International Research Journal of Pure Algebra-7(4), 2017, 542-547
Available online through www.rjpa.info ISSN 2248–9037

INITIAL PARAMETERS ANALYSIS FOR MIXTURE OF GAUSSIAN MODEL

LIHONG WANG*, SHANLIN WANG, RUIMIN WU

Longqiao College of Lanzhou University of Finance and Economics, Lanzhou, China, 730101.

(Received On: 24-03-17; Revised & Accepted On: 31-03-17)

ABSTRACT

The mixture of Gaussian is a common model for background subtraction. There are several parameters in such a model. Obviously, the assignment of initial values to these parameters affects the accuracy of background subtraction. In this paper, we analyze in detail the impact of different initial parameter values based on the EM algorithm. The tested results of waving trees video sequences have illustrated. This parameter values analysis provides suggestions how to choose suitable initial parameter values while using a mixture of Gaussian model in video surveillance application

Keywords: *background subtraction, mixture of Gaussian, parameters analysis, video surveillance.*

I. INTRODUCTION

Moving target detection is the key technology of the video processing in the field of computer vision, the background precise model and accurate segment the target from the video sequence is the basis of target tracking, recognition and behavior analysis. Background subtraction approach [1], optical flow [2] and inter-frame difference method [3] are mostly used in background modeling. However, the computational complexity of the optical flow method is very high; it is difficult to meet the requirements for real-time image processing [4], when the scene is complex, for example: illumination change, water wave and branches shaking, the accuracy and robustness are greatly reduced in background modeling with background subtraction approach and inter-frame difference method.

Scholars have proposed many background modeling and target detection method in complex scenes. The most used model is certainly the pixel-wise MOG one proposed by Stauffer and Grimson [5] due to a good compromise between robustness to the critical situations and constraints (CT, MR). There are many improvements of this MOG model like GMM [6], TLGMM [7], STGMM [8], SKMGM [9], TAPPMOG [10] and TAPPMOG [11]. All the developed strategies attempt to be more rigorous statistically or to introduce spatial and/or temporal constraints. The motivation of this paper concerns these initial parameters. The rest of this paper is organized as follows: in section 2 we firstly remind the original pixel-wise MOG model and the EM algorithm. Initial analysis of mixture of Gaussian model using EM algorithm is discussed in section 3. Experiments and conclusion are presented in section 4 and section 5, respectively.

II. RELATED WORK

A. Mixture of Gaussian model

In the context of a traffic surveillance system, Friedman and Russel[12] proposed to model each background pixel using a mixture of Gaussians corresponding to road, vehicle and shadows. This model is initialized using an EM algorithm [13]. Then, the Gaussians are manually labeled in a heuristic manner as followers: the darkest component is labeled as shadows; in the remaining two component, the one with the largest variance is labeled as vehicle and the other one as road. This remains fixed for all the process giving lack of adaptation to changes over time. For the foreground detection, each pixel is compared with each Gaussian and is classified according to it corresponding Gaussian. The maintenance is made using an incremental EM algorithm for real time consideration. Stauffer and Grimson[14][15]generalized this ideas by modeling the recent history of the color features of each pixel $\{X_1,...,X_t\}$ by a mixture of K Gaussians.

*Corresponding Author: Lihong Wang**
Longqiao College of Lanzhou University of Finance and Economics, Lanzhou, China, 730101.

Each pixel is characterized by its intensity in RGB color space. The pixel intensity is modeled by a mixture of K Gaussian distributions to model variations in the background like tree branch motion and similar small motion in outdoor scenes. The probability that a certain pixel has intensity x_t at time t is estimated as:

$$P(x_t)=\sum_{j=1}^{K}\omega_j(2\pi)^{-d/2}\left|\sum\nolimits_j\right|^{-1/2}\times\exp(-\frac{1}{2}(x_t-\mu_j)^T\sum\nolimits_j^{-1}(x_t-\mu_j)) \tag{1}$$

Where ω_j is the weight, μ_j is the mean and $\sum_j$ is the covariance for the jth distribution, $\sum=\sigma^2 I$ is used in this paper. Normally, K equals 3, 4 or 5 in practice. Every new pixel value x_t is checked against the existing K Gaussian distribution until a match is found. Based on the matching results, the background is updated as followings:

x_t matches component i if x_t is within 2.5 standard deviation of this distribution, in case of such a match, the parameters of the ith component are updated as followings:

$$\omega_{i,t}=(1-\alpha)\omega_{i,t-1}+\alpha \tag{2}$$

$$\mu_{i,t}=(1-\rho)\mu_{i,t-1}+\rho x_t \tag{3}$$

$$\sigma_{i,t}^2=(1-\rho)\sigma_{i,t-1}^2+\rho(x_t-\mu_{i,t})^T(x_t-\mu_{i,t}) \tag{4}$$

Where $\rho=\alpha P(x_t\mid\mu_{i,t-1},\sum_{i,t-1})$, α is the predefined learning parameter, where $1/\alpha$ controls the speed at which the model adapts to change, $\sigma_{i,t}^2$ is the variance of the ith Gaussian in the mixture at time t, μ_t is the mean of the pixel at time t, x_t is the recent pixel at time t.

The parameters for all the unmatched distributions remain unchanged, what means that

$$\mu_{i,t}=\mu_{i,t-1} \tag{5}$$

$$\sigma_{i,t}^2=\sigma_{i,t-1}^2 \tag{6}$$

But the corresponding weights $\omega_{i,t}$ need to be adjusted using the formulation:

$$\omega_{i,t}=(1-\alpha)\omega_{i,t-1} \tag{7}$$

If x_t matches none of the K distribution, then the least probable distribution is replaced by a distribution where the current value acts as its mean value, the variance is chosen to be "height" and the a-priori weight is "low".

The background estimation problem is solved by specifying the Gaussian distribution, which have the most supporting evidence and the least variance. Because the moving object has large variance than a background pixel, so the K distribution are ordered based on ω_j/σ_j^2 and the first B distribution are used as a model of the background of the scene where B is estimated as

$$B=\arg\min_b\left(\frac{\sum_{j=1}^{b}\omega_j}{\sum_{j=1}^{K}\omega_j}>T\right) \tag{8}$$

The threshold T is the fraction of the total weight given to the background model. Note that the denominator is supposed to be equal to 1 in case of proper normalization.

B. EM algorithm

The basic ideas of EM algorithm is: set Y is the set of observation values, complete data set is $X=(Y,Z)$, Z is the missing data, θ is the parameter, the a posteriori distribution of $p(\theta|Y)$ is too complex to some kind of statistical calculate. If the missing data Z is given, the adding posteriori distribution $p(\theta\mid y,z)$ will be known. The distribution $p(\theta\mid y,z)$ will be simple to do some statistical calculate. Conversely, it can check and improve the hypothesis of the missing data. Thus, a complicated maximum or sampling problems convert into a series of simple maximum or sampling problems.

Algorithm 1: EM-algorithm

Input: observation data Y , Latent variable Z , simultaneous distribution $p(Y,Z|\theta)$, conditional distribution $p(Z|Y,\theta)$

Output: model parameters

Step-1: choose initial values of the parameters $\theta^{(0)}$, start the iteration

Step-2: Set $\theta^{(i)}$ is the ith estimation of the parameter θ, calculate the Q function

$$\begin{aligned} Q(\theta,\theta^{(i)}) &= E_Z[\log p(Y,Z|\theta)|Y,\theta^{(i)}] \\ &= \sum_Z \log p(Y,Z|\theta)p(Z|Y,\theta^{(i)}) \end{aligned} \tag{9}$$

Where $p(Z|Y,\theta^{(i)})$ is conditional distribution under the given observation data Y and the current parameter estimation

Step-3: find θ that maximizes the function $Q(\theta,\theta^{(i)})$, Determine the parameter $\theta^{(i+1)}$ of the $(\mathrm{i}+1)$th iteration. That is：

$$\theta^{(i+1)} = \arg\max_{\theta} Q(\theta,\theta^{(i)})$$

Step-4: repeat step2 and step3 until convergence.

III. INITIAL PARAMETERS ANALYSIS

There are K components in the random variable spaceX, that is,$\mathrm{X} = \mathrm{X}_1 \cup X_2 \cup \ldots \cup X_k$, each component have probability density function $p(x;\theta_k)$, and the probability that each component selected is π_k, so, the probability density function of a random variable is:

$$p(x;\pi,\theta) = \sum_{k=1}^{K} \pi_k p(x;\theta_k) \tag{10}$$

Where $\pi = (\pi_1,\pi_2,...,\pi_k)$,

$\theta = (\theta_1,\theta_2,...,\theta_k)$ are parameters, it is the sum of weighted density or mixture of density. Obviously, when the function $p(x;\theta_k)$ selected Gaussian function, this model is named GOM, there are two parameters, mean value and variance for each Gaussian function, $\theta_k = (\mu_k,\sigma_k^2)$ the formulate can be rewrite:

$$P(y|\theta) = \sum_{k=1}^{K} \alpha_k \phi(y|\theta_k) \tag{11}$$

Where $\alpha_k \geq 0$, $\sum_{k=1}^{K} \alpha_k = 1$， $\phi(y|\theta_k)$ is the Gaussian function, $\theta_k = (\mu_k,\sigma_k^2)$, the kth component is formulated:

$$\phi(y|\theta_k) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{(y-\mu_k)^2}{2\sigma_k^2}\right) \tag{12}$$

Algorithm 2. Initialed for MOG using EM algorithm
Input：MOG， observation data: $y_1, y_2, ..., y_N$
Output：Parameters of MOG

Step-1: choose initial values of the parameters， start the iteration

Step-2: calculate the responsibility of the kth model to the observation data y_j under the current model parameters

$$\hat{\gamma}_{jk} = \frac{\alpha_k \phi(y_j|\theta_k)}{\sum_{k=1}^{K} \alpha_k \phi(y_j|\theta_k)} \quad j = 1,2,...,N;\ K = 1,2,...,K \tag{13}$$

Step-3: calculate the parameters for next iteration

$$\hat{\mu}_k = \frac{\sum_{j=1}^{N} \hat{\gamma}_{jk}\, y_j}{\sum_{j=1}^{N} \hat{\gamma}_{jk}},\ K = 1, 2, ..., K \tag{14}$$

$$\hat{\sigma}_k^2 = \frac{\sum_{j=1}^{N} \hat{\gamma}_{jk}(y_j - \mu_k)^2}{\sum_{j=1}^{N} \hat{\gamma}_{jk}},\ K = 1, 2, ..., K \tag{15}$$

$$\hat{\alpha}_k = \frac{\sum_{j=1}^{N} \hat{\gamma}_{jk}}{N} \tag{16}$$

Step-4: repeat step2 and step3 until convergence.

IV. EXPERIMENT

In many visual surveillance applications that work with outdoor scenes, the background of the scene contains many non-static objects such as tree branches and bushes whose movement depends on the wind in scene. This kind of background motion causes the pixel intensity values to vary significantly with time. For example, one pixel can be image of the sky at one frame, tree leaf at another frame, tree branch on third frame and some mixture subsequently, in another words, the pixel value have different color in each situation. Fig.1show changes over a short period of time about a gray level of vegetation pixel from an outdoor scene, it's located at (9, 90). The scene is shown at Fig.2.

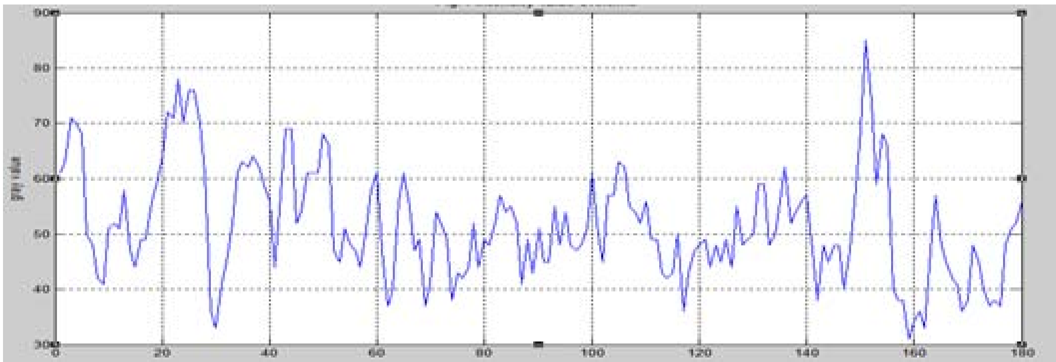

Figure-1: Intensity value overtime

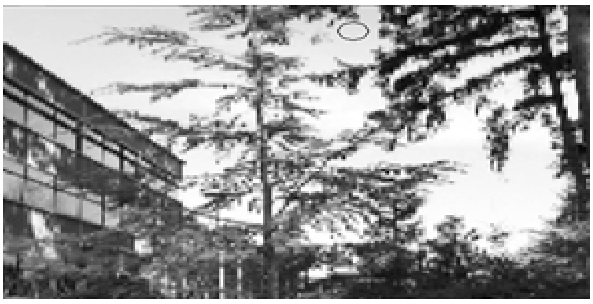

Figure-2: Outdoor scene with a circle showing the location of the sample pixel in Fig 1

There are two important factors in the moving objects detection in practices. Firstly, it is always desirable to achieve very high sensitivity in the detection of moving objects with lowest possible false alarm rates; another important factor is how fast the background model adapts to change. Fig.3.left shows the intensity histogram for this pixel, it is clear that intensity distribution is multi-modal, so using the mixture of Gaussian model to fitting is right. Fig.3. also shows 6 histogram of the same pixel obtained by dividing the original time interval into six equal length subintervals. Form this partial histogram we notice that the intensity distribution is changing dramatically over very short periods of time.

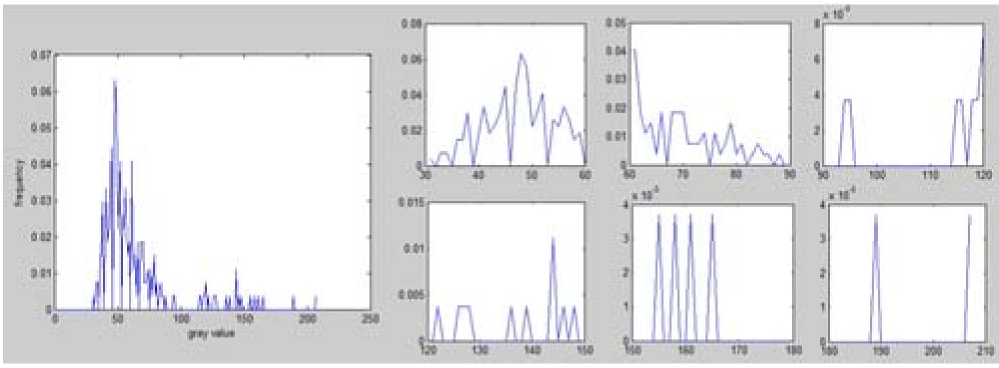

Figure-3: Iintensity histogram for this pixel

So, we can draw a conclusion that using more "short-term" distribution will allow us to obtain better detection sensitivity. In our experiment, we use firstly 1 second images of a given video to initialize for MOG using EM algorithm.

In our experiment, we choose $K=3$, a show background model with first six images using EM algorithm, b with first nine images, c with first twelve images, d with first eighteen images, e with first twenty-seven images, f show the background that build using MOG. The white edge denotes the region that waving tree can reach. From these images we can see that MOG background contain more noise, EM-MOG background less noise. The cloud of this video as background is eliminated on the EM-MOG background. As K increase, the conclusion is similar.

V. CONCLUSIONS

An important property of Gaussian distributions is that they still remain Gaussian distributions after any linear transformation. This property is one of the reasons that the Gaussian models are very commonly used for solving estimation problems. A Gaussian mixture is a probability density function, point distribution function consisting of a weighted sum of Gaussian densities. The value of each parameter influences the preference of the MOG. We presented a new method to initialize parameter for MOG using EM algorithm, and according to the experiment a conclusion that using more "short-term" distribution will allow us to obtain better detection sensitivity is draw. The experiment shows that this new methods can get better background. Of course, how to assign suitable values to parameters during an initialization period will also depend on specific applications. One needs to balance out all conditions according to different applications and environments.

REFERENCES

1. H.Y.S.K.C.J.S., "Robust background maintenance by estimating global intensity level changes for dynamic scenes", Intelligent Service Robtics, vol.2, no.3, pp.187-194, 2009.
2. H.D.Chalidabhongse T H.Kim K, "A perturbation method for evaluating background subtraction algorithm", In proceedings of the Joint IEEE International workshop on visual surveillance and performance Evaluation of Tracking and Surveillance, vol.10, 2003, pp.11-15.
3. L.A.J.F.H.P.R.S.,"Moving target classification and tracking from real-time video", in Proceedings of the 4thIEEEWorkshop on Applications of Computer Vision. Los Alamitos: IEEE Computer Society Press, 1998, pp.8-14.
4. Y.L.Z.Feng., "Research on state prediction based on multi-model fusion", Journal of Software, vol.7, no.2, pp.352-357, 2012.
5. Stauffer and Grimson, "Adaptive background mixture models for real-time tracking", in Computer Visio and Pattern Recognition.CO, vol.2, 1999, pp.246-250.
6. Han B, Lin X. "Update the GMMs via adaptive Kalman filtering". Proceedings of SPIE - The Int Society for Optical Engineering, 2005; 5960(3): 1506-1515.
7. Yang H, Tan Y, Tian J, Liu J. "Accurate dynamic scene model for moving object detection". IntConf on Image Processing (ICIP2007), 2007; 6: 157-160.
8. Zhang W, Fang X, Yang X, Wu Q. "Spatiotemporal Gaussian mixture model to detect moving objects in dynamic scenes". J Electronic Imaging 2007; 16(2).
9. Tang P, Gao L, Liu Z. "Salient moving object detection using stochastic approach filtering". Fourth IntConf on Image and Graphics (ICIG 2007), 2007; 530-535.
10. Harville M. "A framework for high-level feedback to adaptive, perpixel, mixture-of-Gaussian background models". Proc of the 7th European Conf on Computer Vision (ECCV 2002), Copenhagen, Denmark, May 2002, 543-560.
11. Cristani M, Murino V. "A spatial sampling mechanism for effective background subtraction". 2nd IntConf on Computer Vision Theory and Applications (VISAPP 2007), Barcelona, Spain, March 2007, 2: 403-410.

12. Friedman N, Russell S.” Image Segmentation in Video Sequences: A Probabilistic Approach”. Proceedings Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI 1997), 1997; 175-181.
13. Dempster A, Laird N, Rubin D. “Maximum likelihood from incomplete data via the EM algorithm”. J Royal Statistical Society, Series B (Methodological) 1977; 39(1): 1-38.
14. Stauffer and Grimson, “Learning patterns of activity using real-time tracking”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.22, no.8, pp.747-757, 2000.
15. Grimson and Stauffer, “Using adaptive tracking to classify and monitor activities in a site”, Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference, pp. 22 – 29.

Source of Support: Nil, Conflict of interest: None Declared