Smart Video Surveillance System using Android Smart Phones with GCM Alert

Rupa Kesavan[1] R. Anbarasu[2]
[1]Final Year (ME –CSE) [2] Assistant Professor
[1, 2]Computer Science & Engineering Dept.
[1, 2]Selvam College of Technology, Namakkal

Abstract—Video surveillance systems are becoming increasingly important for crime investigation and the number of cameras installed in public space is increasing. However, many cameras installed at fixed positions are required to observe a wide and complex area. In order to efficiently observe such a wide area at lower cost, mobile robots are an attractive option. According to the result of moving object detection research on video sequences, the movement of the people is tracked using video surveillance. The moving object is identified using the Cauchy distribution model. The Cauchy distribution model will compare the current frame with the previous frame. The threshold value is calculated to find the moving image. Using threshold value the detected pixel is identified. Hence the movement of the object is identified accurately. After motion detection it will send GCM alert to the android mobile application. Experimental results demonstrate the effectiveness of the proposed method in providing a promising detection outcome and a low computational cost.

Keywords: Background matching framework, background model, video surveillance, motion detection.

I. INTRODUCTION

VIDEO surveillance systems have become widely available to ensure safety and security in both the public and private sectors due to incidents of terrorist activity and other social problems. In general, the video surveillance system is usually developed with the motion detection task applied to determine the motion region. Moreover, motion detection has been used for many computer vision applications, including recognition of traffic situations [1], detection of foreign bodies in food [2], event recognition of human actions [3], visualization of traffic flow [4], detection and classification of highway lanes [5], driver assistance [6], face detection [7], adaptive response to critical motion of nearby vehicles [8], human–machine in-teraction [9], detection of driver intentions [10], remote image processing [11], collision prediction of pedestrians [12], etc. According to a survey of relevant research [13], motion detection methods can be categorized into three major classes, i.e., temporal difference [14], [15], optical flow [16], [17], and background subtraction [18]–[20]. Temporal difference methods readily adapt to sudden changes in the environment, but the resulting shapes of moving objects are often incom-plete [20]. Optical flow methods generally show the projected motion on the image plane with good approximation based on the characteristics of flow vectors. Unfortunately, flow vectors of moving objects only indicate streams of moving objects, thus detecting a sparse form of object regions. Moreover, the computational complexity of optical flow methods is usually too high to easily implement the motion task in the general video surveillance system [20]. Out of these three categories, background subtraction methods received the most attention due to the moderate time complexity and the accurate detection of moving entities [20].

In general, the existing background subtraction methods can detect moving objects by estimating the absolute difference between each incoming video frame and the background model, which is applied to calculate the binary moving objects' de-tection mask with the object threshold function. However, the generated background model may not be applicable in some scenes with some specified issues, including but not limited to six terms.

1) *Adaptability to illumination change:* The background model should adapt to gradual illumination changes.
2) *Dynamic textures adaptation:* The background model should be able to adapt to dynamic background movements, which are not of interest for visual surveillance, such as moving curtains.
3) *Noise tolerance:* The background model should exhibit appropriate noise immunity.
4) *Sensitivity to clutter motion:* The background model should not be sensitive to repetitive clutter motion.
5) *Bootstrapping:* The background model should be properly generated at the beginning of the sequence.
6) *Convenient implementation:* The background model should be able to be set up fast and reliably.

This paper proposes a novel motion detection method to extract moving objects using the Cauchy distribution with the proposed high-quality background model. Our proposed method is organized here.

1) A self-adaptive background matching framework is pro-posed to select suitable background candidates of back-ground model generation.
2) The conditional Cauchy distribution model is applied to extract moving objects of the video sequence.

Based on both qualitative and quantitative evaluations, experiments will verify that the proposed method is more efficient than other state-of-the-art methods in terms of motion detection in a wide range of natural video sequences.

The remainder of this paper is organized as follows: Section II describes some representational motion detection methods that are implemented in our experiments. Section III presents our motion detection method in detail. In Section IV, we present the experimental results and discussions. Finally, our conclusion is provided in Section V.

II. RELATED WORKS

Here, we describe and discuss three representative mo-tion detection methods, including a multitemporal difference (MTD) method [15], a Gaussian mixture model (GMM)

method [18], and a multiple Σ− estimation (MSDE) method [19]. Note that each method is described by assuming that $I_t(x, y)$ refers to an original video frame with size M by N, where t represents the frame index, and (x, y) represents the coordinate of each pixel.

A. MTD method

Compared with the traditional temporal difference method, the MTD method holds several previous reference frames to reduce holes inside moving entities for motion detection [15]. As mentioned in [15], seven previous reference frames are used to calculate the difference image. The previous frames can be held by the following function:

$$Bn = I_t - 1 \text{ if } count = 5^n, \text{ then } n = n + 1 \quad (1)$$

where B_n represents each reference frame, n represents the index of reference frame, and *count* represents the short-term counter, which should be progressively increased *(Count ++)* at each frame. Note that both n and *count* are initialized to zero. Furthermore, n and *count* should be reset to zero when n is greater than n_{max}, which is empirically set to 6 according to [15].

After generating the reference frames, the absolute difference image $_t(x, y)$ can be calculated as follows:

$$t(x, y) = \sum_{n=0}^{n_{max}} |l_t(x, y) - B_n(x, y)| \quad (2)$$

The binary mask of moving objects is calculated as follows:

$$D_t(x, y) = \begin{cases} 1 & if\ t(x, y) > 250 \\ 0 & otherwise \end{cases} \quad (3)$$

where $D_t(x, y)$ is equal to 1 to represent a motion pixel and is equal to 0 to represent a background pixel.

B. GMM Method

Suppose that each pixel can be represented as a mixture with more than one particular distribution, the GMM method maintains the persistence of K Gaussians that correspond to the time series of background pixel values [18]. For each pixel, the probability can be calculated by the following function:

$$p(X_t) = w_{i.t} \cdot \eta(X_t, \mu_{i.t}, \sum_{i=1}^{k} i.t) \quad (4)$$

where X_t is each pixel value of the tth image frame, and $w_{i,t}$ is a weight vector for the corresponding Gaussian distribution that can be expressed as follows:

$$\eta\left(X_t, \mu_{i.t}, \sum_{i=1}^{k} i.t\right) = \frac{\exp\left(\frac{X_t - \mu_t}{2}\right) \cdot \Sigma^{-1}(X_t, \mu_t)}{2\pi - \frac{n}{2} |\Sigma|^{\frac{1}{2}t}} \quad (5)$$

where $\mu_{i,t}$ is the mean value of the ith Gaussian, and $\Sigma_{i,t}$ is the covariance matrix of the ith Gaussian. For computational reasons, covariance matrix $\Sigma_{i,t}$ can be assumed as

$$\Sigma_{k,t} = {\sigma_k}^2 \cdot l \quad (6)$$

Note that the statement of each pixel can be assumed as a match if the pixel value is within 2.5 (τ_1) standard deviations of a distribution. Otherwise, each pixel value can be checked against the existing κ Gaussian distributions.

The adaptive parameters of the mixture model that matches the new observation can be updated as follows:

$$w_{k,t} = (1 - \alpha) \cdot w_{k,t-1} + \alpha \cdot M_{k,t} \quad (7)$$

$$\mu_t = (1 - \rho) \cdot \mu_{t-1} + \rho \cdot X_t \quad (8)$$

$${\delta_t}^2 = (1 - \rho) \cdot {\delta_{t-1}}^2 + \rho \cdot (X_t - \mu_t)^T \cdot (X_t - \mu_t) \quad (9)$$

For the pixel that matches with models, parameter $M_{k,t}$ is 1; otherwise, it is 0. The adaptive background models are then estimated using the value w/σ of each Gaussian, and the initial B distributions can be selected as follows:

$$B = {}_{k=1}^{b} argmin(w_k) > T_2 \quad (10)$$

where T_2 is a measure of the minimum portion of the data that should be categorized as the background.

C. MSDE Method

Based on a constant sign function, the MSDE method generates several reference images to calculate a mixture background model [19]. If the current frame index t is a multiple of α_i, each reference image $b^i_t(x, y)$ can be generated as follows:

$$b_t^i(x, y) = b_{t-1}^i(x, y) + \text{sgn}\ b_t^{i-1}(x, y) - b_{t-1}^i(x, y) \quad (11)$$

where i is the index of each reference image that is a real number from 1 to R. Note that $b^0_t(x, y)$ is regarded as the original frame $I_t(x, y)$. According to [19], R is set to 3, α_1 is set to 1, α_2 is set to 8, and α_3 is set to 16. In addition to reference images, each corresponding weight value $v_t^i(x, y)$ is also calculated as follows:

$$v_t^i(x, y) = v_{t-1}^i(x, y) + \text{sgn}\ N\ {}_t^i(x, y) - v_{t-1}^i(x, y) \quad (12)$$

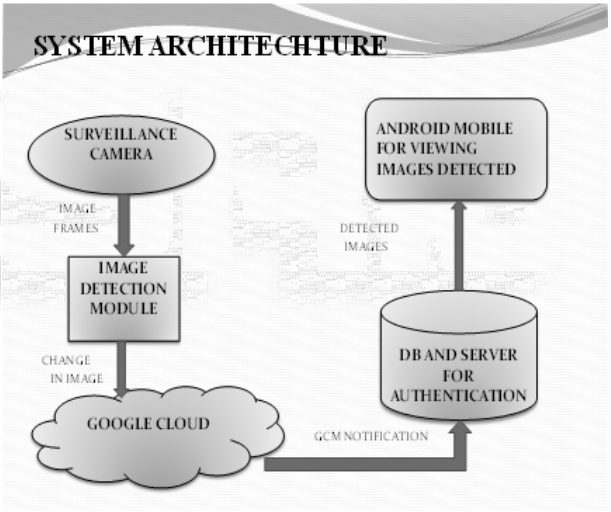

Fig. 1: Flowchart of the proposed method.

where $_i(x, y)$ represents the difference between $b^i_t(x, y)$ and $I_t(x, y)$.

After generating reference images and weight values, a mixture background model $B_t(x, y)$ can be calculated as

$$B_t(x, y) = \frac{\sum_{i=1}^{R} \alpha_i b_t^i(x, y) / v_t^i(x, y)}{\sum_{i=1}^{R} \alpha_i / v_t^i(x, y)}. \quad (13)$$

Finally, the first weight value $_t^1(x, y)$ can be used as the threshold parameter to detect each motion pixel.

D. Discussion

Since the MTD method holds previous reference frames in-stead of background generation, the computational cost is low, and holes are reduced inside moving entities [15].

However, each held frame may include moving objects to drag ghost artifacts in the detection results.

Compared with the single background model, GMM and MSDE methods use several temporal terms to generate multiple reference images, which can be applied to calculate a mixture background model [18], [19]. Unfortunately, these methods cannot detect moving objects well since the multiple and clutter motion may drag ghost artifacts to the background model.

To compensate for the limitations of these methods, a method must be developed to create a background model without ghost artifacts in terms of motion detection. To achieve this objective, a background matching framework should be designed to select the possible background pixel at each frame by using the global trend of the video stream.

III. Proposed Method

Here, we propose a novel motion detection method with a new background model and a Cauchy distribution model. Fig. 1 shows the flowchart of the proposed method. To facilitate the quick determination of the suitable background region, each pixel is checked by the temporal match method in the proposed background model at each frame. Subtracting the generated background model from each input frame and we can obtain the absolute differential values. Finally, we develop a conditional Cauchy distribution model to generate an accurate motion mask for the accurate detection of moving objects at each frame.

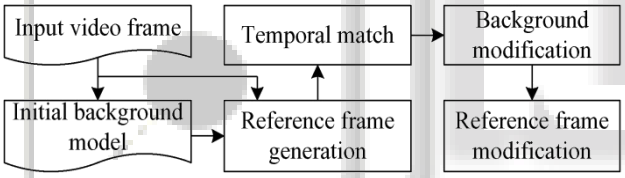

Fig. 2 : Flowchart of the proposed BMM.

A. Background Model

1) Initial Background Model:

For employment of back-ground model generation, adequate initialization is necessary for the proposed background model module; the average ob-tained at the beginning of the input video sequence well suffices as the initial approximation of the background. Therefore, the average of the prior K frames must be first calculated for initial background model generation by

$$B_t(x,y) = \frac{1}{k}\sum_{t=0}^{k-1} I_t(x,y) \quad (14)$$

where $I_t(x, y)$ represents the incoming video frame, $B_t(x, y)$ represents the background model, t represents the frame num-ber, and K is experimentally set at 30 to calculate the initial background model.

2) BMM:

After generating the initial background model, the proposed background matching mechanism (BMM) can be used to modify the background image at each frame. Fig. 2 shows the flowchart of the proposed BMM.

Reference frame generation: To update the background image without the moving object, the reference frame should be calculated at each frame. For each pixel, the reference frame can be calculated as follows:

$$M_t(x, y) = M_{t-1}(x, y) + \text{sgn}\,(I_t(x, y) - M_{t-1}(x, y)) \quad (15)$$

where $M_t(x, y)$ represents each pixel value of the current reference frame, $M_{t-1}(x, y)$ represents each pixel value of the past reference frame, and $I_t(x, y)$ represents each pixel value of the current input frame. Note that M_t is ini-tialized by the initial background model at the frame with index $K - 1$.

Temporal match: Based on reference frame generation, we can easily decide a current set of background candidates to update the background model. For each pixel, $I_t(x, y)$ is regarded as a background pixel if $M_t(x, y)$ is equal to $I_t(x, y)$. Otherwise, the input pixel is eliminated from the background candidates.

Background modification: After determining the background candidates at each frame, the current background model $B_t(x, y)$ can be updated by the following equation:

$$B_t(x, y) = B_{t-1}(x, y) + \text{sgn}\,(M_t(x, y) - B_{t-1}(x, y)) \quad (16)$$

where $B_t(x, y)$ represents the current background model, $B_{t-1}(x, y)$ represents the previous background model, and $M_t(x, y)$ represents a current set of background candi-dates, which is equivalent to $I_t(x, y)$.

Reference frame modification: Based on temporal match and background modification, each modified pixel of the background model can be further used to update the reference frame $M_t(x, y)$. Since each pixel of the input video frame is strictly matched in terms of determining the background candidates, the background model can be updated without the noise pixel of the input image at each frame. Moreover, the reference frame is also updated by the modified background pixel. Hence, the adaptive back-ground model can be generated by the proposed BMM.

B. Motion Mask Generation

1) Absolute Differential Estimation:

Based on the proposed background model module, the absolute differencing image $_t(x, y)$ can be generated by calculating the modulus of the difference between the background model

$B_t(x, y)$ and the incoming video frame $I_t(x, y)$ at each frame, i.e.,

$$t(x, y) = |B_t(x, y) - I_t(x, y)| \quad (17)$$

2) Cauchy Distribution Model:

The accurate detection of the pixels of moving objects in the absolute differencing image $t(x, y)$ at each frame is conductedvia the Cauchy distribution model. It is expressed as follows:

$$f(\Delta t(x, y); a, b) = \frac{1}{\pi} \cdot \frac{b}{(\Delta t(x,y) - a)^2 + b^2} \quad (18)$$

where a represents the location parameter, and b represents the scale parameter, which is experimentally set to 20.

The differentiation between the moving objects and the back-ground region can be achieved using the Cauchy distribution model at each pixel. For the classification of the background region at each pixel, the first type of the developed conditional probability model $f_1(\Delta_t(x, y); a_1, b)$ is utilized and can be expressed as

$$f(\Delta t(x, y); a1, b) = \frac{1}{\pi} \cdot \frac{b}{(\Delta t(x,y) - a1)^2 + b^2} \quad (19)$$

where a_1 represents the location parameter of the first type of the developed conditional probability model and is

calculated as follows:

$$a1 = {}_{l=0}^{b} l \times n_l$$

where b represents the scale parameter,
l represents the arbitrary gray level within the absolute differencing image $t(x, y)$, and
nl represents the pixel number corresponding to the arbitrary gray level l
where a_2 represents the location parameter of the second type of the developed conditional probability model and is calculated as follows:

$$a2 = \frac{{}^{l_{max}}_{l=b+1} l \times n_l}{{}^{l_{max}}_{l=b+1} n_l} \tag{22}$$

where b represents the scale parameter, l represents the arbitrary gray level within the absolute differencing image $t(x, y)$, $l_{\max}$ represents the maximum gray level within the absolute dif-ferencing image $t(x, y)$, and n_l represents the pixel number corresponding to the arbitrary gray level l.

Finally, the binary object binarization mask $D_t(x, y)$ can be formed as follows:

$$D_t(x, y) = \begin{cases} 0 & if\ f_1 > f_2 \\ 1 & otherwise \end{cases} \tag{23}$$

Notice that each pixel of $D_t(x, y)$ is equal to 0 when it belongs to the background region of the incoming video frame $I_t(x, y)$ and that each pixel of $D_t(x, y)$ is equal to 1 when it belongs to moving objects in the incoming video frame $I_t(x, y)$.

IV. Experimental Results

This section presents a comprehensive comparison between our BMM-based Cauchy distribution (BMMC) method and other motion detection methods, including the MTD [15], GMM [18], and MSDE [19] methods. We use six video sequences to test each method by qualitative visual inspection. On the other hand, quantitative evaluation is also applied to measure the detection accuracy by using four metrics, including *Recall*, *Precision*, F_1, and *Similarity* [21]–[25]. Finally, the time consumption of each method is reported in terms of performance evaluation.

A. Qualitative Measurement

Figs. 3–8 show four representative motion masks of each method by using six video sequences, which are "express way (EW)," "hall (HA)," "intelligent room (IR)," "university (MSA)," "road (RD)," and "street (ST)." Note that these sequences feature both indoor and outdoor environments. Slight oscillations of the camera caused by wind, gradual illumination changes because of changes in the cloud cover or the location of the sun, and many natural phenomena of the outdoor environments may act as sources of complexity for processing the obtained video sequences. For the indoor environments, such as office, meeting rooms, and passages, the accuracy of the detection rate may be decreased when the moving object is slowing down or stopping.

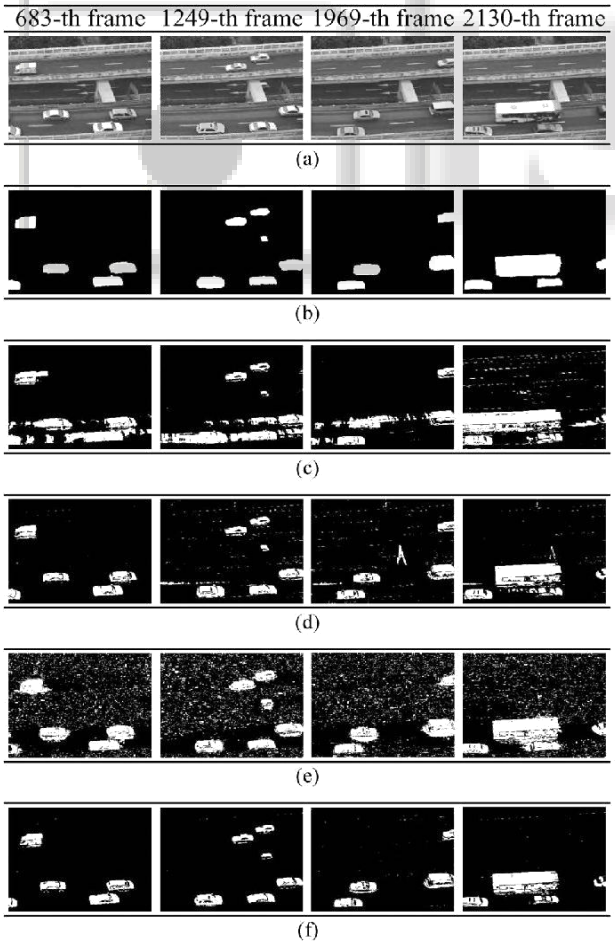

Fig. 3. EW video sequence. (a) Four original EW frames. (b) Ground truths. The remaining four subpictures show the motion masks generated by the (c) MTD [15], (d) GMM [18], (e) MSDE [19], and (f) BMMC methods.

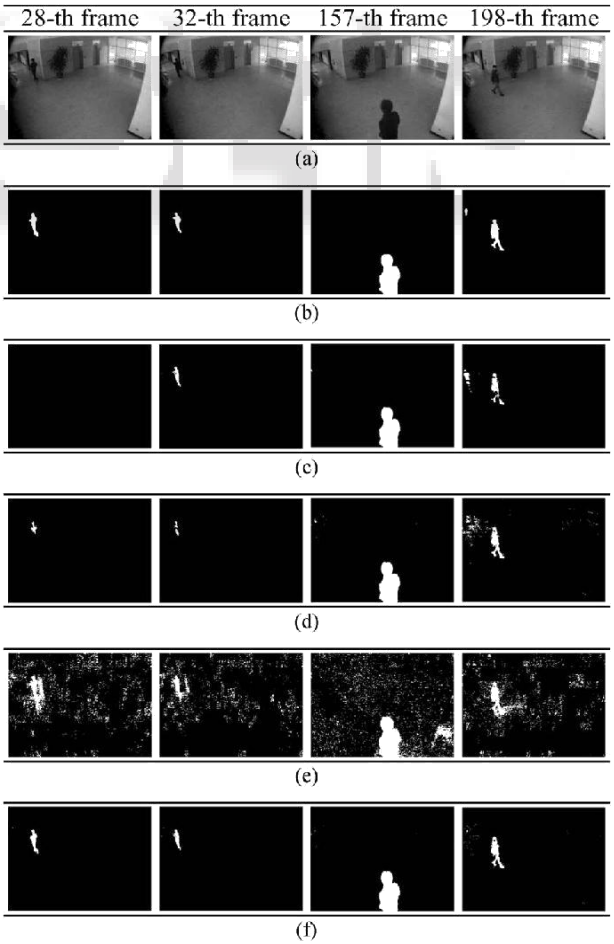

Fig. 4. HA video sequence. (a) Four original EW frames. (b) Ground truths. The remaining four subpictures show the motion masks generated by the (c) MTD [15], (d) GMM [18], (e) MSDE [19], and (f) BMMC methods.

Fig. 3 shows four representative frames with ground truth of the EW sequence and motion masks obtained by the MTD [15], GMM [18], MSDE [19], and BMMC methods.

GMM method uses several Gaussian kernels to generate a mixture background model, thus generating the acceptable motion masks, as shown in Fig. 3(d). Compared with the GMM model, the MSDE method also generates a mixture background model by constant sign calculation. Moreover, constant sign calculation is also applied to calculate the threshold parame-ter in terms of motion detection. Unfortunately, the threshold parameter is still modified when the incoming pixel is the background, thus detecting the wrong object pixels, as shown in Fig. 3(e). The proposed BMMC method detects the moving object by comparing the parameters of the background and the foreground, which are calculated by two conditional Cauchy-based kernels instead of a single threshold function. Hence, noise artifacts are reduced in the detection results shown in Fig. 3(f).

Fig. 4(a) shows four representative frames of the HA video sequence, and Fig. 4(b) shows its ground truths, whereas other subpictures show the binary mask of moving objects obtained by the MTD [15], GMM [18], MSDE [19], and BMMC meth-ods, respectively. This sequence shows a scene of one floor at a hall. As shown in Fig. 4(c)–(f), all methods can be applicable because few people walk in the scene. In other words, the refer-ence frames of the MTD method always contain the background region, and other methods can easily generate the background model. However, the MTD method cannot detect the moving object at the 28th frame since this method should use 30 frames for initialization. On the other hand, the MSDE method still generates motion masks with the noises.

Fig. 5 shows four representative frames of the IR video sequence, ground truths, and binary mask of moving objects ob-tained by the MTD [15], GMM [18], MSDE [19], and BMMC methods, respectively. In Fig. 5(a), one person is working in a room. Due to the low quality of the device, some system noises are generated in this sequence. As shown in Fig. 5(c), the MTD method generates motion masks with artificial ghost trails because the previous reference frames also involve the consecutive motion pixels of a moving object. In Fig. 5(d) and (e), both GMM and MSDE approaches generate serious noises due to the system noises. The proposed BMMC approach applies the Cauchy model twice to calculate the parameter for the probabilities of the background and the foreground. Therefore, the motion masks can be generated without noises, as shown in Fig. 5(f).

Fig. 6 shows four representative frames of the MSA video sequence, ground truths, and binary mask of moving objects obtained by the MTD [15], GMM [18], MSDE [19], and BMMC methods, respectively. Some resemblances exist in the motion masks of each method between IR and MSA sequences since the characteristic of the MSA sequence is very similar to that of the IR sequence. Therefore, we can easily observe that the proposed BMMC method still generates better results than those of other methods.

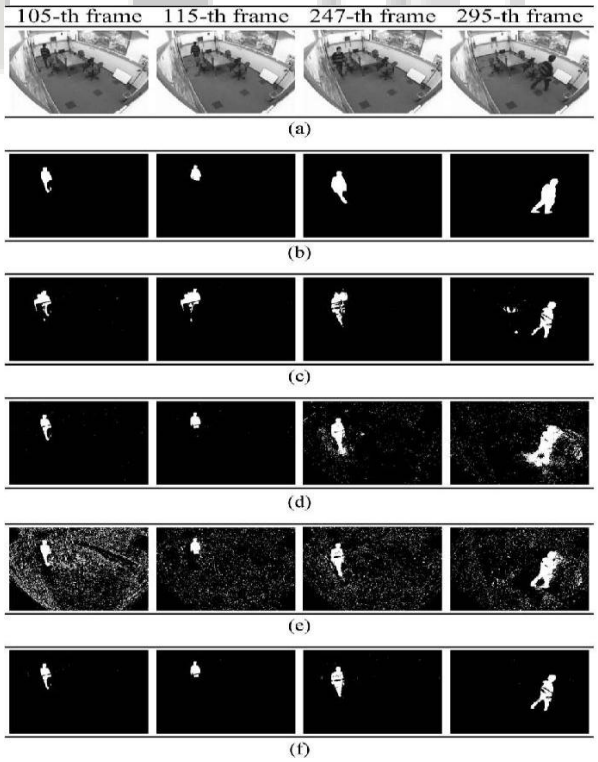

Fig. 5: IR video sequence. (a) Four original IR frames. (b) Ground truths. The remaining four subpictures show the motion masks generated by the (c) MTD [15], (d) GMM [18], (e) MSDE [19], and (f) BMMC methods.

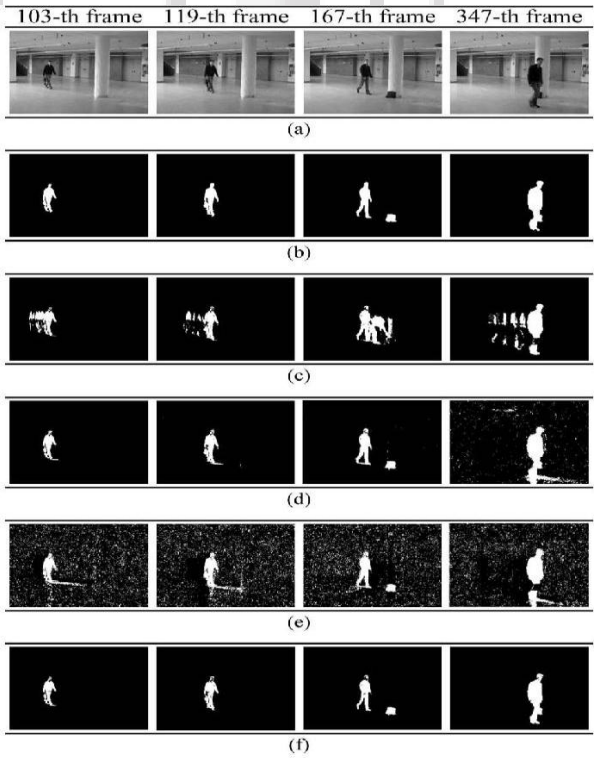

Fig. 6: MSA video sequence. (a) Four original MSA frames. (b) Ground truths. The remaining four subpictures show the motion masks generated by the (c) MTD [15], (d) GMM [18], (e) MSDE [19], and (f) BMMC methods.

Fig. 7(a) shows four representative frames of the RD video sequence, and Fig. 7(b) shows its ground truths, whereas other subpictures show the binary mask of moving objects obtained by the MTD [15], GMM [18], MSDE [19], and BMMC methods, respectively. Compared with the EW sequence, this sequence shows the driven vehicles with large sizes because of a short focal distance. Since many vehicles are very active in the road, the artificial ghost trails and serious noises are generated by each method except the proposed method, as shown in Fig. 7(c)–(f).

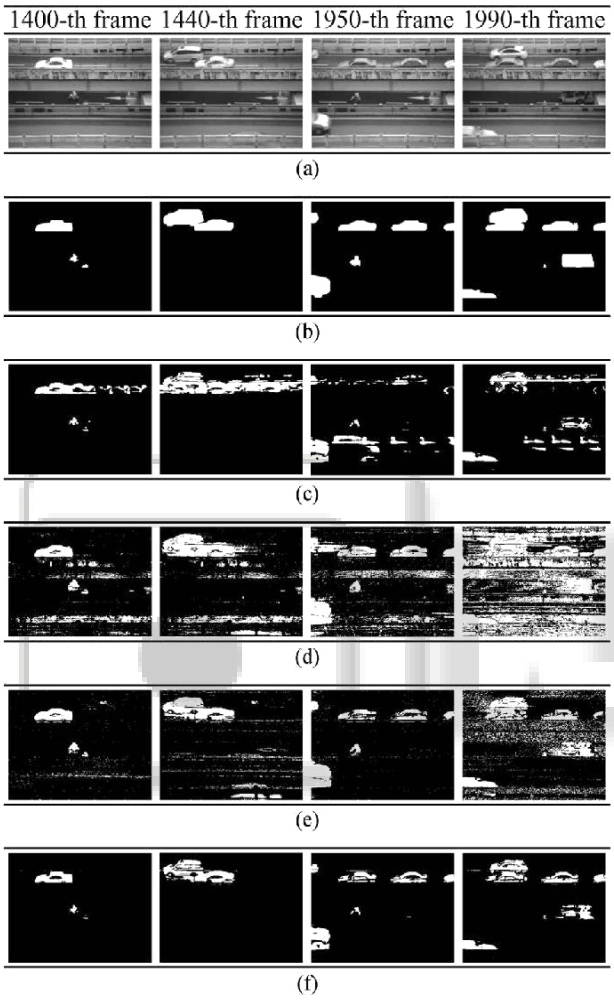

Fig. 7 : RD video sequence. (a) Four original RD frames. (b) Ground truths. The remaining four subpictures show the motion masks generated by the (c) MTD [15], (d) GMM [18], (e) MSDE [19], and (f) BMMC methods.

B. *Quantitative Measurement*

In addition to qualitative assessment, quantitative measure-ment is also desirable in experiments. Nevertheless, quantitative measurement of motion detection is not an easy task. In general, *Recall* and *Precision* are usually used to assess the accuracy of the binary motion mask [21]–[25].

Based on the *Recall* metric, the percentage of necessary true positives in the detected binary object mask can be evaluated as follows:

$$Recall = tp/(tp + fn) \quad (24)$$

where tp is the total number of true-positive pixels, fn is the total number of false-negative pixels, and (tp + fn) indicates the total number of true-positive pixels in the ground truth.

Fig. 8 shows four representative frames of the ST video sequence, ground truths, and binary mask of moving objects ob-tained by the MTD [15], GMM [18], MSDE [19], and BMMC methods, respectively. As shown in Fig. 8(a), this sequence features a simple scene with few moving objects on a street. Hence, most of the motion detection method can properly work. Unfortunately, the noises are always generated by the MSDE method, which modifies the threshold parameter whether the incoming pixel belongs to object or not.

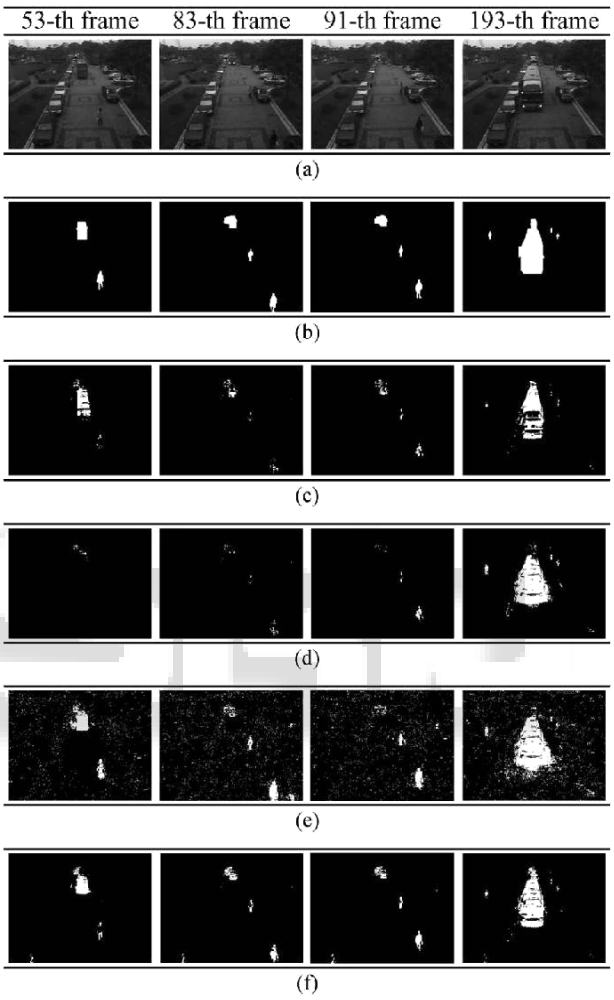

Fig. 8: ST video sequence. (a) Four original ST frames. (b) Ground truths. The remaining four subpictures show the motion masks generated by the (c) MTD [15], (d) GMM [18], (e) MSDE [19], and (f) BMMC methods.

On the other hand, the percentage of unnecessary true posi-tives can be evaluated using the following *Precision* metric:

$$Precision = tp/(tp + fp) \quad (25)$$

where fp is the total number of false-positive pixels, and (tp + fp) indicates the total number of true-positive pixels in the detected binary object mask.

Since *Recall* and *Precision* individually measure the per-centage of necessary and unnecessary true-positive pixels, the excess true-positive pixels may be associated with high *Recall*, and the missing true-positive pixels may be associated with high *Precision*. To facilitate an effective accuracy measurement, the harmonic means of *Recall* and

Precision can be achieved by using the F_1 and *Similarity* metrics, which are, respectively, expressed as follows:

$$F1 = 2(Recall)(Precision)/(Recall + Precision) \quad (26)$$

$$Similarity = tp/(tp + fp + fn) \quad (27)$$

Note that all metric-attained values range from 0 to 1, with higher values indicating greater accuracy.

Table. 1 lists the average accuracy rate of each method for all test sequences. As a result, the proposed BMMC method attains the highest *Similarity* and F_1 values compared with other motion detection methods. In particular, we can easily observe that our BMMC method attains greater accuracy rates of all metrics than 80% with regard to motion detection of "MSA" and "HA" sequences.

C. Computational Complexity

According to the Big-O notation, all implemented motion detection methods belong to $O(n)$, where n is the resolution of the video sequence. However, the Big-O notation can be only applied to measure the time complexity in mathematics, computer science, and related fields. This implies that the time consumption of each method should be reported in terms of exact performance evaluation.

To further demonstrate the efficiency of the proposed BMMC method, we evaluate the processing time of each motion de-tection method for each video. Note that all methods are implemented by the prototype C programming language on a desktop computer with a Core2 Duo central processing unit, 3.25 GB of random access memory, and running the Windows XP operation system.

Sequence	Evaluation	MTD	GMM	MSDE	BMMC
EW	*Similarity*	0.5695	0.5880	0.3878	0.7776
	F_1	0.7240	0.7200	0.5567	0.8745
	Precision	0.6633	0.6866	0.4014	0.9827
	Recall	0.8149	0.8136	0.9227	0.7882
HA	*Similarity*	0.7453	0.4509	0.1670	0.8277
	F_1	0.8415	0.5791	0.2765	0.9039
	Precision	0.8588	0.8442	0.1682	0.9173
	Recall	0.8279	0.5879	0.9563	0.8916
IR	*Similarity*	0.4412	0.6213	0.1822	0.6870
	F_1	0.6040	0.7556	0.2988	0.8122
	Precision	0.5493	0.7233	0.1915	0.8924
	Recall	0.7248	0.8268	0.8392	0.7476
MSA	*Similarity*	0.5618	0.8484	0.2058	0.8719
	F_1	0.7146	0.9141	0.3376	0.9312
	Precision	0.6045	0.9025	0.2084	0.9880
	Recall	0.8789	0.9414	0.9656	0.8812
RD	*Similarity*	0.4601	0.2480	0.4732	0.6757
	F_1	0.6153	0.3911	0.6303	0.8046
	Precision	0.5761	0.2539	0.5230	0.9133
	Recall	0.6904	0.9274	0.8426	0.7201
ST	*Similarity*	0.4536	0.4448	0.4112	0.6678
	F_1	0.5943	0.5511	0.5510	0.7945
	Precision	0.7444	0.7327	0.4598	0.8568
	Recall	0.5169	0.5375	0.7531	0.7437

Table 1 : Quantitative Measurement of Various Methods

Illustration

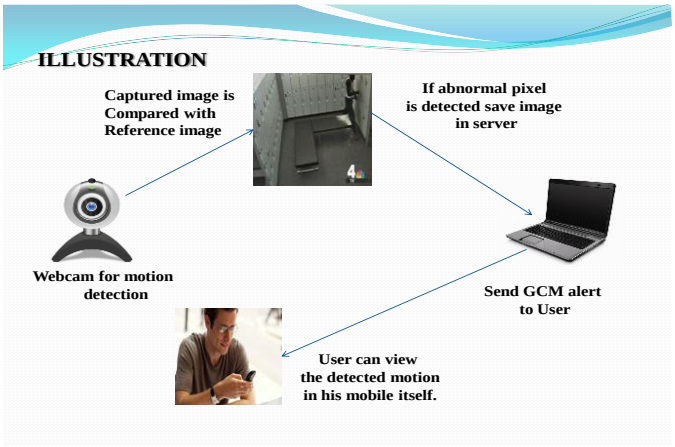

Fig. 9: Illustration of System

V. CONCLUSION

This paper introduced an approach for an effective video surveillance in the current system; this overcomes the traditional Surveying where Human intervention is needed and has to watch keenly for keeping track of the entire system. But now with this project we have introduced a unique technique which is a Major advantage to the old system. The accuracy of the surveillance is also increased with the implication of the Cauchy distribution model. The usage of Threshold level also enhances the accuracy of motion detection. This project also has an unique feature in which it sends GCM alert at once there is any sort of variation in the captured pixel. This paper is aimed at overcoming setbacks of existing system by enhancing the accuracy of motion detection and instant alert system through the use of GCM. As this project is application oriented, it finds many real-time applications in the field of surveillance.

REFERENCES

[1] M. Haag and H. H. Nagel, "Incremental recognition of traffic situations from video image sequences," *Image Vis. Comput.*, vol. 18, no. 2, pp. 137– 153, Jan. 2000.

[2] G. Ginesu, D. D. Giusto, and V. Margner, "Detection of foreign bodies in food by thermal image processing," *IEEE Trans. Ind. Electron.*, vol. 51, no. 2, pp. 480–490, Apr. 2004.

[3] S. Park and J. Aggarwal, "A hierarchical Bayesian network for event recognition of human actions and interactions," *Multimedia Syst.*, vol. 10, no. 2, pp. 164–179, Aug. 2004.

[4] A. C. Shastry and R. A. Schowengerdt, "Airborne video registration and traffic-flow parameter estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 4, pp. 391–405, Dec. 2005.

[5] J. Melo, A. Naftel, A. Bernardino, and J. Santos-Victor, "Detection and classification of highway lanes using vehicle motion trajecto-ries," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 2, pp. 188–200, Jun. 2006.

[6] H. Cheng, N. Zheng, X. Zhang, J. Qin, and H. Wetering, "Interactive road situation analysis for driver assistance and safety warning systems: Framework and algorithms," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 1, pp. 157–167, Mar. 2007.

[7] M. Castrillon, O. Deniz, C. Guerra, and M. Hernandez, "ENCARA2: Real-time detection of multiple faces at different resolutions in video streams," *J. Vis. Commun. Image R.*, vol. 18, no. 2, pp. 130–140, Apr. 2007.
[8] S. Cherng, C.-Y. Fang, C.-P. Chen, and S.-W. Chen, "Critical motion detection of nearby moving vehicles in a vision-based driver-assistance system," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 1, pp. 70–82, Mar. 2009.
[9] A. Licsar, T. Sziranyi, and L. Czuni, "Trainable blotch detection on high resolution archive films minimizing the human interaction," *Mach. Vis. Appl.*, vol. 21, no. 5, pp. 767–777, Aug. 2010.
[10] A. Amditis, E. Bertolazzi, M. Bimpas, F. Biral, P. Bosetti, M. Da Lio, L. Danielsson, A. Gallione, H. Lind, A. Saroldi, and A. Sjögren, "A holis-tic approach to the integration of safety applications: The INSAFES sub-project within the European framework programme 6 integrating project PReVENT," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 554–566, Sep. 2010.
[11] M. Ding, Z. Tian, Z. Jin, M. Xu, and C. Cao, "Registration using robust kernel principal component for object-based change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 761–765, Oct. 2010.
[12] J. L. Castro, M. Delgado, J. Medina, and M. D. Ruiz-Lozano, "An expert fuzzy system for predicting object collisions. Its application for avoiding pedestrian accidents," *Expert Syst. Appl.*, vol. 38, no. 1, pp. 486–494, Jan. 2011.
[13] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 34, no. 3, pp. 334–352, Aug. 2004.
[14] C.-C. Chang, T.-L. Chia, and C.-K. Yang, "Modified temporal differ-ence method for change detection," *Opt. Eng.*, vol. 44, no. 2, pp. 1–10, Feb. 2005.
[15] J.-E. Ha and W.-H. Lee, "Foreground objects detection using mul-
tiple difference images," *Opt. Eng.*, vol. 49, no. 4, p. 047 201, Apr. 2010.
[16] F. Barranco, J. Diaz, E. Ros, and B. Pino, "Visual system based on artificial retina for motion detection," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 3, pp. 752–762, Jun. 2009.
[17] A. Doshi and A. G. Bors, "Smoothing of optical flow using robustified diffusion kernels," *Image Vis. Comput.*, vol. 28, no. 12, pp. 1575–1589, Dec. 2010.
[18] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. IEEE Comput. Vis. Pattern Recog.*, 1999, pp. 246–252.
[19] A. Manzanera and J. C. Richefeu, "A new motion detection algorithm based on Σ− background estimation," *Pattern Recognit. Lett.*, vol. 28, no. 3, pp. 320–328, Feb. 2007.
[20] D.-M. Tsai and S.-C. Lai, "Independent component analysis-based back-ground subtraction for indoor surveillance," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 158–167, Jan. 2009.
[21] S. Park and J. Aggarwal, "A hierarchical Bayesian network for event recognition of human actions and interactions," *Multimedia Syst.*, vol. 10, no. 2, pp. 164–179, Aug. 2004.
[22] P. Remagnino, T. Tan, and K. Baker, "Multi-agent visual surveillance of dynamic scenes," *Image Vis. Comput.*, vol. 16, no. 8, pp. 529–532, Jun. 1998.
[23] Z. Zhu, G. Xu, B. Yang, D. Shi, and X. Lin, "VISATRAM: A real-time vision system for automatic traffic monitoring," *Image Vis. Comput.*, vol. 18, no. 10, pp. 781–794, Jul. 2000.
[24] S. L. Dockstader and A. M. Tekalp, "Multiple camera tracking of interact-ing and occluded human motion," *Proc. IEEE*, vol. 89, no. 10, pp. 1441– 1455, Oct. 2001.
[25] L. Maddalena and A. Petrosino, "A self-organizing approach to back-ground subtraction for visual surveillance applications," *IEEE Trans. Im-age Process.*, vol. 17, no. 7, pp. 1168–1177, Jul. 2008.

