International Journal of Computer and Information Technology (ISSN: 2279 – 0764)
Volume 05 – Issue 06, November 2016

Empirical Analysis of GMM over Wallflower Dataset

Lajari S. Alandkar
Electronics Department,
Walchand Institute of Technology Solapur
Solapur, MH, India
Email: lajari.alandkar [AT] gmail.com

Sachin Gengaje
Electronics Department,
Walchand Institute of Technology Solapur
Solapur, MH, India

Abstract—**Gaussian mixture Model (GMM) based Background Subtraction (BS) is widely popular method for detection of moving object. Selection of important GMM parameters such as learning rate (α) and threshold (T) is crucial step for satisfactory performance of GMM. Present study primarily focuses on empirical analysis of GMM w.r.t theses parameters. This study analyzes performance of GMM using precision recall curve. GMM is evaluated over Wallflower dataset which cover different critical scenarios. This empirical study reveals strength and weaknesses of GMM to handle various critical situations. It also helps researchers for appropriate selection of GMM parameters based on applications.**

Keywords- Background Subtraction, GMM, Precision Recall Curve

I. INTRODUCTION

Stauffer and Grimson [1] defined Gaussian Mixture Model (GMM) for background representation. This model has ability to model multiple colors in the background. Adaptive nature of background model is capable to adjust with dynamic changes in background [1, 7]. These two features of GMM increase the robustness of Background Subtraction (BS) method against various critical scenarios such as: rippling water, waving tree, flickering monitor, rain, snow fall, daylight changes in outdoor scenario etc. However, satisfactory performance of BS requires proper tuning of important parameters learning rate (α) and threshold (T) [1, 2]. Manual tuning of these parameter is difficult if background is complex i.e. contains large amount of dynamic changes [2]. Traditionally, theses parameters are selected empirically but it is difficult for new researcher to implement large number of repetitive experimentation to find optimum settings [4].

This paper implemented empirical study of GMM with different setting of α and T. Precision Recall Curve (PRC) is compact way to analyses the empirical behavior of any algorithm [3]. In this context, PRC has been used to represent performance of GMM for different values of learning rate (α). GMM has been evaluated over Wallflower dataset that covers different critical situations. Therefore, this study provides basic guidance to new researcher for appropriate selection of parameters during critical background condition. Empirical analysis also depicts the ability of GMM based BS to handle different background situation.

This paper is divided in five sections. Second section describe GMM based BS in brief. Section three provides material and methods. Fourth section contains detail empirical analysis of GMM over seven canonical sequences in Wallflower Dataset. Fifth section includes analysis of GMM performance w.r.t various background situations. Sixth section concludes the paper.

II. GMM BASED BS

To account various colors in background due to frequent dynamic changes, Stauffer and Grimson [1] used mixture of Gaussians for background modelling. Every pixel has been modeled by mixture of K Gaussians. Probability of observing current pixel value is

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t}\eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \qquad (1)$$

Where, K is the number of gaussians, $\omega_{i,t}$ is weight of i^{th} Gaussian in the mixture at time t, $\mu_{i,t}$ and $\Sigma_{i,t}$ is mean value and covariance respectively of i^{th} Gaussian in the mixture at time t, and where η is a Gaussian probability density function. Covariance is assumed to be diagonal matrix for simplified computation.

In this method [1, 4], parameters of each matched component (i.e. the Gaussian model for which X_t is within 2.5 standard deviations of its mean) is updated as follows

$$\omega_{k,t} = (1-\alpha)\omega_{k,t-1} + \alpha(M_{k,t}) \qquad (2)$$

$$\mu_t = (1-\rho)\mu_{t-1} + \rho X_t \qquad (3)$$

$$\sigma_t^2 = (1-\rho)\sigma_{t-1}^2 + \rho(X_t - \mu_t)^T(X_t - \mu_t) \qquad (4)$$

Where $\rho = \alpha\eta(X_t|\mu_k, \sigma_k)$

Where α is learning rate and $M_{k,t}$ is 1 for model which matched and 0 for remaining models. The μ and σ parameters of unmatched distributions remain the same while their weight is reduced by equation (1). Whenever no component matches $\boldsymbol{X_t}$ the one with lowest weight is replaced by a Gaussian with mean $\boldsymbol{X_t}$ a large initial variance σ_0 and a small weight ω_0. Once every Gaussian has been updated, the K weights $\omega_{i,t}$ are normalized so they sum up to 1. Then, the K distributions are ordered based on a ratio $\omega_{i,t}/\sigma_{i,t}$ and only the B most reliable are chosen as part of the background :

$$B = argmin_b\left(\sum_{k=1}^{b} \omega_k > T\right) \qquad (5)$$

T is threshold which is known as minimum background ratio [2]. T is kept low for simple unimodal background while it must be sufficiently high for complex background containing repetitive motion such as waving tree, rain, snow, flickering monitor etc. [2, 4]

III. Materials and Methods

In this section, details of video dataset and experimental setup are explained.

A. Video Dataset

Wallflower dataset has been chosen for GMM evaluation since dataset covers common critical situations in background scenario. Wallflower database is open source database [6]. It includes seven set of image sequences covering possible critical situations in background such as: Moved Object (MO), Waving tree (W), Camouflage (C), Bootstrap (B), Foreground Aperture (FA), Time of Day (TOD), Light Switch (LS). All sequences are of size 160x120 pixels, sampled at 4Hz [5, 6]. Sequence MO is excluded while experimentation as its ground truth image doesn't involve single foreground pixel and therefore it is impossible to obtain PRC.

B. Experimental Setup

Total experimental framework is developed over MATLAB 2012 using computer vision toolbox. Object detection system is implemented using GMM based BS as in [1]. Speckle noise removal is applied to the output of object detection which eliminates 4-connected components of area less than 8.

Our aim is to evaluate the ability of setting of α and T to correctly detect foreground. Ground truth provided in the database allows the evaluation of true positives (TP), false positives (FP) and false negatives (FN) numbers. Those values are combined into a (Precision/Recall) couple defined as:

$$Precision = \frac{TP}{TP+FP} \qquad Recall = \frac{TP}{TP+FN} \qquad (5)$$

A good setting is one which produces high precision and recall value. Since single (α, T) setting produces a single Precision / Recall couple per video, 10 different thresholds are used to produce curve for fixed value of α. Five values of α are considered to draw five different curves. In this way, the comparison between different settings has been implemented for each video in Wallflower dataset. In this context, setting (α, T) is assumed to be good if it provides at least 0.5 precision and recall value.

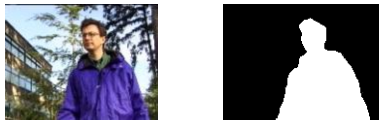
Other parameter i.e. number of Gaussians (K) and initial variance σ_0 are fixed to 4 and 0.0137.

IV. Empirical Analysis of GMM

GMM is evaluated with 50 different settings of (α, T) over each sequence in Wall flower dataset. For this experimentation, range of α is chosen from 0.001 to 0.1; T is from 0.5 to 0.95. Therefore it is important to note that, in this context, 'Any/all value of α and T' is referred as any values in their corresponding range. Empirical analysis corresponds to each video is presented in following subsections.

A. For Waving Tree (WT)

This video mainly involves swaying tree and a person walks in front of the tree (Refer Fig 1). Hand segmented ground truth image is provide in Dataset.



(a) (b)

Figure 1 Waving Tree Video. (a) sample frame (b) its ground truth image

PRC curves for different value of α such as: 0.001, 0.005, 0.01, 0.05, and 0.1 are shown in Fig 2.

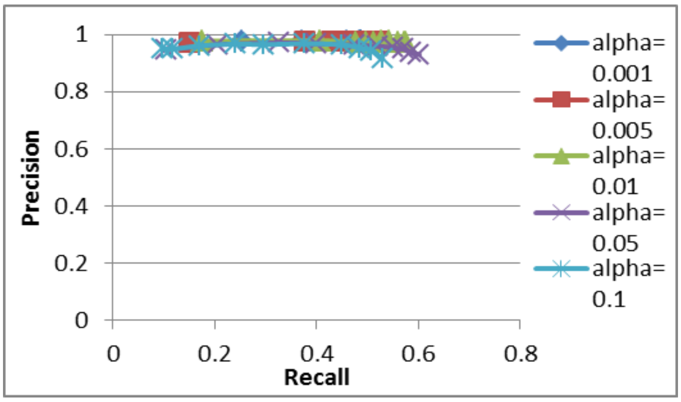


Figure 2 PRC of GMM for Waving Tree

It is observed that precision for all α's is always greater than 0.8 for any value of T whereas; recall is nearly or above 0.5 for most cases of T. It shows that careful selection of T is important to account modes in the background rather than α to adapt background. Therefore, T plays important role to produce desired result over multimodal background. It also reveals that GMM performs well for multimodal background by providing satisfactory detection of foreground object for most settings.

B. For Camouflage (C)

This video involves a person walks in front of a monitor, which has rolling interference bars on the screen. The bars include similar color to the person's clothing (Refer Fig 3)

(a) (b)

Figure. 3 Camouflage video (a) sample frame (b) its ground truth image

PRC curve corresponds to this video is shown in Fig. 4

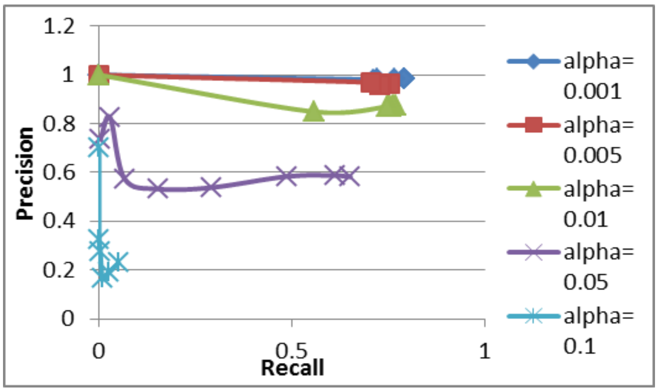
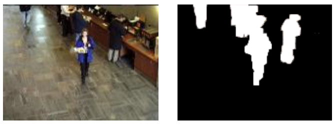


Figure 4 PRC of GMM for Camouflage

PRC corresponds to α's <= 0.01 produce desired result w.r.t precision. Desired recall would be obtained for most values of T. Therefore, appropriate selection of both parameters is important to produce desired detection. This analysis also indicates that GMM performs well for camouflage like situation with high detection rate.

C. *For Bootstrap(B)*

This image sequence shows a busy cafeteria and each frame contains people (Refer Fig 5). No single frame is available in training sequence without moving object.



(a) (b)

Figure. 5 Bootstrap video (a) sample frame (b) its ground truth image

PRC curve corresponds to bootstrap video is shown in Fig 6

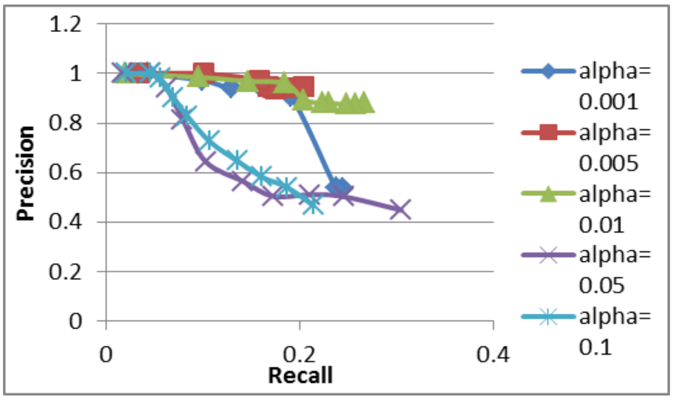
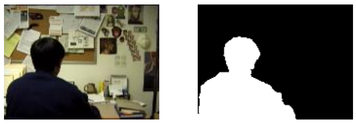


Figure 6 PRC of GMM for Bootstrap

It is observed that overall performance of GMM is worst as no single setting could be able to produce recall greater than 0.5 even if acquired precision is high. This is caused due to lack of pure training sequence during background initialization. Maximum possible recall obtained by one of the setting is just 0.3 for which precision is too low i.e. below 0.5. Therefore, it is concluded that GMM is not suitable for bootstrap situation.

D. *For Foreground Aperture (FA)*

This video includes a person with uniformly color shirt wakes up and begins to move slowly (Refer Fig 7).



(a) (b)

Figure. 7 Foreground aperture video (a) sample frame (b) its ground truth image

PRC curve corresponds to Foreground Aperture is shown in Fig 8.

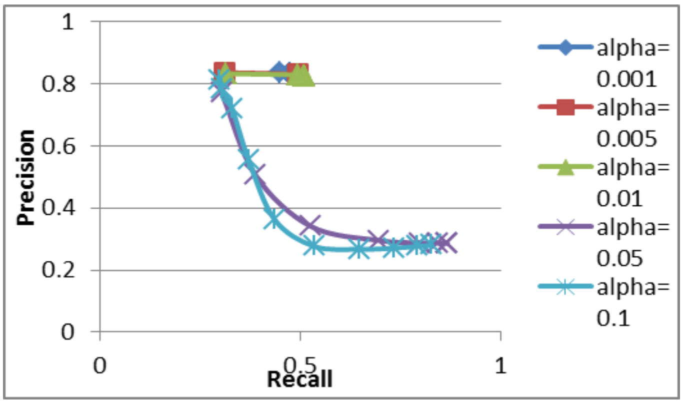
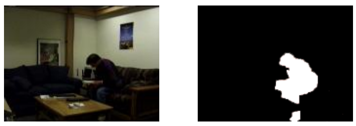


Figure 8 PRC of GMM for Foreground Aperture

GMM performs moderate for foreground aperture situation as only few settings are able to provide recall nearly 0.5 with high precision. It is caused due to foreground object present in training sequence during the background initialization. Therefore, it is concluded that, programmer may require to perform large number of experimentation to find optimum settings.

E. *For Time of Day (TOD)*

In this scenario, the light in a room gradually changes from dark to bright. Then, a person enters the room and sits down. (Refer Fig 9)



(a) (b)

Figure 9 Foreground aperture video (a) sample frame (b) its ground truth image

PRC curve corresponds to time of day is shown in Fig 10.

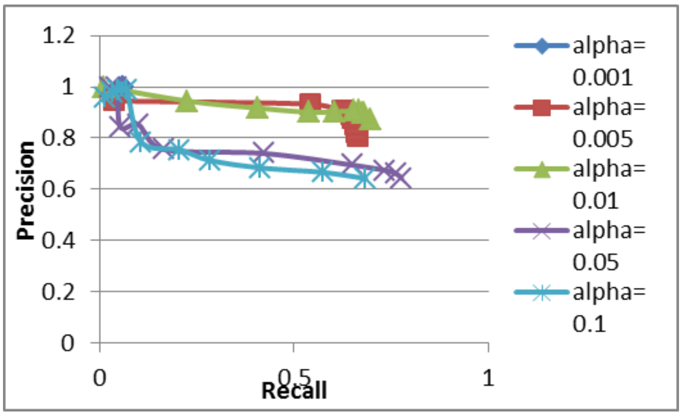
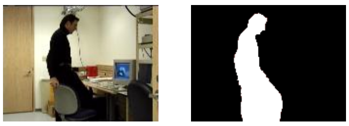


Figure 10 PRC of GMM for Foreground Aperture

Most of the settings are able to provide satisfactory performance of GMM as precision and recall are both greater than 0.5. It is due to the adaptive nature of GMM to cope up with gradual changes. Therefore, smaller setting of α less than 0.01 is suitable to deal with such kind of situations. It is concluded that, GMM deals robustly with the slower illumination change and able to produce satisfactory performance on most different settings.

F. For Light Switch(LS)

In this image sequence, a room scene begins with the lights on. Then a person enters the room and turns off the lights for a long period. Later, a person walks in the room, switches on the light, and moves the chair, while the door is closed. Refer Fig 11 for sample frame.



(a) (b)

Fig 11 Light Switch Video. (a) sample frame (b) its ground truth image

PRC curve corresponds to Light Switch is shown in Fig 12.

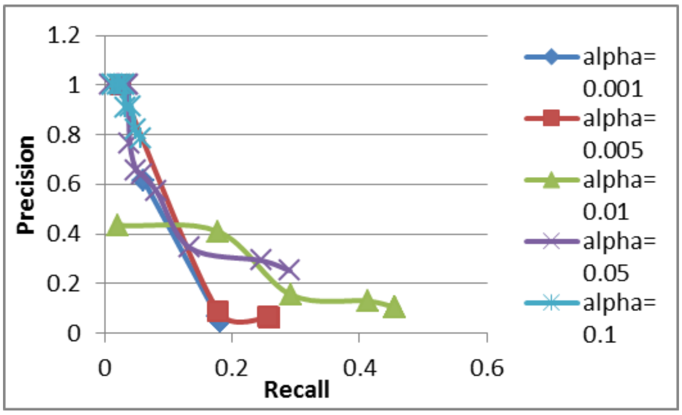


Figure 12 PRC for Light Switch

GMM is worst to deal sudden light change since no single setting is providing high precision and high recall value. Performance is worst since background initialization is implemented over training sequence that contains absence of light and after initialization when sudden light change occurs in a sequence then that background is appears as totally new background for GMM. This analysis helps to reach to the conclusion that GMM is not at all suitable to handle sudden illumination change.

Table I shows foreground masks for all sequences for two different settings with Highest Precision (HP), Recall (R) and Highest Recall (HR), precision(P) respectively.

V. GMM Performance Analysis

To conclude about GMM performance to handle six typical cases each setting individually analyzed for Precision (P) and Recall (R) value. For particular setting, if (P, R) ≥ (0.5, 0.5) then it is assumed as good setting; otherwise it is considered as bad setting. GMM is executed with 50 different settings. All are classified in good/ bad category based on their corresponding P, R value (Refer Table II). GMM is assumed to be good to handle particular situation if it has more than 50 percent of total settings are good. It is assumed to be perform moderate if 30 to 50 percent of total settings are classified as good. Except these two conditions GMM is assumed to be worst to deal particular situation.

TABLE II. GMM performance based on setting (alpha, T)

Video	Number of Bad Settings	Number of Good Settings	Remark about GMM performance
WT	32	18	Moderate
C	21	29	Good
B	50	0	Worst
FA	34	16	Moderate
TOD	18	32	Good
LS	50	0	Worst

VI. Conclusion

Empirical study of GMM based background subtraction method w.r.t learning rate (α) and threshold (T) is implemented in this paper. Wallflower dataset is chosen for GMM evaluation as it covers wide variety of background conditions. Study mainly involves evaluation of GMM with 50 different setting of (α, T) for each video in Wallflower dataset. Performance of GMM for each video is analyzed using precision recall curve. This analysis helps researcher to decide appropriate selection of (α, T) based on their application. GMM is analyzed for robustness to handle different critical situations using precision recall values obtained by GMM. This analysis indicates that GMM is well suited for slower illumination change and camouflage. It performs moderately for multimodal background and foreground aperture whereas gives worst performance for sudden illumination change and bootstrap like situations.

REFERENCES

[1] Stauffer C, Grimson W. "Adaptive background mixture models for real-time tracking". Proc. IEEE Conf. on Comp Vision and Patt. Recog.(CVPR 1999) 1999; 246-252

[2] Qi Zang and Reinhard Klette, "Parameter Analysis for Mixture of Gaussians Model", www.citr.auckland.ac.nz/researchreports/CITR-TR-188.pdf

[3] Yannick Benezeth, Pierre-Marc Jodoin, Bruno Emile, H Laurent, Christophe Rosenberger. Review and evaluation of commonly-implemented background subtraction algorithms. International Conference on Pattern Recognition, Dec 2008, Tampa, United States. IEEE, 2008

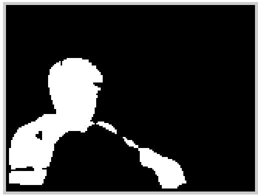
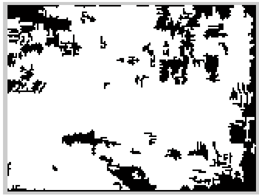
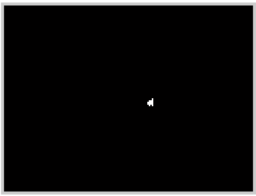
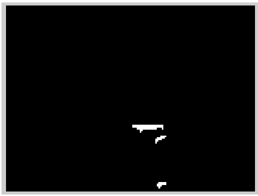
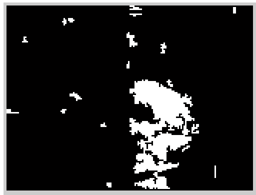
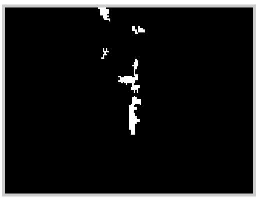
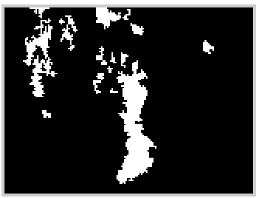
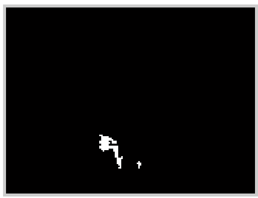
[4] Brandyn White B, Shah M. Automatically tuning background subtraction parameters using particle swarm optimization. IEEE Int Conf on Multimedia & Expo (ICME 2007), Beijing, China, 2007; 1826-1829.

[5] Wallflower Dataset: http://research.microsoft.com/users/jckrumm/WallFlower/TestImages.htm

[6] Toyama K, Krumm J, Brumitt B, Meyers B. Wallflower: Principles and practice of background maintenance. Int. Conf. on Computer Vision, (ICCV 1999), Corfu, Greece, September 1999; 255 -261

[7] Thierry Bouwmans, Fida El Baf, Bertrand Vachon. "Background Modeling using Mixture of Gaussians for Foreground Detection - A Survey". Recent Patents on Computer Science, Bentham Science Publishers, 2008, 1 (3), pp.219-237

TABLE I. Foreground mask using GMM for different settings (A, T) based (HP, R) and (P,HR). P- Precision, R-Recall, HP- High Precision, HR-High Recall

Foreground Mask using GMM **(α, T)**			
Waving Tree (W)		**Foreground Aperture (FA)**	
			
(HP, R) = (0.978, 0.401) (α, T) = (0.01, 0.9)	(P, HR) = (0.929,0.601) (α, T) = (0.05, 0.5)	(HP, R) = (0.834,0.476) (α, T) = (0.001,0.7)	(P, HR) = (0.286,0.867) (α, T) = (0.005,0.55)
Camouflage (C)		**Time of day (TOD)**	
			
(HP, R) = (1, 0.001) (α, T) = (0.001, 0.95)	(P, HR) = (0.982,0.791) (α, T) =(0.7,0.001)	(HP, R) = (1,0.061) (α, T) = (0.001,0.7)	(P, HR) = (0.644,0.775) (α, T) =(0.05,0.5)
Bootstrap (B)		**Light Switch (LS)**	
			
(HP, R) = (1,0.101) (α, T) = (0.005,0.9)	(P, HR) = (0.448,0.304) (α, T) =(0.05,0.5)	(HP, R) = (1,0.0368) (α, T) = (0.05,0.85)	(P, HR) = (0.456,0.251) (α, T) =(0.01,0.5)