Optik 126 (2015) 2256–2259

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Research on moving object detection based on improved mixture Gaussian model

Xiaorong Chen*, Chuanli Xi, Jianghui Cao

Shanghai Key Laboratory of Modern Optical System, School of Optical–electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

ARTICLE INFO

Article history:
Received 17 April 2014
Accepted 25 May 2015

Keywords:
Gaussian mixture model
Object detection
Parameter estimation

ABSTRACT

For the deficiency of the mixture Gaussian model (GMM), an improved GMM algorithm is proposed, which can be applied to moving object detection. Combined with codebook detection algorithm, the GMM model is initialized by video images pixels statistical mean and variance, and updated model using parameters confidence interval. The experiment results indicate that the proposed update model can detect moving objects in complex background effectively and has good robustness.

1. Introduction

Moving target detection can detect the moving object in video sequences, and then positioning and segment moving targets in order to extract the target from video image. As the first step in video analysis, object detection plays an important role in system. There are mainly two modes of moving object detection: static camera, background is also static. It is easy to detect the moving object with background subtraction method, it usually used in video surveillance tasks; moving camera, in this case, the background move. It usually detects and follows moving object and person on a moving platform [1]. This paper mainly discusses the moving object detection under static background.

At present, motion detection includes three methods: light flow method [2], frame differential method [3], the background subtraction method [4]. Light flow method is complicated to calculation. The method is not often used for its bad real-time performance and requires specially configured hardware support. Inter-frame differential method is not suitable for detecting a slow moving target and it is not sensitive to detect gray uniformity moving object. The background subtraction method is a common method for moving target detection. It uses the current frame minus the reference background image to detect moving objects. The mixture of Gaussians method is widely used for the background modeling.

This paper made two improvements to mixed Gaussian background model. First, codebook algorithm is used to initialize model parameters in order to improve the accuracy of the model. Second, interval estimation are used to match pixels, it solved the problem of pixel perturbation and misjudgment brought by foreground pixels slow movement.

2. Mixture Gaussian background model

Mixture Gaussian background model is also called expectation maximization algorithm [5]. Each pixel in the scene create K Gaussian model, using the model counts the change of the pixels in the scene. Larger K means that the model is stronger to deal with fluctuations. Of course, the complexity of the algorithm increases.

The probability of the current pixel value can be written as:

$$p(X_T) = \sum_{k=1}^{K} w_{k,t}\eta(X_t, \mu_{k,t}, \sum\nolimits_{k,t}) \quad (1)$$

K is the number of Gaussian model, $w_{k,t}$ is the weight of Kth Gaussian distribution at the current pixel, $\sum_{i=1}^{K} w_{i,t} = 1$. η is the kth Gaussian probability density function at the current pixel:

$$\eta(X_t, \mu_{k,t}, \sum\nolimits_{k,t}) = \frac{1}{(2\pi)^{n/2}\left|\sum_{k,t}\right|^{1/2}} e^{-\frac{1}{2}(X_t-\mu_{k,t})^T \sum_{k,t}^{-1}(X_t-\mu_{k,t})} \quad (2)$$

$\mu_{k,t}$ is the mean of the kth Gaussian at the pixel. $\sum_{k,t}$ is the covariance matrix of the kth Gaussian. In order to simplify the calculation, we suppose that three channels of RGB are independent and covariance equal, $\sum_{k,t} = \sigma_{k,t}^2 I$.

K Gaussian models of each pixel are sorted by $w_{k,t}/\sigma_{k,t}$. $w_{k,t}/\sigma_{k,t}$ represents the possibility of the Gaussian model is background. In

* Corresponding author. Tel.: +86 13816365571.
E-mail address: piaomiaoxi@163.com (X. Chen).

http://dx.doi.org/10.1016/j.ijleo.2015.05.122
0030-4026/

order to eliminate the interference of isolated pixel, this paper set a background condition B:

$$B = \arg\min_{b}\left(\sum_{j=1}^{b} w_{j,t} > T\right) \tag{3}$$

T is the threshold, the former B models that can satisfy Eq. (3) are background model.

At the moment t, the acquired pixel value match with K Gaussian distributions, if the pixel value X_t and a model satisfy Eq. (4), then the pixel matches the model.

$$\left|X_t - \mu_{i,t-1}\right| < D\sigma_{i,t-1} \tag{4}$$

$\mu_{i,t-1}$ and $\sigma_{i,t-1}$ are mean and variance of the ith Gaussian distribution at the moment $t-1$; D is a user-defined parameter, generally takes 2.5.

If the Gaussian model backgrounds satisfy Eq. (4), model parameters are updated as follows:

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

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

$$\sigma_{i,t}^2 = (1-\rho)\sigma_{i,t-1}^2 + \rho(I_t - \mu_{i,t})^2 \tag{7}$$

α is learning rate, it ranges from 0 to 1. α determines the speed of the background updating. When α takes a larger value, the background update faster. If pixel value X_t and background model does not match, then following formula updates $w_{i,t}$:

$$w_{i,t} = (1-\alpha)w_{i,t-1} \tag{8}$$

3. Improved Gaussian mixture background model

Although Gaussian mixture model has many advantages in complex scene modeling applications, but this algorithm still exist some problems.

The first is model parameter initialization problem. As mentioned earlier, this model uses the mean and variance of the image pixels in a period of time as the initial model parameters. If the moving object appears in initial stage of the frames, this method will take their pixels into account. Although in the following update, as the number of frames increase the mean and variance will be close to the "true" value. But, this method will cause the background model is not accurate.

Secondly, Eq. (4) are used to distinguish foreground and background pixels. In the initial stage σ is a large value. If the disturbance of background pixels is small, model will not detect interference pixels as foreground pixels. But, for a slow-moving object, if object pixel and background pixel values are similar, disturbance of object pixels is small. The model will make it as the background. In addition, according to Eq. (7), with the background model update. Variance will converge to a minimum value. Model will take a part of the background as foreground.

3.1. Parameters initialization based on codebook model

In order to avoid background pixel and the foreground pixel mixture counting, we propose a pixel value classification model based on codebook model [6,7]. Codebook model classify all pixels based on a color distortion and brightness distortion. A code word represents a class to determine whether the pixel matching codebook model.

In video sequence, pixels of the same location at different times can be expressed as $X = \{X_1, X_2, ..., X_N\}$. Every sequence creates a codebook model: $C = \{c_1, c_2, \ldots, c_l\}$. Each codebook containing L codewords c_i. Codeword consists of a RGB vector $V_i = (\bar{R}_i, \bar{G}_i, \bar{B}_i)$ and a six-tuple $aux_i = (I_l, I_h, f_i, \lambda_i, p_i, q_i)$. I_l and I_h are minimum and maximum brightness variables; f_i can calculate codeword frequency; λ_i means the maximum negative run-length (MNRL) of the codeword. p_i and q_i means the time of the first and last activation of the codeword.

Detailed codebook model algorithm is as follows:

(1) Set codebook for each pixel are empty, $L = 0$.
(2) For the pixel sequence, $X = \{X_1, X_2, \ldots, X_N\}$, $X_t = (R_t, G_t, B_t)$, $t = 1, \ldots, N$.
(3) If codebook is empty, then create a new codeword:

$$L = L + 1, \quad I = \sqrt{R_t^2 + G_t^2 + B_t^2} \tag{9}$$

$$v_L = (R_t, G_t, B_t), \quad aux_L = (I, I, 1, t-1, t, t) \tag{10}$$

(4) If codebook is not empty, then the constraints:

$$colorim(x_t, v) \geq \varepsilon_1 \tag{11}$$

$$brightness(I_l < I < I_h) = ture \tag{12}$$

If a codeword c_m satisfy the above conditions, then update the matched codeword:

$$v_m = \left(\left(f_m\bar{R_m} + R_t\right)/(f_m+1), \left(f_m\bar{G_m} + G_t\right)/(f_m+1), \left(f_m\bar{B_m} + B_t\right)/(f_m+1)\right) \tag{13}$$

$$aux_m = (\min\{I, I_l\}, \max\{I, I_h\}, f_m + 1, \max\{\lambda_m, t - q_m\}, p_m, t) \tag{14}$$

(5) After training, we statistics the maximum negative run-length of the codeword:

$$\lambda_i = \max(\lambda_i, (N - q_i + p_i - 1))$$

(6) Eliminate redundant code using λ can get best initial codebook:

$$M = \{c_k | c_k \in C, \lambda_k \leq T_M\} \tag{15}$$

T_M is a training threshold, usually take half of frame number, means that all backgrounds codeword appear at least in the $N/2$ frame images. After the codebook model, this algorithm statistics background pixels mean and variance value. Since the background pixel has ruled out interference of moving target and noise, so it can represent the true background.

This paper counts distribution of pixel values of a point in a standard test video. The distribution of the pixel value is multimodal. As shown in Fig. 1, we use the (180, 120) pixel as the sample pixel to gather statistics the values in the number of frame images. Fig. 2 shows the distraction of codebook model method and the general calculation method.

Fig. 2 shows R channel pixel value curve about sample pixel (180,120) in 100 frame images. In the figure pixel value curve has obvious fluctuation and the sharp troughs are caused by vehicles. Mean value 1 is calculated by this paper's method and mean value 2 is calculated by all pixels. Mean value 1 could represent more precisely the background pixels distribution. Because of exclude the interference of foreground pixels, pixel variance becomes very small. This paper's method variance is calculated 5.2 and mean value 2 method variance is 1921.16.

3.2. Interval estimation of parameters

After the initialization of background, pixels in the current image and the corresponding pixel in the model will be matched. If the