Improved Algorithm of Adaptive Gaussian Mixture Model for Moving Target Detection

Kaibi Zhang[a], Subo Wan[b] and Yangchuan Zhang [c]

School of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

[a]739569588@qq.com, [b]892821409@qq.com, [c]2293806388@qq.com

Abstract

The background modeling algorithm based on Gaussian mixture model (GMM) is a widely used method in moving objects detection with static cameras. Base on the situation that traditional Gaussian mixture model is very sensitive to sudden illumination variation and is slow for convergence speed, this paper proposed a method to detect the illumination variation and update the single learning rate, in order to build the adaptive updating background model. Using the algorithm of color histogram matching, the proposed method can adaptively adjust the learning rate by introducing the illumination variation factor and the counter for model parameters updating. Meanwhile, adaptive adjustment of the number of GMM components reduced the computational cost and improved the real-time performance.Experimental results show that this proposed approach can adapt scene changes efficiently, and has better accuracy and robustness than traditional Gaussian mixture model.

Keywords

Gaussian Mixture Model; Moving Target Detection; Adaptive; Illumination Variation; Background Modeling.

1. Introduction

Moving target detection belongs to the low-level stage of processing of computer vision, and is as well the basic step to further achieve target tracking, feature extraction, behavior analysis and understanding, etc[1]. Therefore ,effective testing results are vital to subsequent high-level stage of processing. Detection methods of moving targets are mainly divided into point inspection, picture segmentation, inter-frame difference method, background subtraction, clustering methodology and motion vector field method[2]. Among them, background subtraction method is widely put into use especially in the scenes where the camera is fixed. It works by comparing every single frame with real-time updated background model to segment out background areas and foreground areas.

Gaussian mixture model (GMM) is a kind of background modeling based on pixel-level .Literature 3 applies it to background modeling and movement segmentation of video surveillance, adopting k (usually 3 to 5) Gaussian distribution to describe the transformation law of background pixel. Literature 4 divide background study and update into two parts based on the original GMM. The learning rate at the initial stage of the background is 1/N, later the common way of iteration will be employed to update. Literature 5 put forward an adaptive method to provide different learning rate for every Gaussian distribution, making the speed of convergence greatly promoted. Literature 6 analyze the saturation pixel phenomenon occurring when mean and variance of modeling update and result in severe deterioration of convergence of variance. Besides, literature 6 raises two different learning factor and separately update the mean and variance .In some recent researches,literature 7 comes up with a renewing plan of window weight number to decrease the systematic performance period and strengthen the instantaneity. Through many scholars' research, The robustness and stability of GMM has both been promoted to some extent, but the adaptive capacity towards changes of background is still not satisfying, particularly in some certain phenomenon like abrupt variation of illumination. Literature 8 proposes judging whether abrupt variation of illumination exists by counting foreground pixel number and whole image pixels in total to get the scale value, and on this point the present changing image frame can not manage to get effective update in time.

This passage ,based on traditional GMM, through introducing illumination change factor, solves the problem of decline in dynamic adaptability caused by illumination change, and by updating the self-adaption of learning rate in the mean while adaptively select the number of Gaussian components to cut down the operand, the robustness and stability have been enhanced at a certain degree.

2. Background modeling based on GMM

For simple scene, the application of single Gaussian model can express the color vector change of each pixel. However, for complicated scene, the application of single model distribution cannot fit the data effectively. GMM is the weighted sum of finite Gaussian functions,it can describe the multimodal state of pixel and model the complicated dynamic background.

Traditional GMM supposes that each pixel in the image is independent mutually and the change in the time domain is simulated by using K multi-dimensional Gaussian distribution,the sampling value of a pixel P(x,y) is $\{X_1,X_2,\ldots,X_t\}$, then the probability of current pixel value xt observed at the moment of t is:

$$P(X_t)=\sum_{i=1}^{k}\omega_{i,t}N(X_t,u_{i,t},\Sigma_{i,t}) \tag{1}$$

Hereinto, K is the number of modular component; ωi,t, μi,t and Σi,t are the weight, mean and covariance matrix of Gaussian distribution i in the model at the time t. the probability density function of Gaussian distribution i is:

$$N(X,u,\Sigma)=\frac{1}{(2\pi)^{n/2}\left|\Sigma\right|^{1/2}}\times\exp(-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)) \tag{2}$$

When K Gaussian distributions are ranked the order according to ω/σ, then the top B distributions are taken as the background model:

$$B=\arg\min_b(\sum_{k=1}^{b}\omega_k>T) \tag{3}$$

If the difference between the current pixel value and the background model is within certain range, it can be judged as background, namely

$$\left|X_t-\mu_{i,t-1}\right|\le\beta\times\sigma_{i,t-1} \tag{4}$$

Hereinto, β is 2. 5~3, if meeting Formula (4), then it shall be updated as per the following formula:

$$\omega_{i,t+1}=(1-\alpha)\omega_{i,t}+\alpha\times M_t \tag{5}$$

$$\mu_{i,t+1}=(1-\rho)\mu_{i,t}+\rho\times X_{t+1} \tag{6}$$

$$\sigma^2_{i,t+1}=(1-\rho)\sigma^2_{i,t}+\rho\times(X_{t+1}-\mu_{i,t+1})\times(X_{t+1}-\mu_{i,t+1})^T \tag{7}$$

Hereinto, learning rate α is a constant, ρ=α×N(Xt+1, μi, ∑i). None of the current background model can meet the distribution in Formula (4), then a smaller weight and mean as the current pixel value and lager variation distribution are initialized to replace the distribution with smallest weight in the model.

The updating speed of GMM mainly depends on learning rate α. If α is smaller, the initialization and updating speed of background model is slow and it takes longer time to adapt to the environmental change; otherwise, if α is bigger, the initialization and updating speed of background model is rapid and it has powerful capability to adapt at the environmental change, but it may produce noise[9].

3. The method of self-adaptive Gaussian Mixture Model

3.1 Illumination Change Detection

GMM supposes each pixel is independent. Thus, when the outside world illumination variation causes scene change, it is probable to cause big area of false target and thus produce erroneous judgment [10]. Through observing illumination variation in the practical video application, the paper classifies illumination variation as two types, namely sudden variation and slow variation. Through analyzing such two kinds of variations, the paper introduces illumination variation factor θt so as to eliminate the impact of illumination variation on the detection of moving object.

$$\theta_t = 1 - \frac{E_{t-1}}{E_t} \tag{8}$$

$$E_t = \sqrt{E_R^2 + E_G^2 + E_B^2} / \sqrt{3} \tag{9}$$

Hereinto, Et expresses the information entropy of the current frame. ER,EG and EB represents the information entropy of RGB components of the current frame. The color in the image is closely related to the object. Different luminance of pixel distribution can reflect the ambient illumination variation. Hence, the paper adopts color histogram method to extract the color feature. When the characteristic values of two tested image samples are different, but the difference is smaller than one threshold value, it indicates they have high similarity and the statistic distributions of these two images are similar. By utilizing the theory, the paper presents illumination variation detection method, namely histogram matching algorithm, to distinguish the slow variation and sudden variation of illumination. The mathematic expression is:

$$D_{(t,t-1)} = 1 - \sum_{m=1}^{M} \min(H_t(m), H_{t-1}(m)) \tag{10}$$

Hereinto: H(t) means the histogram of the image at the time t. through normalizing Formula (10), it can be got:

$$D(t,t-1) = 1 - \frac{\sum_{m=1}^{M} \min(H_t(m), H_{t-1}(m))}{\sum_{m=1}^{M} \min(H_t(m))} \tag{11}$$

The paper distinguishes sudden variation from slow variation based on D(t,t-1) and the judgment method is:

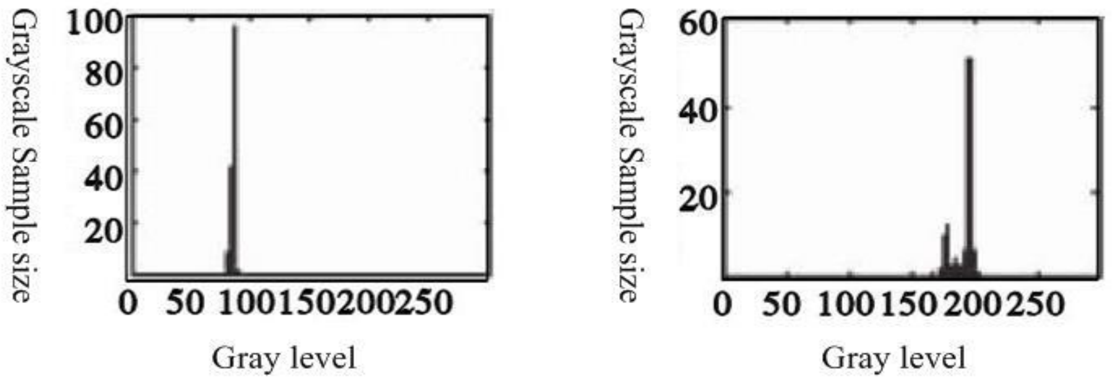
$$\begin{cases} \text{slow variation}, & \text{if } D(t,t-1) \le (T) \\ \text{sudden variation, otherwise} & \end{cases} \tag{12}$$

Hereinto, T2 is similarity matching threshold value. After distinguishing illumination variation, learning rate α shall be made self-adaptive updating as per the following Formula (13).

$$\begin{cases} \alpha = \alpha + \theta_t & \text{if slow variation} \\ \alpha = 2\alpha & \text{if sudden variation} \end{cases} \tag{13}$$

3.2 Self-adaptive Gaussian Mixture Model

K is a fixed constant in traditional GMM and establishes K distributions for each pixel. However, in realistic background, the modal distribution number of each pixel is not equal. The stable region may be single mode and it can model with a Gaussian component; the busy region may need multiple Gaussian components to model[11]. As shown in Diagram 1, it expresses grey level statistic histogram of two pixels of one video in time domain. It can be seen the aggregation characteristic appeared by sample value may be single-peak or multiple-peak value.



(a)pixel(146,60) (b)pixel(136,60)

Fig. 1 Statistic histogram of the pixel gray-levels of video sequences

In order to reduce redundant Gaussian component so as to reduce calculated quantity, literature [12] presents an online iterative algorithm. Through importing modal Dirichlet prior probability and according to the result of maximum posterior probability or the minimum message length standard, giving up or adding the number of Gaussian components automatically at the time of parameter estimation, the algorithm makes K value adapt to multi peak of each pixel dynamically. The iterative approach of modal weight is

$$\omega_{i,t+1} = (1-\alpha)\omega_{i,t} + \alpha \times M_t - \alpha c_T \tag{14}$$

Hereinto, cT is a constant, reflecting the dimension of model parameter. Self-adaptive selection of proper number of mode is the target of model design. Not only can it enhance the modal stability, but also save operation time effectively, besides, it improve the real-time performance of detection system greatly.

Based on the algorithm mentioned in literature [12], through self-adaptive adjustment of recursive learning rate, the paper presents a self-adaptive Gaussian Mixture Model. Learning rate α in the traditional Gaussian Mixture Model is a constant, but single learning rate cannot adapt to the scene change dynamically. The paper combines illumination variation factor θt to have a real-time adjustment of learning rate α(Formula (13)). Meanwhile, it updates the learning rate ρ in Formula (6) and (7) automatically. For a new sample, the updating method of learning rate ρ is:

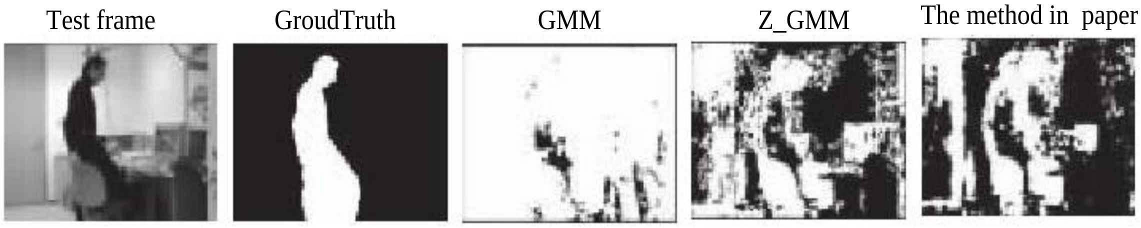
$$\rho_{i,t} = \frac{\alpha}{\omega_{i,t}} \times \frac{1+c_i}{c_i} \tag{15}$$

Hereinto, ci is a counter to record the updating of corresponding component parameter. When a new component is created, the initial value of ci is 1, and increased gradually by 1 when the corresponding component is updated; when a component is abandoned, the initial value of ci is set as 1. Weight is updated according to Formula (14), mean and variance should be updated according to Formula (6) and (7).

4. Experimental result and analysis

In order to verify the validity of algorithm proposed in this paper, the paper conducts contrast experiment of several video sequences in different scenes. The experimental environment is 2. 5 GHz PC. Firstly, the paper conducts contrast experiment specific to the impact of illumination variation on the detection system. Secondly, the paper conducts comparison test of the running time of algorithm.

The paper selects image sequence that can represent sudden variation and slow variation of illumination [10,11] to verify the self-adaption of the algorithm proposed to the illumination variation, as shown in Diagram 2. Hereinto, GroudTruth is the truth value of the corresponding moving object binary image of test frame in various video databases. In Diagram 2, sequence of image (a) is LightSwitch; sequence of image (b) is Time of day; sequence of image (c) is high-Way.



(a)LightSwitch video files

(b)Time of day video files

(c)Highway video files

Fig. 2 Experimental comparison of illumination variation

In LightSwitch, the tester conducts turn on/off the light experiment indoors to simulate the sudden variation of illumination; in Time of day, tester simulates the weather gradual variation process indoors; in highway, the reflection of flickering leaves, car and road toward the illumination causes sudden variation of local illumination. From Diagram 2, it can be seen that GMM has low adaptability to illumination variation and produce large-area false target. The adaptability of the method proposed in this paper and the algorithm presented in literature 12 (called Z-GMM) is enhanced to some extent. With the guidance of self-adaptive learning rate α, the method proposed in this paper has a superior foreground division effect than Z_ GMM algorithm at the time of sudden illumination change; the foreground division effect of the method proposed in this paper is similar to Z_GMM at the time of slow illumination variation.

Z_ GMM algorithm can select to describe the number of Gaussian component of each pixel automatically, so the processing time is similar. As shown in Diagram 1, through the image sequences of three different scenes, the paper compares the average processing time of each frame by using the method presented in the paper and GMM. Hereinto, the number of fixed Gaussian component of GMM is 4, the maximum Gaussian component of the method presented in this paper is also 4. According to Diagram 1, it can be seen that in the crowd and complicated scene, the average processing time of GMM algorithm is slightly higher than that of method presented in this paper; in the simple scene, the average processing time of the method proposed in this paper is enhanced to some extent.

Table 1.The comparison of per frame's average processing time

Algorithm	Canoe	Highway	Campus
GMM	18.86	18.67	18.06
The method in paper	15.09	8.56	5.38

5. Conclusion

The method presented in this paper improves GMM algorithm from two aspects: a) to mitigate the impact of illumination variation on detection system, the paper, through color histogram matching algorithm, distinguishes sudden variation from slow variation of illumination and update learning rate α self-adaptively through importing illumination variation factor; b) through importing model parameter updating counter ci, it updates learning rate ρ self-adaptively at the time of updating the model so as to enhance the convergent speed. By combining literature [12], the paper adjusts the number of nt of GMM automatically and enhances the real-time performance of detection algofixed Gaussian componerithm. The robustness and stability of detection effect have been verified in the experiment.

References

[1] Lin Lizhong, Liu Zhiguo, Zhang Yubin. Research on Detection and Tracking of Moving Target in Intelligent Video Surveillance [C].In Proceedings of the IEEE International Conference on Computer Science and Electronics Engineering. 2012.477-481.

[2] Zhao Wenzhe, Qin Shiyin. Comparative Study on Detection Methods for Video Motion Targets[J].Science & Technology Review,2009,27(10):64-70.

[3] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking[C] //Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. MA. USA. 1999. 2(6). 23-25.

[4] Bouwmans T, Baf F E, Vachon B. Background modeling using mixture of Gaussians for foreground detection—a survey [J]. Recent Patents on Computer Science,2008, 1(3): 219-237.

[5] Lee D S. Effective Gaussian mixture learning for video background subtraction[C]//Proc of IEEE Tran on Pattern Analysis and Machine Intelligence. CA.USA. 2005. 27(5). 827-832.

[6] Greggio N, Bernardino A, Laschi C.Self-adaptive Gaussian Mixture Models for Real-time video Segmentation and Background Subtraction[C]//Proc. Of the 10th International Conference on Intelligent Systems Design and Application.Cario,Egypt:[s.n.],2010:325-330.

[7] Ruben H E, Michael P, Thomas S. Splitting Gaussians in Mixture Models[C]//Proc of IEEE International Conference on Advanced Video and Signal Based Surveillance.Berlin. Germany. 2012. 300-305 .

[8] Mingzhi,Ma Zhiqiang,Shan Yong.Adaptive background update based on Gaussian mixture model under complex condition[J]. Journal of Computer Applications,2009,27(10):64-70.

[9] Lin Lizhong, Liu Zhiguo,Zhang Yubin. Research on Detection and Tracking of Moving Target in Intelligent Video Surveillance [C]. In Proceedings of the IEEE International Conference on Computer Science and Electronics Engineering. 2012.477-481.

[10] Zang Q, Klette R. Evaluation of an adaptive composite gaussian model in video surveillance. CITR Technical Report 114.Auckland University. August 2002.

[11] Wang H, Suter D. A Re-Evaluation of Mixture -of-Gaussian background modeling.30th IEEE Int Conf on Acoustics, Speech, and Signal Processing (ICASSP 2005).Pennsylvania. USA. March 2005. 1017-1020.

[12] Zivkovic Z, Improved adaptive Gaussian mixture model for background subtraction[C]//Proc of International Conference on Pattern Recognition. Amsterdam Univ. Netherlands. 2004. 28-31.