Improved Gaussian Mixture Model in Video Motion Detection

Xie Yong

Modern Education Technology Center, Xi’an International University, 408 Zhangba North Road, Xi’an 710077, China,
Email: xiey_7948@163.com

Abstract—As the classical Gaussian mixture model has some problems of not considering it self’s matching degree of Gaussian density functions, model updating and the background in real video motion detection, made improvements on the three aspects. Optimized Gaussian mixture model’s overall architecture and proposed an improved algorithm according to the analysis of the definition and disadvantages of classical Gaussian mixture model. Finally, through detailed experiment, the result showed: the improved Gaussian mixture model is faster in model convergence rate in video motion detection, can quickly adapt to the changes of background and greatly decreases the fall-out ratio.

Index Terms—Matching Modal; Video Motion; Framework Optimization; Gaussian Mixture

I. INTRODUCTION

Efficiently and accurately abstracts the no stationary objects is a key to video processing in intelligent video surveillance system [1]. The motion detection algorithm is multifarious, the traditional algorithms are optical flow field method, frame difference method and background subtraction. The first two has some obviously disadvantages. The optical flow field method needs repeatedly iterative operation, the calculations are completing and time-consuming, it’s hard to realize real-time detection without specific hardware support; the algorithm has poor ability in anti-jamming like noise so it’s rarely used in practical application [2-5]. The position of motion objects detected by the frame difference method is inaccuracy, its enclosing rectangle extrudes in motion direction and affected by the object’s velocity and time interval between adjacent images, usually cannot abstract all the relatively feature pixel point, inside of the motion entity easily generates cavitations, when the moving object in the scene has no obviously movement, the pixel value in the moving object’s area returns to zero, the moving object cannot be detected [6-9]. Now the generally used method is background subtraction, different the inputted video sequence’s current frame and established background model and abstracts the moving object. Comparing with other motion detection algorithm, it has the advantages like small data calculation, easy to realize and can abstracts complete moving object, as an common method of motion detection, it’s wildly used. Now the motion detection’s major research direction is to establish appropriate background model to adopt to different change of scene and application environment [10-13].

The Gaussian mixture model is a wildly used background model and it’s a improvement of single Gaussian background model. The single Gaussian background model uses individual Gaussian distribution to express background, can process the simple scene with small change and secular change [14]. But its not applicable to those scenes with big change or sudden change of background, or the background pixel value is multimodal distributed (like small and repeatedly motion). Because then the background changes fast, its distribution is not a transition from a relatively steady unimodal distribution to another [15]. Considered the distribution of background pixel value is multimodal, on the basis of the method of unimodal, can uses unimodal set to describe the change of pixel value in complex scene, the Gaussian mixture model exactly uses multiple single Gaussian function to describe multi-modal scene background. It better describe the background model distribution in complex scene, has good application effect on processing sudden change of background and background disturbance. The classical Gaussian mixture proposed by STAUFFER C can solve the the problems like sudden change of background and background disturbance. KAEWTRAKULPONG P realized real-time application on the basis of theory, and had a better result in sudden change of background and illumination variation. But in real application, because of the deficiencies of motion detection algorithm, there are some problems: firstly, matching Gaussian modal selection only considers the modal’s weight without considering the modal it self’s matching degree of Gaussian density function. Secondly, it uses totally different update mechanism in the background model’s initialization phrase and steady phrase, when the background change is very big, the background model’s re-establish time is too long which affects the system’s detection. Thirdly, the final background display only shows the modal with the biggest weight without considering the effect of other modals.

The Gaussian mixture background’s modeling process permits the existence of video moving object, which is appropriate for the detection of small and fast moving object outdoor and with weather change. But the algorithm has its disadvantages, the detailed analysis is as following: firstly, when matching Gaussian modal...
selection, this algorithm only considers the weight of modal without considering its self’s matching degree of Gaussian density function, the selected modal may not be the best matching. Secondly, the background model's initialization phrase and the steady phrase uses two different updating mechanisms in the Gaussian mixture proposed by STAUFFER C, it has a better effect when the scene change is not big and the background disturbance is small.

The paper based on classical Gaussian mixture algorithm and aimed at the disadvantages of Gaussian mixture model algorithm, respectively made a series of improvement about matching modal selection method, model’s updating algorithm and background display; this paper mainly made explorative and innovative works as following aspects: Aimed at the problems of matching degree, model updating and background display in motion detection of classical Gaussian mixture model, when the classical Gaussian mixture model makes modal matching, compares to each modal by turns, once meets \[ |X_t-| \leq j \], then regards the sample matches this modal, the selection patterns has obviously problems, when the modal updating, the convergence rate of system is relatively slow, when the classical Gaussian mixture model algorithm makes background display, directly takes the modal component value with the biggest weight as background display, without considering other (B-1) component value is not accurate enough as background. So there made improvements of Gaussian mixture model in these three aspects. The specific method of improvement is: in modal matching, considers the effect of weight, Gaussian modal itself mean value and variance towards distribution, makes twice compare of the new samples and the Gaussian modal of the model; takes unified expression in model updating; reunifies the weight of the background’s B modal in background display, distributes the background’s each model component value by the weight after uniformization.

In order to further prove the accuracy and efficiency of improved Gaussian mixture model, integrated considered every modal’s distribution of background model, made the displayed background closer to the background distribution of real scene. Through detailed experiment, the result showed, compared with previous method, the model in this text is faster in convergence rate, can better adopt to the exchange of background and greatly decrease the fall-out ratio.

II. ANALYSIS ON GAUSSIAN MIXTURE MODEL

The background modeling’s fundamental of Gaussian mixture model is: according to the different modal matching’s frequency of sample set (the order of pixel point’s gray value formed in time shaft) and background model, constantly updates the Gaussian distribution’s parameter of every modal in the model, ie, trains the parameter like weight, mean value and covariance of every Gaussian distribution, makes background pixel value distributed convergence at one or several Gaussian distribution, realizes the cluster of background pixel value, and realizes the modeling of background.

A. Model Definition

Based on the background modeling method of Gaussian mixture model, its modeling object is every pixel point’s pixel value. In the model, every pixel point’s gray value in the image is regards as a statistics and stochastic process, the pixel value of pixel point can be regarded as a vector sequence, arbitrarily pixel point (x,y) its history pixel value can be expressed as:

\[
\{x_1, ..., x_r \} = \{I_i(x,y) : 1 \leq i \leq r \}
\]

There \(I_i(x,y)\) refers to the gray value of pixel in time i. Gaussian mixture model uses K Gaussian distribution to stand for those history value, so the proportion of pixel \(x_i\) as current value is:

\[
p(x_i) = \sum_{j=1}^{K} w_{ij} \eta(u_i, r_i, \Sigma_{ij})
\]

Among them, \(w_{ij}\) is the i weight of Gaussian distribution in time t, which reflects the appearance proportion of Gaussian distribution; \(\eta(u_i, r_i, \Sigma_{ij})\) is the Gaussian proportion density function when the i time’s mean value \(u_i(t)\) and the covariance is \(\Sigma_{ij}\);

\(p(x_i)\) refers to the proportion of pixel value as X; K is the number of distribution.

III. PROPOSED SCHEME

Figure 1 is a flow diagram of improved Gaussian mixture model, mainly improves the Gaussian mixture algorithm in three aspects: the selection of matching Gaussian modal, the updating of the model and the display of the background.

Figure 1. The flow diagram of improved Gaussian mixture algorithm

The process module of blue checks background is the flow diagram of classical Gaussian mixture algorithm, the orange process module is the detailed improvement of Gaussian mixture. The process module of blue checks background are respectively matching the select and judge process of the algorithm’s initialization, modal updating, modal matching and background modal. The
orange process module is the major improve points of this improved algorithm which includes: the matching modal’s selection, updating means and background display patterns.

Aimed at the disadvantages of the Gaussian mixture algorithm, the improved algorithm mainly makes a series of improvements on the matching modal selection, updating means and background display patterns. Firstly, on the matching modal selection module, combines the modal weight and it self’s matching condition, makes the judge of Gaussian modal matching; secondly, on updating means, unifies the updating means of initialization phrase and detection progress, makes the model a smooth transition from initialization phrase to steady phrase; finally, on the background display patterns, integrated considers the background modal’s every modal distribution and makes the background’s display can clearly reflect the background model’s detail distribution.

A. The Selection of Matching Modal

Each modal of the Gaussian mixture model is ordered as \( w_j/a_j \) from the smallest to the biggest, when the modal matching is processing, comes a new sample \( X_t \), compares with each modal, if only meets \( |X_t - u_j|/\sigma_j \leq 2.6 \) then regards the sample is matching to the modal, doesn’t consider other not compared modals. This selection pattern has some obviously problems: when the modal matching is processing, the matching modal may not be the best modal by the current pattern, because the algorithm will out of the loop when finds one modal, if other rest modal is more suitable than the selected one for the weight, this selection pattern is not the best evidently. For example: three Gaussian modal’s distribution and weight of certain pixel is showed as figure 2. The Gaussian modal’s mean value and variance is different, the weight decreases successively.

\[ \text{Figure 2. The Gaussian modal and weight distribution figure} \]

The pixel distribution at certain time is \( x_t \), if only considers weight \( w_j \), and the weight’s biggest modal 1 as matching model, showed as the first line in figure 3; if only considers without considering current pixel distribution, then the biggest modal 2 as matching modal, showed as the second line in figure 2; if both considers the weight and itself’s matching degree, according to the value of \( w_j/(|x_t - u_j|/a_j) \), the biggest modal 3 as matching modal, showed as the third line in figure 3. This shows different matching modal selection mechanism can led to different modal matching result, considering multiple factors affected by modal matching, to select the best modal matching is the key to the algorithm improvement.

B. Updating Means

The updating of Gaussian mixture model’s parameter can be expressed by a uniformed formula (3),

\[ \theta(t) = (1 - \eta(t))\theta(t - 1) + \eta(t)\nabla(y(x(t);\theta(t - 1))) \]  

Among them, \( \theta(t) \) generally refers to parameter. At time \( t \), rate \( \eta(t) \) and control factor \( \nabla(y(x(t);\theta(t - 1))) \) determines the updating. \( K(t) \) depends on updating rate.
Method 1: $\eta(t) = 1/t$, the parameter estimation is related to the number of observed samples, with the increase of sample’s number, the parameter quickly reaches to expectation value, which is suitable for the updating in system’s initialization phrase;

Method 2: when $\eta(t) = a$, the system’s parameter is related to the sample number in the window which nearest length is $L = 1/a$, has no relations with previous samples, compares with initialization phrase, the computing speed is increased, but the corresponding parameter’s convergence rate is also decreased, which is suitable for the parameter updating when the system reaches steady phrase.

The left side in figure 4 is the taken updating, which shows a speedily convergence of the system in initialization phrase and reaches steady phrase; the right side shows the system’s convergence rate is slower in initialization phrase. The classical MOG algorithm respectively takes two different updating mechanism in initialization phrase and after the system stabled to reach the expectation result.

The improved algorithm unifies the two updating method, which not only considers the speedily convergence of system parameter in initialization phrase, but also researches the system parameter updating’s calculation after stable, makes it can quickly get expectation background parameter when the background has big changes. Figure 7 is the simulated updating process uses improved algorithm, shows it can quickly convergent to stable when the system has sudden changes.

The updating rate $\eta(t)$ is calculated as formula:

$$\eta_k = q_k \left( \frac{1 - \alpha}{c_k} + \alpha \right)$$

When the sample is matched to the Gaussian modal, $q_k = 1$, or $q_k = 0$, $c_k$ is a counting for the number of matching the modal and sample, $c_k = c_k + g_k$ when the Gaussian modal with the smallest weight is replaced by new modal, this modal $c_k$ resets as 0. The weight updating:

$$w_k(t) = (1 - \delta)w_k(t-1) + aq_k$$  (4)
The weight updating,
\[ \mu(t) = (1-\eta k) \mu(t-1) + \mu_k \cdot x \]

(5)

The variance updating,
\[ \sigma^2_k(t) = (1-\eta_k) \sigma^2_k(t-1) + \eta_k \cdot (\mu_k(t-1) - \mu_k)^2 \]

(6)

When system initializes or the background has big changes, the number of sample is small, \( \eta_k \approx 1/ck \), the system can quickly converge in a short period of time; when the system is stable, the number of sample is big, \( \eta_k = q_k \cdot (1-\alpha) / c_k + \alpha \rightarrow \alpha \), i.e., \( \eta_k \approx \alpha \), the system’s calculation is small, its performance is improved.

IV. BACKGROUND DISPLAY PATTERNS

The classical MOG algorithm directly uses the modal component value with the biggest weight as background display, without considering other (B-1) component value as background, it’s not accurately in display. At time t, the proportion of pixel gray value as x is:
\[ p(x) = \sum_{k=1}^K p(g_k) p(x | g_k) = \sum_{k=1}^K \omega_k g(x | \mu_k, \sigma_k) \]

(7)

According to Bayes formula, the proportion of pixel x as background is:
\[ p(b | x) = \frac{\sum_{k=1}^K p(b | g_k) p(h_k | x) p(h_k)}{\sum_{k=1}^K p(h_k | x) p(g_k)} \]

(8)

The background expectation value \( E(X | B) \) as background display is
\[ E(X | B) = \sum_{k=1}^K E(G_k | B) \]
\[ = \sum_{k=1}^K \mu_k \cdot P(B | G_k) P(G_k) \]
\[ = \sum_{k=1}^K \omega_k \cdot P(B | G_k) P(G_k) \]

(9)

\( P(G_k) \) is the weight of each modal, re-unifies the weight of background B modal in application, distributes the background’s each modal component by uniformed weight, the background display uses the mixture modal re-distributed model as the basis of array display.

V. EXPERIMENTAL RESULTS

A. Test Platform and Data Sources

In order to test the improved algorithm’s accuracy, tests it in multiple video and compares with classical Gaussian mixture background modeling method (MOG). All the test sequences uses the same parameter setting:
\[ K = 4, \alpha = 0.02, \lambda = 2.6, \tau = 0.76 \]

The platform for testing hardware: CPU is Intel Pentium Dual E2160, basic frequency is 1.80GHz, RAM is 1G. The platform for software: Linux Suse11.2 Opencv. The data used in the test contains three set of testing video: 1) uses the standard sequences offered by http://cvrr.ucsd.edu/aton/shadow, selects its indoor scene: Intelligent-room; 2) the testing video Tree comes with Opencv and a set of indoor scene Curtain; 3) offered simulated synthesizing dataset: http://www.vis.uni-stuttgart.de/index.php?id=sabs, simulated synthesizes artificially multiple motion detection’s common scene like breaches shaking, covering, shadowing, mirror reflecting, illumination varying. Makes motion detection test to each video in same experiment condition.

B. Evaluation Criterion of Performance

The recall and pression is the important indicator of reflecting detection performance. Recall, the percentage of correct result to all the detected result. Pression is the percentage of detected correct result to all the detected result. The recall:

\[ \text{Recall} = \frac{TP}{TP + FN} \]

(10)

And the pression is expressed as:

\[ \text{Precision} = \frac{TP}{TP + FP} \]

(11)

The two’s integrate indicator:

\[ F - \text{Score} = \frac{2 \cdot \text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}} \]

(12)

The background’s error rate:

\[ \text{FPR} = 1 - \frac{TN}{TN + FP} \]

(13)

TP: foreground of correct detection. TN: background of correct detection, FP: foreground of false detection (which should be background), FN: background of false detection (which should be foreground).

C. Compare and Analysis of the Experiment Result

Figure 8 is the result for the testing video tree. (a) is the initial image frame of testing video, (b) is the foreground image detected by classical MOG, (c) is the detected foreground image detected by improved Gaussian mixture algorithm. The experiment result shows, parts of the foreground detected by MOG is false detected shaking pixel points of the branches, and the foreground detected by improved algorithm has less pixel points, which shows the disturbance robustness to background of
improved Gaussian mixture algorithm is better than initial MOG algorithm in the aspects like overcoming background disturbance (branches shaking):

Figure 9 is the detection result of indoor scene Curtain. In this video, the curtain is floating with wind, which is a big disturbance to the motion detection. From the test result, the classical MOG would detects parts of the floating curtain as foreground, but the improved algorithm has little this kind of phenomenon, which only detected the motion object. It shows the improved algorithm has better robustness to background disturbance.

Figure 10 is the detection result of the scene Intelligent-room. This video is a indoor scene, involved to the background change condition like scene illumination changing. The selected frame is at the time when motion object just came to monitor area, from the detection result, in MOG algorithm, when the object just came in, because of the change of background, the detection result is not good; the improved algorithm, as its background model can convergent quickly, even the object just came in, it can quickly detect the motion object.

Figure 11 is the background frame’s display image structured by the two algorithms in the Tree video, among them, (a) is the initial image frame with motion object at its top right corner, (b) is the single Gaussian background display structured by the Gaussian modal with the biggest weight in classical MOG algorithm background model, (c) is the background frame structured by Gaussian mixture unified by each modal weight in background model. From the experiment. The background model of MOG algorithm has more noisy points, the background model of improved algorithm basically reaches the demand, the established background model includes due background, because the improved algorithm considers other background modal, its structured model can better overcome the disturbance made by background disturbance like shaking branches, the detailed color is not distorted, which is more fitted the real scene.

Figure 12 is the simulated dataset detection result offered by BRUTZER, S. This video is a simulation scene related to the background change condition of multiple motion detections like scene illumination changing, shadowing, covering and background disturbance. The branches is shaking in this scene, because of the cover of buildings, the shadow, cover and illumination change is obviously. The selected frame is the time when the two motion objects shows together, from the detection result, in MOG algorithm, when the objects just came in, because of the background’s change, the detection result is not good, which is affect a lot by the environment like shadowing and covering; the improved algorithm, because of the background model can convergent quickly, even the objects just came, it can detect the motion objects and detect them pretty completely.

Figure 13 is the index analysis curves of MOG algorithm and improved algorithm, the testing data is the simulation data in common scene offered by BRUTZER, S, selects the first 1400 frames and tests, draws the analysis curves. The analysis curves is corresponding to Recall, Precision, FPR and F-Score, the blue chain line is the curve of MOG, the black solid line is for improved algorithm. From the figure, the
Recall 0 0 0 0 0 0 0 0.

shadowing, and researches on the transplantation method texture information, eliminating the adverse effects of the next step of work, the research will focus on harmony algorithms and has high detection accuracy and recall. In rule of algorithm, the experiment showed, this algorithm made full use of sample information, designed the fuse convergence rate and detection accuracy. The algorithm techniques, to overcome the contradiction of its while reducing energy intensity and volume.

According to application feedbacks, the unusual video event is too slow (<40km/h). According to application feedbacks, the algorithm operates well in the system.

Figure 14 is the real-time screenshot detection result when the algorithm is used in highway real-time monitor system. The experiment chooses the real-time video of Huning highway, PEG-2, resolution is D1(704×576); seen as the figure, detected the unusual video event is too slow (<40km/h). According to application feedbacks, the algorithm performs well in this system.

VI. CONCLUSION

This text improves classical Gaussian (MOG) on the basis of deeply research on the background modeling techniques, to overcome the contradiction of its convergence rate and detection accuracy. The algorithm made full use of sample information, designed the fuse rule of algorithm, the experiment showed, this algorithm can better fitted the complex scene changes than previous algorithms and has high detection accuracy and recall. In the next step of work, the research will focus on harmony texture information, eliminating the adverse effects of shadowing, and researches on the transplantation method from algorithm to DSP and CUDA platform, realizes the promoting of system’s stability and processing capacity while reducing energy intensity and volume.

REFERENCES