A REAL-TIME TRACKING SYSTEM FOR TAILGATING BEHAVIOR DETECTION

Yingxiang Zhang, Qiang Chen and Yuncai Liu
Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University
Frogren2002, Chenqiang, China

Keyword: IGMM, Color histogram similarity, Tailgating detection.

Abstract: It is a challenging problem to detect human and recognize their behaviors in video sequence due to the variations of background and the uncertainty of pose, appearance and motion. In this paper, we propose a systematic method to detect the behavior of tailgating. Firstly, in order to make the tracking process robust in complex situation, we propose an improved Gaussian Mixture Model (IGMM) for background and combine the Deterministic Nonmodel-Based approach with Gaussian Mixture Shadow Model (GMSM) to remove shadows. Secondly, we have developed an algorithm of object tracking by establishing tracking strategy and computing the similarity of color histograms. Having known door position in the scene, we specify tailgating behavior definition to detect tailgater. Experiments show that our system is robust in complex environment, cost-effective in computation and practical in real-time application.

1 INTRODUCTION

Extensive researches have been conducted in video surveillance to improve social and private safeties. In our surveillance system, we detect behavior of tailgating in the scene. Detecting tailgating from video heavily depends on the detection of moving objects involved in each frame and the integration of this frame-based information to recognize special behavior. This high level behavior description relies on exact foreground detection and track of moving objects.

Much work has been done in background modeling and foreground detection. Karmann and Brandt, Kilger modeled each pixel with a Kalman Filter (Karmann & A. Brandt, 1990) to make the system robust with lightning changes in the scene. M.Piccardi used mean-shift approach (T. Jan & M. Piccardi, 2004) to express the mixed distribution of each pixel. Stauffer and Grimson (C. Stauffer & W. E. L Grimson, 1999) presented modeling each pixel as a mixture of Gaussians and using an on-line approximation to update the model. Shadow detection and removing is critical for accurate object detection in video streams since shadow is often misclassified as an object, causing errors in segmentation and tracking. The most common algorithms are Statistical Nonparametric approach, Statistical Parametric approach and Deterministic Nonmodel-Based approach (R. Cucchiara, M. Piccardi and A. Prati, 2003). Tracking requires matching the same moving objects in consecutive frame based on the extracted character. M. Isard and A. Blake proposed Condensation algorithm (M. Isard and A. Blake, 1998) which uses Monte Carlo approach to obtain statistic nonparametric model to track objects in real-time application. S.Birchfield studied color histogram models combining with Gradient model of density function (S.T. Birchfield 1998) to track.

In this paper, firstly, we propose an improved adaptive mixture Gaussian model (IGMM). In classic Gaussian Mixture Model (GMM) (C. Stauffer & W. E. L Grimson, 1999), Gaussian background model of each pixel is independent. Also, due to the fact that RGB of each pixel includes 3-5 Gaussian distributions and GMM updates every frame, high computational cost of GMM makes surveillance system difficult to accommodate real-time operation. In our approach, we improved GMM's updated speed and reduced the number of Gaussian model by segmenting different regions of the scene. We also use GMM to remove shadows by updating parameters of Gaussian distribution of suspicious shadow pixels. Secondly, we use color histogram to track and match moving objects (S. T. Birchfield, 1998). Object occlusion is a difficult problem in understanding of human behavior. We utilize seven

Zhang Y., Chen Q. and Liu Y. (2009).
A REAL-TIME TRACKING SYSTEM FOR TAILGATING BEHAVIOR DETECTION.
In *Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*, pages 398-402
DOI: 10.5220/0001665703980402

possible hypothetical situations to obtain a better matching. Lastly, by illustrating definition of tailgating, we sign every object stepping into the scene and constitute logical rules to catch candidate tailgater. Tested our algorithm using real scene video, the system gave satisfactory detection results.

This paper is organized as follows: Section 2-4 respectively describes foreground extraction, tracking scheme and definition of tailgating event. In section 5 we discuss the application environment and analyze the results in real-time surveillance scenes. Section 6 gives conclusion.

2 FOREGROUND EXTRACTION

2.1 Background Modeling

Gaussian Mixture Model (GMM) (C. Stauffer and W. E. L Grimson, 1999) has been widely used for background modeling as it adapts to deal robustly with lighting changes, repetitive motions of scene elements. However, to process more complex environment, it should make more Gaussian distributions to resist disturbance. And based on the principle of updating GMM, all the parameters of Gaussian function of each pixel have to be updated in every next frame which increases computation and make real-time work difficult.

We notice that only a few parts of one frame suffer changes while other regions are static in the surveillance scene. Pixels in static regions always match one Gaussian distribution which has a high possibility to match the same pixel in the following frames. After learning, the distribution weight ω has a large value while deviation σ2 becomes comparatively smaller. In a long time the value of ω/σmaximizes and mean of the same Gaussian distribution would be regarded as pixel value of background. It is not necessary to update background model in static regions in every frame. In Improved Gaussian Mixture Model (IGMM) we proposed, when the probability that certain old Gaussian distribution matches every new pixel value in the same region exceeds chosen threshold, we will not update parameters of background distributions of the same pixel in the next 100 frames. After 100 frames, we reset parameter ω of Gaussian distributions to be equal and update GMM until that probability reaches chosen threshold again.

2.2 Shadow Removing

The principle of GMM can also be used for removing shadow. In this paper, we combine Deterministic Nonmodel-Based approach (R. Cucchiara, M. Piccardi and A. Prati, 2003) with GMM to remove shadow.

In the first step, we detect shadow taking advantage of relationship between the reference frame and background in HSV space. HSV color space explicitly separates chromaticity and luminosity and has been proved easier than the RGB space to set a mathematical formulation for shadow detection. A shadow mask $SP(k)$ at frame k defined as follows (R. Cucchiara, M. Piccardi & A. Prati 2003):

$$SP_{i,j}(k)=\begin{cases}1 & if\ \alpha\le\frac{I_{i,j}(k).V}{B_{i,j}(k).V}\le\beta\, and\left|I_{i,j}(k).S-B_{i,j}(k).S\right|\le\gamma\, and\, D_H(k)\le\tau \\ 0 & otherwise\end{cases} \quad (1)$$

Where, $I_{i,j}(k)$ and $B_{i,j}(k)$ are the pixel values at coordinate (i,j) in the input image(frame k) and in the background model (computed in the frame k) respectively, H,S ,V means hue, Saturation and Value of one pixel, α ,β ,γ ,ζ denotes threshold,

$$D_H(k)=\min\left(\left|I_{i,j}(k).H-B_{i,j}(k).H\right|,360-\left|I_{i,j}(k).H-B_{i,j}(k).H\right|\right) \quad (2)$$

In the second step, Gaussian Mixture Shadow Model (GMSM) learns and updates its parameter based on those candidate pixels of shadow detecting in HSV space. If the relationship between candidate pixels of shadow and certain Gaussian distribution of GMSM meets:

$$|I_t-\mu^s_{i,t-1}|\le D^s\times\sigma^s_{i,t-1},\ \text{i=1,2,...K}, \quad (3)$$

Where, $G\sim(\mu^s,\sigma^s)$, S means GMSM and D^s=2.5,K is number of Gaussian distributions. Then parameters updated as follows:

$$\omega^s_{i,t}=(1-\alpha^s)\omega^s_{i,t-1}+\alpha^s \quad (4)$$

$$\mu^s_{i,t}=(1-\rho^s)\mu^s_{i,t-1}+\rho^s I_t \quad (5)$$

$$(\sigma^s_{i,t})^2=(1-\rho^s)(\sigma^s_{i,t-1})^2+\rho^s(I_t-\mu^s_{i,t})^2 \quad (6)$$

Here, α is learning ratio which relates to background updating speed and $0\le\alpha\le1$. ρ is learning ratio of parameters and $\rho\approx\alpha/\omega^s_{i,t}$.

If there is no Gaussian distribution matching candidate pixel of shadow I_t, new Gaussian distribution will replace the distribution with the smallest weight. New distribution will have a comparatively larger standard deviation σ_0^s and smaller weight ω^s_0 and mean I_t. Other Gaussian distributions keep the same mean and deviation, but the weights will decrease:

$$\omega^s_{i,t}=(1-\alpha^s)\omega^s_{i,t-1} \quad (7)$$

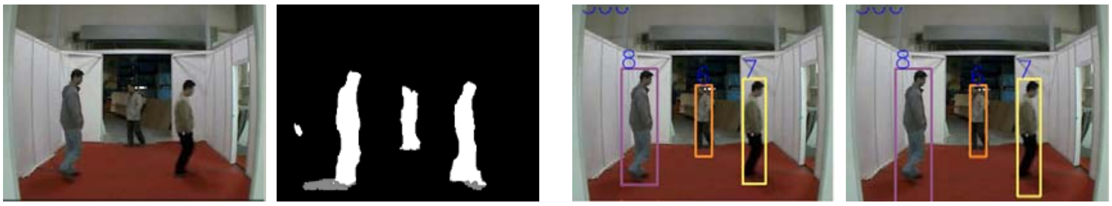

Figure 1: From left to right, the first figure is original image. The second one is foreground with foreground as white and shadow pixels as grey. The last two compares the accuracy of object tracking with and without shadow removing respectively.

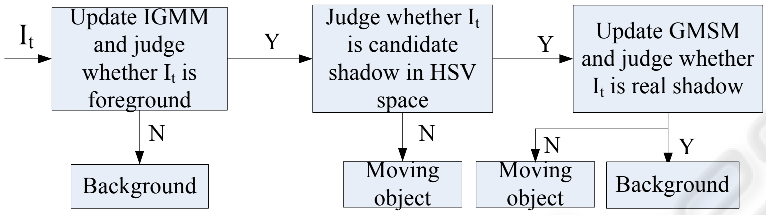

Figure 2: Flow chart of foreground extraction.

Finally, after normalization of all the weights of Gaussian distributions, we array Gaussian distributions as $i_1^s, i_2^s, \ldots, i_{ks}^s$ in accordance with the value of $\omega_{i,t}^s / \sigma_{i,t}^s$. If preceding N distributions meet following principle:

$$\sum_{K=i_1^s}^{i_N^s} \omega_{k,t}^s \geq \tau^s \tag{8}$$

Then these N distributions can be regarded as shadow distribution. If value of candidate shadow pixel I_t does not match every shadow distribution, this pixel is defined as moving objects. However, if it matches one shadow distribution, it would be eliminated from foreground as shadow. The result of shadow removing can be seen in Figure 1 and the detailed backgrounding flow chart is shown in Figure 2.

3 TRACKING METHOD

3.1 Tracking Strategy

After motion detection, surveillance systems generally track moving objects from one frame to another in an image sequence. One of the main difficulties of the tracking process concerns the partial or total temporal occlusions of the objects. Therefore, the splitting of the blobs in order to separate or isolate the objects has to be considered in this section. Next are seven possible situations in matching strategy:

1) Normal Event: If two blobs of two consecutive frames are overlapped, we think these two blobs belong to one motion object.
2) Entry Event: A blob at the current frame with no matching blob of previous frame is thought as an entering object.
3) Exiting Event: Finding no matching blob at the current frame for a blob of previous frame, and the place of this blob is at the edge of frame.
4) Stopping Event: The place of this blob is not at the edge of image and the color similarity for the places of blob at the consecutive frame is larger than a threshold.
5) Vanishing Event: Finding no matching blob at the current frame for a blob of previous frame.
6) Splitting Event: More than one blob match a previous blob.
7) Merging Event: A blob at current frame matches more than one previous blob.

This strategy helps effectively handle problems such as disappearance, occlusion and crowd.

3.2 Color Histogram Similarity

Objects can be divided and merged. When objects merge, the labels are inherited from the oldest parent region. When an object splits, all separating objects inherit their parent’s labels. In order to track people consistently as they enter and leave groups, each person’s appearance must be

Figure 3: The two rows display the processed sequence using our method. The first row displays outdoor scene with camera placed at side of the site while the second row gives indoor tailgating scene e with camera placed at top of the site.

Table 1: Detecting result of tailgating event.

	GMM K=3	IGMM K=3	GMM K=5	IGMM K=5
Scene A	9	12	7	10
Scene B	8	11	6	8

Table 2: Comparison of average processing speed between IGMM and GMM, unit (frame/s).

Frame frequency	Real events num	Detected events num	event detection rate
10frame/s	30	28	93.3%
15frame/s	30	27	90%
20frame/s	30	25	83.3%

modeled. A color model (R. P. Perez, C. Hue and J. Vermaak, 2002) is built and adapted for each person being tracked. A histogram H_i (X | k) counts the number of occurrence of X=(R, G, B) within the person i at frame k. Because the same person may have different sizes in different frames, we should normalize H_i(X | k).

$$P_{i\_new}(X \mid k) = H_i(X \mid k) / A_i(k) \quad (9)$$

Where, A_i (k) means the number of pixels within the mask for person i at frame k.

Histogram models are adaptively updated by storing the histograms as probability distributions:

$$P_i(X \mid k+1) = \alpha P_{i\_new}(X \mid k+1) + (1-\alpha) P_i(X \mid k) \quad (10)$$

To determine whether two blobs at frame m and n belong to one object, similar degree (similarity) of color is computed.

$$degree_{i,j}(X \mid m,n) = \frac{P_i(X \mid m) \times P_j(X \mid n)}{|P_i(X \mid m)| \, |P_j(X \mid n)|} \quad (11)$$

Where, P_i(X | m) means normalized histogram of X=(R, G, B) for person i at frame k. If $degree_{i,j}$ (X| m, n) is larger than a threshold, we think two blobs belong to one object.

4 TAILGATING DETECTION

On the basis of tracking, effective recognition of Tailgating behavior requires specific definition of this behavior. Tailgating means one person steps behind former one and slides into the entrance without allowance. The tailgating behavior contains two steps. Firstly, we define Euclidean distance between two walkers as a standard distance to judge tailgater. If one person steps behind the former with the same direction and a higher speed, he or she may gradually approach the former and when the distance between them is smaller than the standard distance, he is regarded as a tailgater

$$\cdot L <= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (12)$$

Here, L means the standard distance user defines and (x1, y1) denotes the center of the former, (x2, y2) denotes the center coordinate of tailgater.

The Second phase is to assure the tailgater entering into the entrance following the authorized person in a short time. When the candidate tailgater tails after the former, the two blobs always merge into one big blob with two labels.

Comparing Euclidean distance between centers of the candidate blobs and door with a threshold, we confirm the range of tailgating. If this distance is smaller than the threshold, we judge whether Vanishing Event happens near the door. If Vanishing Event happens, the behavior will be recorded.

5 RESULTS AND ANALYSIS

To illustrate the proposed method, we experimented in an exhibition hall using one digital camera with image dimension of 320*240. Figure 3 shows a sequence of indoor and outdoor scenes containing walking people who imitates tailgating behavior. The sequence includes lightning changes caused by strobotron, reflections from windows and moving shadows. With traditional tracking method, tracking often failed in the following cases:

Case 1: People walk close to each other. In this case, due to the closely distributed foreground points, extracting single object based on connected components algorithm is difficult.

Case 2: Two people walk across each other or one occludes another. In this case, two different objects merging into one mixed component brings difficulty in tracking the right one with exact trajectory.

Case 3: One connected component disappears in the scene without touching border. In this case, many tracking algorithm can not determine whether this moving object really disappears from the scene or just stays still in the scene.

In our paper, method we proposed can decrease moving shadow and other disturbance comes from environment using IGMM and GMSM to extract single connected component. We have numbered moving objects in the scene and compute the similarity of color histogram between points at current frame and background to judge object's disappearance or stillness. With the direction of velocity and color information, we can process most examples of merging and detaching. Table 1 shows processed results from the indoor sequence Exhibition Hall.

We find the veracity of our algorithm firmly related to the frame frequency of the camera or video. More than 90% of tailgating events can be detected at 10frame/s and more than 80% of the events can be detected at higher frame frequency.

Table 2 provides the evidence that our IGMM has increased processing speed especially in certain surveillance environment where active region of moving objects occupies only a part of the whole scene.

6 CONCLUSIONS

We have proposed an algorithm to detect tailgating behavior using background modeling, tracking strategy and behavior definition. There are three main issues in the process of surveillance. First is acquiring true foreground in complex environment by making use of IGMM and GMSM. Second is effectively tracking by means of considering different situation and matching objects in consecutive frames through similarity computation of color histograms. Third is anti-tailgating taking advantage of definition of tailgater. Compared with other methods in surveillance, our novel algorithm is cost-effective and useful in real-time application.

ACKNOWLEDGEMENTS

This research work is supported by National 973 Program of 2006CB303103 and National Natural Science Foundation of China (NSFC) under Grant No.6083099.

REFERENCES

Karmann and A. Brandt. 1990. "Moving object recognition using an adaptive background memory." *Time-Varying Image Processing and Moving Object Recognition*, 2, Elservier, Amsterdam, Netherlands,.

T.Jan, M.Piccardi. 2004. "Mean-shift background image modeling." In *Proceedings of the international conference on image processing*.

C. Stauffer, W. E. L Grimson. 1999. "Adaptive background mixture models for real-time tracking". In *Proc. Conf. Comp. Vision Pattern Rec., Vol*. 2.

R. Cucchiara, M. Piccardi, A. Prati. 2003. "Detecting moving objects, ghosts and shadows in video streams". *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 25(10):1337 ～ 1342.

M. Isard, A. Blake. 1998. "Condensation-conditional density propagation for visual tracking". *International Journal of Computer Vision*, 29(1): 5-28.

S.T. Birchfield. 1998. "Elliptical head tracking using intensity gradients and color histograms". *In Proc. Conf. Comp. Vision Pattern Rec*., pages 232–237, CA.

R. P. Perez, C. Hue, J. Vermaak. 2002. "Color-based probabilistic tracking." In *Proceedings of the 7th European Conference on Computer Vision.* Berlin, Germany, Springer-Verlag, 661-675.