (May 2000), 751-767.
11. Elgammal, A., Duraiswami, R., Harwood, D. and Davis, L. S., "Background and Foreground Modeling using Nonparametric Kernel Density Estimation for Visual Surveillance", ***Proceeding IEEE, Comput. Vision Lab***, Maryland Univ, College Park, MD, Vol. 90, Issue, 7, (2002), 1151-1163.
12. Adam, A., Shimshoni, I. and Rivlin, E., "Aggregated Dynamic Background Modeling", ***IEEE International Conference on Image Processing (ICIP)***, (October 2006), 3313-3316.
13. Kim, K., Chalidabhongse, T. H., Harwood, D. and Davis, L., "Real-Time Foreground-Background Segmentation using Codebook Model", ***Real-Time Imaging***, Vol. 11, (June 2005), 172-185.
14. Rittscher, J., Jakarta, S. J. and Black, A., "A Probabilistic background Model for Tracking", ***Proceeding Europe on Conf. on Computer Vision and Pattern Recognition***, Vol. 2, (2000), 336-350.
15. Stenger, B., Ramesh, V., Paragios, N., Coetzee, F. and Buhmann, J. M., "Topology Free Hidden Markov Models: Application to Background Modeling", ***Proceeding 8th Intl. Conf on Computer Vision***, Vancouver, Canada, (2001), 294-301.
16. Grabner, H., Roth, P. M., Grabner, M. and Bischof, H., "Autonomous Learning of a Robust Background Model for Change Detection", ***Proceeding, IEEE International Workshop on Performance Evaluation of Tracking and Surveillance***, (2006), 39-46.