Proposed Methods for Background Subtraction in Video Sequence

Rana Nour[1], Hossam El-Din Mostafa[2], Walaa Omar[3], Rasheed Mokhtar[4]

[1] Teaching Assistant of Electronics and Communication department, Mansoura Higher Institute for Engineering and Technology, Mansoura, Egypt
ranamnour@yahoo.com

[2] Lecturer of Communication Engineering Dept., Faculty of Engineering, Mansoura University, Al-Gomhoria St, Mansoura, Egypt
Hossam_moustafa@hotmail.com

[3] Lecturer of Electronics and Communication department, Mansoura Higher Institute for Engineering and Technology, Mansoura, Egypt
walaaomar78@yahoo.com

[4] Prof of Communication Engineering Dept., Faculty of Engineering, Mansoura University, Al-Gomhoria St, Mansoura, Egypt
actt_egypt@yahoo.com

***Abstract*—** Background subtraction (BS) is an essential part of many computer visions with several applications, as it is applied to distinguish moving objects in many video sequences. There are several algorithms such as Mixture of Gaussian, Frame Difference and Approximate Median are likened with synthetic and real videos. In this work, we improve the performance of the various algorithms by using a proposed morphological method and compare several background subtraction algorithms after and before the modification process. BS method also can be identified as foreground detection is applied by keeping the shape of the foreground object and eliminating and the noise.

Keywords—Background Subtraction (BS), Frame Difference (FD), Approximate Median Filter (AMF), Mixture of Gaussians (MOG).

I. INTRODUCTION

Background subtraction is often the first stage of more global computer vision systems. The main job of BS is distinguished moving objects from the background frame in images sequence. It has been widely studied since the 1990s and mainly for video sequence applications since they first need to detect persons, vehicles, animals, etc. [4]. Foreground detection also called background subtraction. The main goal of BS is comparing between recent frame against a previous frame that is called background image or background model and extract foreground objects from the difference. Over the years there a lot of algorithms were tested background subtraction and several algorithms were developed to fix a lot of problem in the previous methods. There are various types of datasets are implemented to test BS algorithms and to determine the performance of these algorithms. Many algorithms can be described by an easy flowchart shown in Fig.1 [2]. Preprocessing, background modeling, foreground detection and data validation are the four main basic steps in background subtraction algorithm.

Background modeling: Many studies have been applied to improving a model for background which is sturdy in the face of nature changes in the background, but sufficiently delicate to know all foreground objects [5]. This model gives a statistical explanation of the complete background scene.

Data validation: The method of developing the candidate foreground mask relied on data gained from external background model to decrease the falsely pixels.

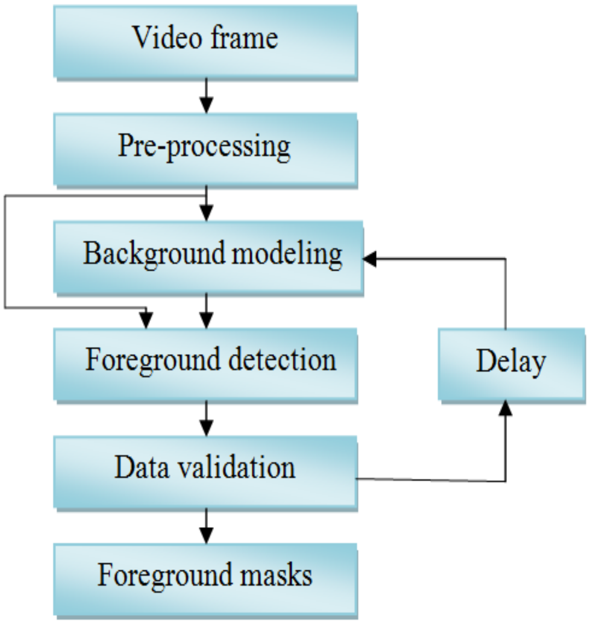


Fig.1. flow chart of Background Subtraction Algorithm

II. PREVIOUS WORK

Many of evaluations and reviews about BS are available in up to date literature.Few papers carry out evaluation experiments on background subtraction algorithms, on their own, used datasets, while some other did not evaluate or review Background Subtraction algorithms for the identified challenges. statistical background subtraction can be evaluated on this papers [16,17]. The main job of BS algorithms is to obtain the difference between the input frame and the estimated background image without moving objects to extract objects which are moving.Depending on this, Arguably that the definition of background subtraction is an uncomplicated method which subtracts the noticed frame from the estimated background frame and thresholded the result by generating the interest objects [18]. The process of comparison divides the frame into two complementary sets of pixels – a) moving objects in the foreground and b) the background, which is closely related. There are some of the previous methods which discuss various techniques of background subtraction algorithms like:

- Principal Component Analysis (PCA) has been presented by Oliver et al. [22]. This is a model where background subtraction is carried out by thresholding the result of subtraction between the generated background image and the recent image

- The codebook algorithm has been presented by M. Seki et al. [19]. In this algorithm, the pixel was considered as a codebook. The process of this algorithm was a compacted type of background model for along video frame.

- ViBe (Visual Background Extractor) algorithm has been presented by Barnich and Van Droogenbroeck [20].The main idea of this technique has stored a value of each pixel which it took in the past and then liken this to recent pixel to evaluate which pixel belongs to background. This technique performed computational cost and detection rate.

- Aspatio-temporal saliency algorithm has been presented by Mahadevan and Vasconcelos [8]. This algorithm is suitable for frames with extremely dynamic backgrounds, that it's obtained to achieve foreground detection.

- GRASTA (Grassmannian Robust Adaptive Subspace Tracking Algorithm) algorithm has been presented by He et al.[11].This algorithm was used to evaluate background subtraction.

- BLWS (block Lanczos with warm start) algorithm has been presented by Lin and Wei.[12], this algorithm is applied to distinguished changes in image sequences to evaluate foreground detection.

III. BACKGROUND SUBTRACTIONALGORITHMS

In this paper, we are studied and implemented three different methods of background subtraction algorithms that performed by MATLAB program. We represent approaches variable from a simplistic method such as mixture of Gaussians, approximate median filtering and frame difference.

A. FRAME DIFFERENCEING (FD)

Frame difference is an easy model in BS algorithm. The basic idea of this model is the difference between the current frame and the reference frame, and if the result of difference is greater than a threshold *Ths* then the pixel is categorized as foreground otherwise background [3]. The formula below describes the frame difference method:

$$\left| F(z,y)_T - B(z,y)_{T-1} \right| > Ths \qquad (1)$$

where $B(x,y)_{T-1}$ is the reference frame, $F(z,y)_T$ is the recent frame and is Ths the predefined threshold.

The result relies on threshold values. And also helps distinguish between real movement and noise.The advantage of this method is very easier and faster than other method but it has a disadvantage that it depends on the threshold value and the threshold is typically found empirically.

B. APPROXIMATION MEDIAN FILTER (AMF)

Arguably median filter is the one of the widest used techniques of background modeling. The basic idea of this technique is buffering the previous N frames of video, and the background is determine as the median of buffered frames. The background is subtracted from the recent frame and foreground pixels can be distinguished from threshold [13]. The approximate median is a development technique to median filtering. This technique works as follows: the pixel in background frame is increased by one if the pixel belong to background frame has a value samller than the pixel belong to recent frame, Likewise,the pixel belong to background frame is decreased by one if the pixel belong to background frame is larger than the recent pixel [1]. The formula below describes the approximate median method:

$$\text{If } (f(Z, X)_{i-1} > B(Z, X)_{i-1} \longrightarrow B(Z, X)_i = B(Z, X)_{i-1} + 1 \quad (2)$$

$$\text{If } (f(Z, X)_{i-1} < B(Z, X)_{i-1} \longrightarrow B(Z, X)_i = B(Z, X)_{i-1} - 1 \quad (3)$$

where $f(Z, X)_{i-1}$ is the pixel in recent frame and $B(Z, X)_{i-1}$ is the pixel in background frame.

The advantage of this method is not expensive method more in calculation and memory than frame difference. Thedisadvantage of this method is taken more time and comlex than frame difference.

C. MIXTURE OF GAUSSIAN (MOG)

This model is the most complex of all. This model relies on a Gaussian probability density function (PDF) to determine the pixel's value. Gaussian functions' number is defined as the location of each pixel which adds in one gathering to perform a pdf. We defined the pixel as mixture of Gaussian. Each pixel can be parameterized by mean µ, weight wi and the variance σ2 of each Gaussians function, we evaluate if it matches to background colors or foreground. Mixture of Gaussian is known as the weighted sum of M component Gaussian densities [1].

$$p(x|\lambda) = \sum_{i=1}^{M} w_i \, .g(x| \mu_i), \Sigma_i) \quad (4)$$

where p refers to the pdf of Gaussian, x refers to a continuous-valued D-dimensional data vector, wi refer to weight, $g(x|\mu i, \Sigma i)$ refer to Gaussian model and I ranges from 1 to M, µi refer to average (mean) vector and Σi refers to a covariance matrix.

The condition 1 can satisfied sum of the weights. Mixture of Gaussian depend on the parameters like mixture weights, the vector mean and covariance matrices from all densities function. The collective parameters are defined as

$$\lambda=\{wi,\mu i ,\Sigma i \} \; i = 1, \ldots M \quad (5)$$

It is considered foreground if the pixel does not like the background then.Mog based on probability density function (pdf) to determine the pixel value. This method compares between the cumulative average of the previous values against the current pixel's value. So it can keep an accumulative mean (µi) of the current pixel's values. If the result of subtraction between the pixel value of recent image and the accumulative pixel value is bigger than the result of (kσ) then it can be considered as foreground, The value of pixel is considered as foreground pixel otherwise, it can be considered as Background [5]. The formula below describes it :

$$|I - \mu_i| > k\,\sigma \longrightarrow \text{Foreground} \quad (6)$$

$$|I - \mu_i| < k\,\sigma \longrightarrow \text{Background} \quad (7)$$

where σ consider as standard deviation and k refer to a constant value.

At this point updating background as the running average:

$$\mu_{i+1} = \mu * I + (1- \alpha) * \mu_i \quad (8)$$

$$\Sigma^2_{i+1} = (I_t - \mu)^2 + (1- \alpha)\, t\, \sigma^2 \quad (9)$$

where α refer to the learning rate, is usually 0.05, I is recent pixel's value and μ_i indicate to the previous mean.

Table I shows that the different parameters which it can uses in background subtraction algorithms. Some of the parameters are variables and others are static (test parameters).

TABLEI : shows different parameters used in algorithms

Algorithms	Static parameters	Evaluate Parameters
Frame difference (FD)	No	Threshold of foreground (Ts)
Approximate median filter (AMF)	No	Threshold of foreground (Ts)
Mixture of Gaussian (MoG)	K=3 Variance = 36 wi = 0.1	Adaptation rate (σ) Weight threshold (T) Deviation threshold (D)

IV. PROPOSED METHODS

The technique of background subtraction is used to distinguish foreground objects with uncomplicated algorithms. After obtaining the background image, there are a great number of noises because of the dynamic background and holes in foreground objects. So we need to remove noise and enhance the image for detecting object correctly. Our proposed method will solve this problem by creating a filter mask. This filter merges between the morphological closing operation to eliminate small holes and median filter to remove noise. We will use this filter to refine the boundaries of moving objects.

V. EXPEREMENTAL RESULTS

A. *DATA SETS VIDEO*

Many datasets have been developed to compare the result and evaluate the performance of background subtraction algorithms.In this approach, we use BMC (Background Models Challenge) datasets. This dataset proposed by Vacavant et al. [6] are collected between real and synthetic videos, annotated with ground-truth data. Ground-truth data is Foreground masks which it is provided with the dataset. BMC dataset was implemented to test BS algorithms and gauge performances this datasets dedicated to the Workshop BMC organized within ACCV(Asian Conference On Computer Vision) 2012 [15]. In this paper, our dataset videos are representations of two scenes: a street and a rotary.

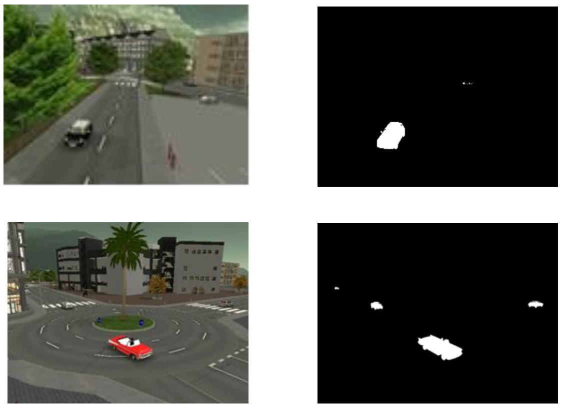


Fig.2 . Snapshots of a street and a rotarydata set and their associated ground truth

B. PERFORMANCE EVALUATION

The comparison of three background subtraction algorithms on MATLAB software has been done, this comparison has been performed using an Intel(R) Core(TM) i5-3230M CPU @ 2.60GHz processors with 4 GB of RAM. Our goal from evaluating performance is to detect the object correctly, we obtain the evaluation of the performance by comparing the foreground which it has been detected by algorithms to ground truth that it came with dataset. As shown in figure 3, green and red refer to the true and detected foreground [14]. For our data, we have four basic parameters for each dataset, false negatives (FN), false positives (FP), true negative (TN) and true positives (TP).These parameters can be described as shown in figure3. They have considered the measurement parameters as follows [7]:

1. True Positive (TP): which refer to the number of pixels in foreground that should have been detected correctly by the algorithms.

2. False Positive (FP): which refer to the number of falsely distinguished pixels in background as foreground.

3. True Negative (TN): which refer to the number of pixels in background that should have been detected correctly as background scene.

4. False Negative (FN): which refer to the number of falsely distinguished foreground pixels as background. Figure 3 can derive easily four basic parameters.

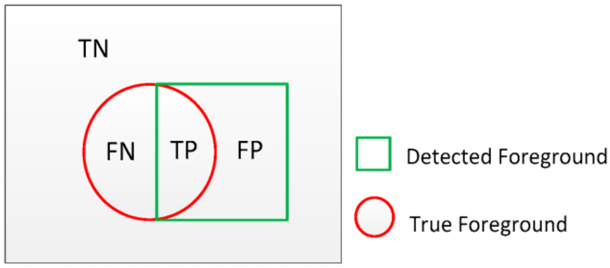


Fig.3. True and detected foreground in background subtraction

We can use this parameteres to evaluate :

➢ Recall can be considered as the number of pixels correctly detected in foreground (true positives) divided by the total number pixels belongs to foreground in the ground truth. In a more formal way:

$$\text{Recall} = \frac{\text{TP}}{TP+FN} \qquad (larger\ is\ better) \tag{10}$$

In other words,it can be rewritten as:

$$\text{Recall} = \frac{number\ of\ correctly\ foreground\ pixels}{number\ of foreground\ pixils\ in\ ground\ truths} \tag{11}$$

➢ Specificity (Spc) or True Negative Rate (TNR) which identify as the proportion of pixels belongs to foreground correctly distinguished that are true positives. In a more formal way:

$$\text{Specificity} = \frac{\text{TN}}{TN+FP} \qquad (larger\ is\ better) \tag{12}$$

➢ Negative Predictive Value (NPV) that consider as the proportion of pixels belong to background correctly detected. In a more formal way:

$$\text{NPV} = \frac{\text{TN}}{TN+FN} \qquad (larger\ is\ better) \tag{13}$$

➢False Alarm Rate (FAR) or False Discovery Rate (FDR) which determine the pixels belong to foreground an erroneously tagged as background. In a more formal way:

$$\text{FDR} = \text{FA} = \frac{\text{FP}}{FP+TP} \qquad (smaller\ is\ better) \tag{14}$$

➢False Positive Rate (FPR) that indicates to pixels belong to background misclassified as foreground.In a more formal way:

$$\text{FPR} = \frac{\text{FP}}{FP+TN} \qquad (smaller\ is\ better) \tag{15}$$

➢False Negative Rate (FNR) measures the pixels belong to background misclassified as background . The formula describes it:

$$\text{FNR} = \frac{\text{FN}}{FN+TP} \qquad (smaller\ is\ better) \qquad (16)$$

C. RESULTS AND DISCUSSION

As we discussed before there is a lot of problem in the result of background subtraction algorithms such as the holes and noise in the image sequences. So we remove the holes appear in the foreground by using the modification that discussed before in proposed method. To gauge the performance of BS algorithms, we choose the BMC dataset. Vacavant et al.[6] presented this dataset. Our methods have been likened with other algorithms such as GRASTA algorithm [11], BLWS algorithm [12] and MOG algorithm [9]. Figure 4 and 5 shows that the results of the three methods (Input image, Ground-truth-MOG image and proposed FD image and proposed, AMF and proposed) we can see the difference between old and proposed methods by our eyes and we can touch it. In Table II,III,IV and V we evaluate performance. Table II and IV reveal that the proposed method is better than the previous methods, with the highest Recall best scores overall the result of other previous methods. The proposed method is done by using Morphological closing filtering. The function of the morphological closing filtering is removing the small regions probably generated by noise, fill up holes. It will give pixel level operations. In Table 3and 5 show that the other performances that have been evaluated.

TABLEII : Average Recall Comparison Of Rotary Dataset

Method	Recall
GRASTA [11]	0.748
BLWS [12]	0.680
MOG [9]	0.59
FD	0.4626
Proposed FD	**0.9165**
AMF	0.6951
Proposed AMF	**0.9507**
MOG	0.9044
Proposed MOG	**0.9906**

TABLEIII : Average Performance Result Of Rotary Dataset

Our Method	SPF	NPV	FDR	FNR	FPR
FD	0.9984	0.9993	0.2292	0.3731	0.0015
Proposed FD	**0.9985**	**0.9959**	**0.1712**	**0.0834**	**0.0011**
AMF	0.9975	0.99368	0.4591	0.3048	0.0063
Proposed AMF	**0.9977**	**0.9995**	**0.2470**	**0.0492**	**0.0022**
MOG	0.9835	0.9991	0.6896	0.0047	0.0156
Proposed MOG	**0.9988**	**0.9999**	**0.2101**	**0.0013**	**0.0018**

TABLEIV : Avarage Recall Comparison Of Street Dataset

Method	Recall
GRASTA [11]	0.787
BLWS [12]	0.786
MOG[9]	0.63
FD	0.49132
Proposed FD	**0.99087**
AMF	0.67439
Proposed AMF	**0.91487**
MOG	0.85944
Proposed MOG	**0.99613**

TABLEV : Avarage Performance Result Of Street Dataset

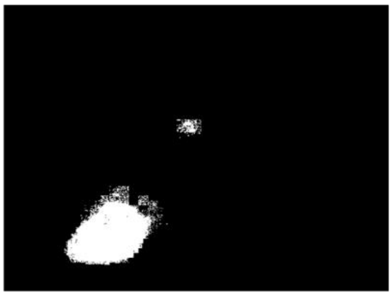
Our Method	SPF	NPV	FDR	FNR	FPR
FD	0.9974	0.9972	0.2848	0.5086	0.0025
Proposed FD	**0.9989**	**0.9998**	**0.2328**	**0.0409**	**0.0010**
AMF	0.9985	0.9983	0.4410	0.3256	0.0014
Proposed AMF	**0.9994**	**0.9996**	**0.3242**	**0.0851**	**0.0035**
MOG	0.9692	0.9795	0.753	0.1235	0.0100
Proposed MOG	**0.9996**	**0.9995**	**0.3134**	**0.0147**	**0.0010**

VI. CONCLUSION

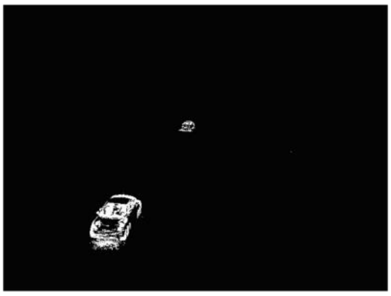
In this paper, three background subtraction Algorithms have been performed. We evaluated them based on the performance parameters such as recall and Specificity. The following algorithms and proposed methods have been tested on BMC dataset: MOG, AMF and FD. The best results of these algorithms are mixture of Gaussians algorithm to the other two algorithms. We use proposed methods to overcome the problems that we faced on the following algorithms. The results show that our proposed methods improved the performance.

:

(a)



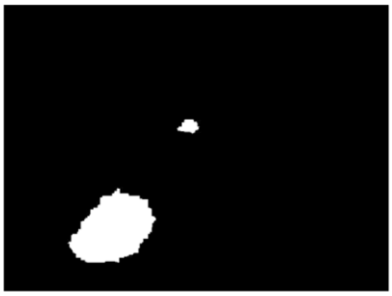
(b)



(c)

(d)

(e)



(f)

(g)

(i)

Fig.4.the background subtraction algorithms output of Street dataset (a) Input image, (b) MOG output,(c) FD output, (d) AMF output, (e) Ground Truth ,(f) Proposed MOG output ,(g) Proposed FD output, (i) Proposed AMF output

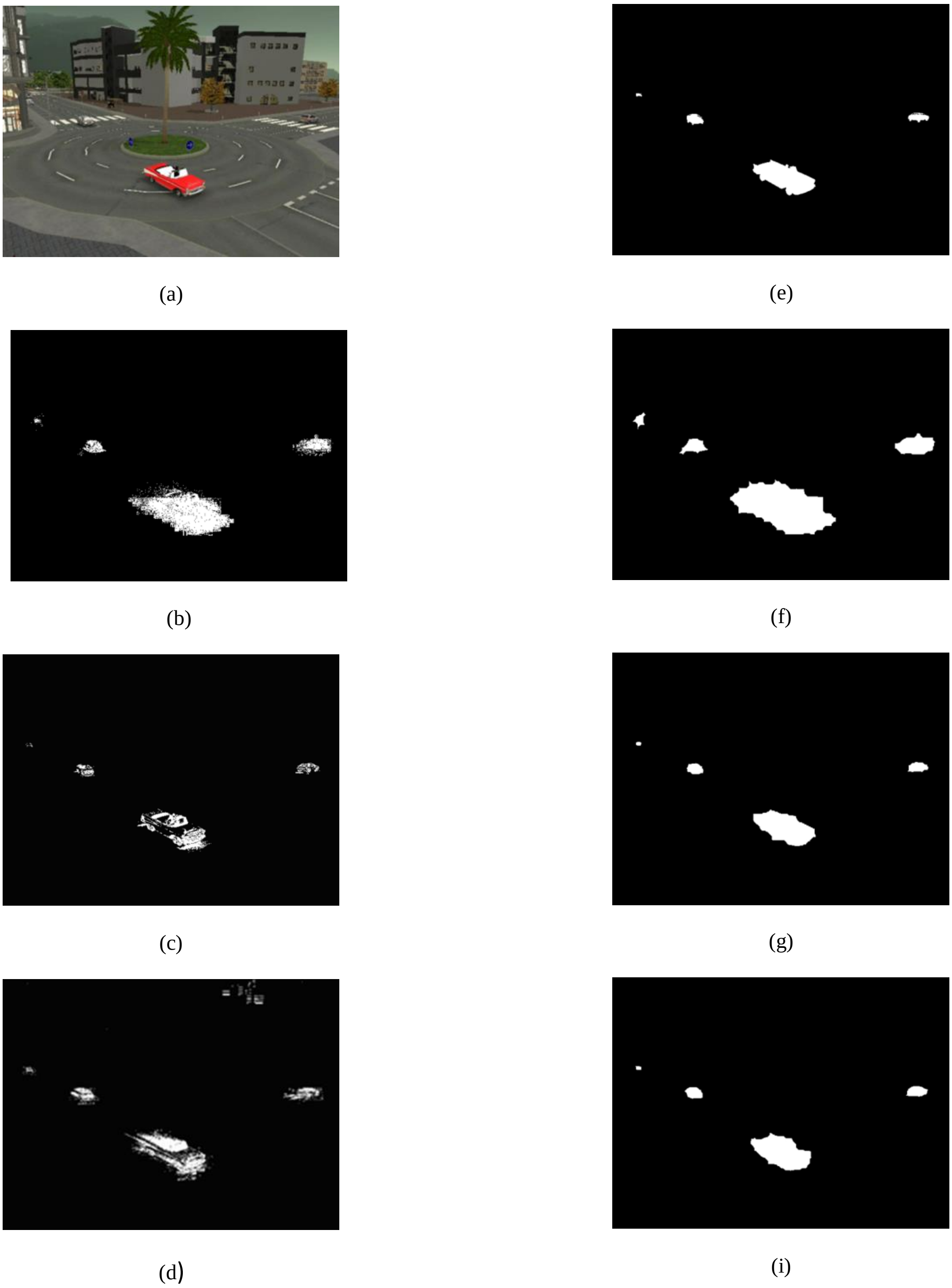


Fig.5.the background subtraction algorithms output of Rotary dataset (a) Input image, (b) MOG output,(c) FD output, (d) AMF output, (e) Ground Truth ,(f) Proposed MOG output ,(g) Proposed FD output, (i) Proposed AMF output

REFERENCES

[1] S. M. DhanyaStudent of M.E VLSI DesignSri Ramakrishna Engineering College Coimbatore, India and N. KirthikaAssistant Professor,Department of ECE (PG- VLSI DesignSri Ramakrishna Engineering CollegeCoimbatore, India. "Survey on Background Identification Using Various Algorithms",(IJERT) Vol. 4 Issue 01,2015algorithms." arXiv preprint arXiv:1405.1815 (2014).

[2] Sobral, Andrews, and Antoine Vacavant."A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos." Computer Vision and Image Understanding 122 (2014): 4-21.

[3] Alawi, Mahmoud Abdulwahab, Othman O. Khalifa, and MD Rafiqul Islam. "Performance Comparison of Background Estimation Algorithms for Detecting Moving Vehicle." World Applied Sciences Journal 21 (2013): 109-114.

[4] A. Vacavant, T. Chateau, A. Wilhelm, L. Lequièvre, " A benchmark dataset for foreground/background extraction, in: Background Models Challenge (BMC) " at Asian Conference on Computer Vision (ACCV), LNCS, vol. 7728, Springer, 2012,pp. 291–300.

[5] R. Manikandan and R. Ramakrishnan. "Video Object Extraction by Using Background Subtraction Techniques for Sports Applications ",CiiT International Journal of Digital Image Processing, Vol 5, No 9, September 2013.

[6] N.M. Oliver, B. Rosario, and A.P. Pentland. A bayesian computer vision system for modeling human interactions. IEEE Trans. on Pattern Analysis and Machine Intelligence, 22(8):831{843, Aug 2000.

[7] Pennisi, Andrea, et al. "Real-time adaptive background modeling in fast changing conditions." Advanced Video and Signal Based Surveillance (AVSS), 2015 12th IEEE International Conference on. IEEE, 2015.

[8] Zhang, Ranran, Xiaoyong Liao, and Jiajun Xu. "A Background Subtraction Algorithm Robust to Intensity Flicker Based on IP Camera." Journal of Multimedia 9.10 (2014).

[9] J. He, L. Balzano, and J. C. S. Lui. Online robust subspace tracking from partial information. CoRR, abs/1109.3827, 2011.

[10] Z. Lin and S. Wei. A block lanczos with warm start technique for accelerating nuclear norm minimization algorithms. CoRR, abs/1012.0365, 2010.

[11] Benton, Seth. "'Background subtraction, part 1: Matlab models'." (2008).

[12] Yao, Guangle, et al. "Comparative Evaluation of Background Subtraction Algorithms in Remote Scene Videos Captured by MWIR Sensors." Sensors 17.9 (2017): 1945.

[13] Vacavant, A., et al. "Computer Vision-ACCV 2012 Workshops: ACCV 2012 International Workshops, Daejeon." Korea, November (2012): 5-6.

[14] Bouwmans, T.; Baf, E.F.; Vachon, B. Background modeling using mixture of Gaussians for foregrounddetection-a survey. Recent Pat. Comput. Sci. 2008, 1, 219– 237. [CrossRef].

[15] Bouwmans, T. Recent Advanced statistical background modeling for foreground detection: A systematic survey. Recent Pat. Comput. Sci. 2011, 4, 147–176.

[16] Zhang, Ranran, Xiaoyong Liao, and Jiajun Xu. "A Background Subtraction Algorithm Robust to Intensity Flicker Based on IP Camera." Journal of Multimedia 9.10 (2014).

[17] M. Seki, T. Wada, H. Fujiwara, et al. "Background subtraction based on co-occurrence of image variations,

Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2003),pp. 65-72, 2003.

.

[18] O. Barnich and M. Van Droogenbroeck."ViBe:a universal background subtraction algorithm for video sequences," IEEE Transaction on Image Processing, vol. 6, pp.1709-1724, June 2011.

[19] Mahadevan and N. Vasconcelos. "Spatiotemporal saliency in dynamic scenes," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, pp. 171-177, 2010.

[20] O. Barnich and M. Van Droogenbroeck."ViBe:a universal background subtraction algorithm for video sequences," IEEE Transaction on Image Processing, vol. 6, pp.1709-1724, June 2011.

[21] V. Mahadevan and N. Vasconcelos. "Spatiotemporal saliency in dynamic scenes," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, pp.171-177, 2010.