Multiple Layer Based Background Maintenance In Complex Environment

Tao Yang[1], Quan Pan[1], Stan Z.Li[2], Jing Li[1]

[1]College of Automatic Control, Northwestern Polytechnical University, Xi'an, China , 710072

[2]National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, P.R.China, 100080

yangtaonwpu@163.com , quanpan@nwpu.edu.cn , szli@nlpr.ia.ac.cn , jinglinwpu@163.com

Abstract

A fast and efficient multiple layer background maintenance model is built to conserve the original and the current background separately. Fusing the properties of object motion in image pixels and the changes between the input video and the multiple background layers, this method could handle various sources of scene changes, including ghosts, abandon objects and illumination changes. An intelligent video surveillance system is developed to test the performance of the algorithm. Experiments are performed using long video sequences under different conditions indoor and outdoor. The results show that the proposed algorithm is effective and efficient in real-time and accurate background maintenance in complex environment.

1. Introduction*

Background maintenance in video sequences is a basic task in many computer vision and video analysis applications, for instance, video surveillance [1,2,3], indexing for multimedia [4,5,6], people detection and tracking [1,2], perceptual human-computer interface. Accurate background maintenance will greatly improve the performance of moving object detection, tracking, recognition, classification and activity analysis.

To design the background maintenance algorithm, several serious problems have to be concerned. The first one comes from fast and accurate background modeling. For instance, if an object stops in the scene for enough time, although it will be updated to the background in almost all the existing background maintenance algorithms, it still has a drawback in that when it starts moving, it will cost the background update algorithms a period of time to deal with the ghost. How to design a precise and fast background update algorithm is a challenge. Other problems such as ghost, abandon object, and illumination changes would also bring great challenges.

The most common approaches to identifying the moving objects are optical flow [7] and background subtraction [1,2,3,8,9,10,11] based ones. Optical flow can be used to detect independently moving objects in the presence of camera motion. However most optical flow based methods are computationally complex and cannot be applied to full-frame video streams in real-time without specialize hardware. Chris Stauffer [8] deal with motion segmentation problem based on an adaptive background subtraction method by modeling pixels as a mixture of Gaussians and uses an on-line approximation to update the model. Several improvements on Gaussian mixture modeling have been made by P.KaewTraKulPong [9,10]. Robert [1] presents a three-frame differencing operation to determine regions of legitimate motion, followed by adaptive background subtraction to extract the entire moving region. Toyama [11] developed a called three level processing (Wallflower) approach for background maintenance. K. Kim [12] presented a codebook (CB) background subtraction algorithm to construct a background model from long observation sequences. Even though many background models have been proposed in the literature, the problem of background maintenance in complex environment is still far from being completely solved.

This paper addresses the problem of real time background maintenance in complex environment. Rather than relying upon the distribution of the pixel value [7,9,10], two background layers is presented to conserve the original and the current background separately. Moreover, through analyzing properties of object motion in image pixels and the background subtraction results, a multiple layer background update model is developed to select and maintain the suitable background layer under different conditions. An intelligent video surveillance system is developed to test the performance of the algorithm. Experiments are performed using video sequences under different conditions indoor and outdoor. The results show that the proposed algorithm is effective and efficient in accurate background maintenance in complex environment. In particular, it is highly computationally cost effective and thus provides enough time for further target detection, tracking and classification.

The remainder of this paper is organized as follows. Section 2 outlines the algorithms. Section 3 describes the background maintenance algorithm. Section 4 and 5 contain the experimental results under various conditions and conclusion.

* The work presented in this paper was performed at NWPU and Microsoft Research Asia. T. Yang, Q. Pan, and J. Li were also sponsored by the foundation of National Laboratory of Pattern Recognition of China.

2. Outline of the algorithm

The flow diagram of the algorithm is shown in Figure 1. There are four major parts: pixel level motion detection, multiple layer background update, background subtraction and object segmentation. Pixel level motion detection identifies each pixel's changing character over a period of time by frame-to-frame difference and analyzes the dynamic matrix presented in this paper. Fusing the detection result of pixel and the background subtraction results, the multiple layer background update model will select and maintain the suitable background layer under different conditions. In background subtraction step, each video frame is compared against the reference background layer, pixels in the current frame that deviate significantly from the background layer will be detected. In the end, an object segmentation unit based on connected blob extraction and image down sampling is used to segment the moving objects.

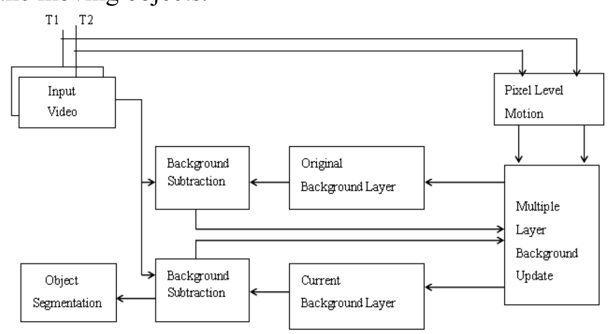

Figure 1. The block diagram of the algorithm

3. Multiple layer maintenance

Background subtraction, the process of subtracting the current image from a reference one, is a simple and fast way to obtain the moving object in the foreground region and has been employed by many surveillance systems [1]. In our multiple layer background maintenance model, the pixel motion character is used to decide which pixel should be updated, and the subtraction results between the input video and the two layers are used to make a decision which layer should be updated.

The basic idea of our pixel level background update method comes from an assumption that the pixel value in the moving object's position changes faster than those in the real background. Fortunately, this is a valid assumption in most application fields such as traffic video analysis, people detection and tracking in intelligent meeting, outdoor security surveillance in residential area, parking lot and entrance. Under this assumption, we can distinguish the foreground and background accurately by a simple frame-to-frame difference method, which could only detect the fast changes of pixel.

However, this method will fail when the inside color of object is uniform. In this situation, pixel values do not vary within the object. To deal with this problem, we present a dynamic matrix $D(k)$ to analyzing the changes detection result of the frame-to-frame difference method, where the motion state of each pixel is stored in the matrix. Only those pixels whose values do not change much can be updated into the background.

Let $I(k)$ denotes the input frame at time k, and the subscript i, j of $I_{i,j}(k)$ represent the pixel position. Equation (1) and (2) show the expression of frame-to-frame difference image $F(k)$ and the dynamic matrix $D(k)$ at time k.

$$F_{i,j}(k) = \begin{cases} 0 & \left| I_{i,j}(k) - I_{i,j}(k-\gamma) \right| \le Tf \\ 1 & otherwise \end{cases} \tag{1}$$

$$D_{i,j}(k) = \begin{cases} D_{i,j}(k-1) - 1 & F_{i,j}(t) = 0, D_{i,j}(k-1) \ne 0 \\ \lambda & F_{i,j}(t) \ne 0 \end{cases} \tag{2}$$

Where γ represent the interval time between the current frame and the old one, Tf is the threshold to make a decision whether the pixel is changing at time k or not, and λ is the time length to record the pixel's moving state, once the $D_{i,j}(k)$ equates to zero, the pixel update method will make a decision that this pixel should be updated into the background.

Although the pixel update method performs well in many application fields, it still has a drawback that once the targets speed changes a lot in the field of view (for instance, the speed of vehicle is quite faster than the human being's), it's hard to unify the time threshold λ in the dynamic matrix $D(k)$. On the one hand, a low threshold λ will increase the false update probability of the low speed target and bring ghost problems. On the other hand, a high threshold λ will increases the detection probability of the low speed target and thus avoid false update, however, this rigor condition make the update step very sensitive to noises and as a result it will not work under this condition.

Here, we present a multiple layer background model to solve the ghost problem. The basic idea of the method is that when a new object is updated to the current background, the previous intensity values of the updated pixels will be conserved in a new layer called original background layer. Once the object leaves, the conserved pixels will be restored to the current background and thus there will have no ghost in the subtraction image.

The differences of the input video and the multiple background layers are used to decide whether the pixel value changes caused by old object leaves or new object entering.

Let $SC(k)$ and $SO(k)$ denote the subtraction images in the current background layer and the original background layer at time k, Ts is the subtraction threshold, and $I.R(k)$, $I.G(k)$, $I.B(k)$, $BC.R(k)$ $BC.G(k)$ $BC.B(k)$, $BO.R(k)$ $BO.G(k)$, $BO.B(k)$ represent R,G,B pixel value of input frame and multiple background layer model separately.

$$SC_{i,j}(k) = \begin{cases} 1 & \begin{array}{l} \left|I.R_{i,j}(k) - BC.R_{i,j}(k)\right| > Ts, or \\ \left|I.G_{i,j}(k) - BC.G_{i,j}(k)\right| > Ts, or \\ \left|I.B_{i,j}(k) - BC.B_{i,j}(k)\right| > Ts \end{array} \\ 0 & \quad otherwise \end{cases} \tag{3}$$

$$SO_{i,j}(k) = \begin{cases} 1 & \begin{array}{l} \left|I.R_{i,j}(k) - BO.R_{i,j}(k)\right| \le Ts, and \\ \left|I.G_{i,j}(k) - BO.G_{i,j}(k)\right| \le Ts, and \\ \left|I.B_{i,j}(k) - BO.B_{i,j}(k)\right| \le Ts \end{array} \\ 0 & \quad otherwise \end{cases} \tag{4}$$

Here, a simple but effective rule is used to distinguish the old object leaving and the new object entering. If $SC_{i,j}(k)$ and $SO_{i,j}(k)$ are all equal to one, it's old object leaving, if $SC_{i,j}(k)$ equals to one and $SO_{i,j}(k)$ equals to zero, it's new object entering.

The main advantage of this multiple layer technique is that it reduces the uncertainty of a pixel update step, and makes the update process fast and robust even in complex environment. Moreover, multiple layers conserve more scene information and will be helpful for application like abandoned bag detection for security surveillance.

4. Experiment results

We have developed a real time intelligent surveillance system based on the presented algorithm. The system is implemented on standard PC hardware (Pentium IV at 1.3GHz). The video image size is 320x240 (24 bits per pixel). The system is tested in typical indoor and outdoor environments for handling ghost situation, abandon object, and object detection and tracking. The average frame rate of our system is 14.5fps. The following presents results.

4.1. Ghost

This situation is due to a background change caused by a flower that is moved in the field of view and becomes part of the background but now starts its motion. Figure 3 measures reactivity in a limit condition when the background reflects changes from a rose that starts its motion after having previously been part of the background. Here, we contrast the segmentation with and without multiple background layers. While the flower is placed on the desk, it is included in the background image in both background model (Figure 3, first column). At frame #514, (Figure 3, second column), a person is moving the flower to another place. In the background model without multiple layer, until frame #553 (Figure 3, third column), the moved flower still substantially covers the area where it was stopped and a ghost can be found clearly in the segmentation result. (Figure 3, second row, second and third column).However, the correct background update and the correct segmentation can be achieved with our approach without time delay (Figure 3,

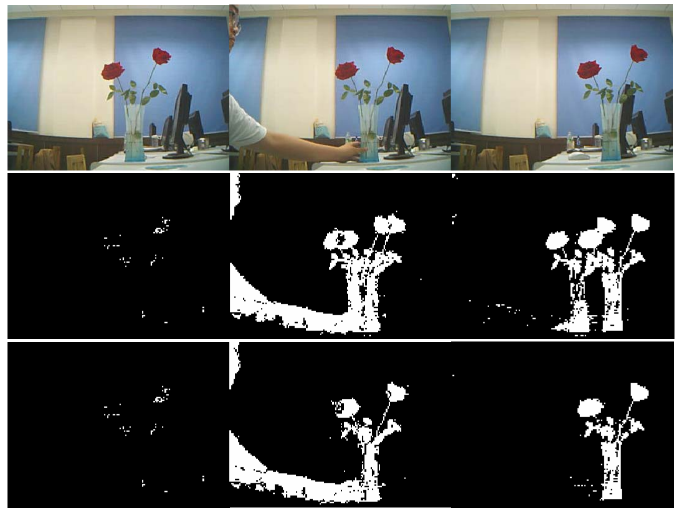

Figure 3. The reactivity of background model when ghost happens. These images are captured at frame #417, frame#514, frame #553 separately. The first row contains the input video. The second row contains the object segmentation result without multiple layer. The third row contains the object segmentation result with multiple layer.

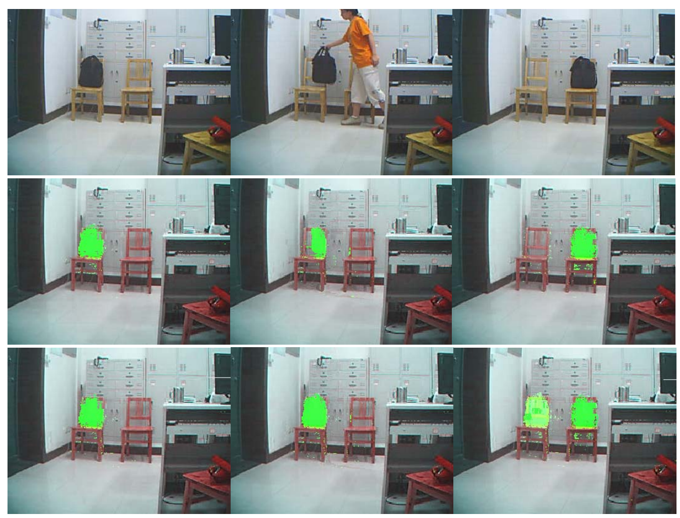

Figure 4. Abandon object detection. These images are captured at frame #262, frame#432, frame#541. The first row contains the input video. The second row contains the abandon object detection result of our method. The third row contains the detection result without multiple background layer. The green area represents the abandon object.

third row, second and third column). The same results can be achieved for the low speed target segmentation which is easily be updated into the background model and bring ghost problem in the traditional background maintenance model.

4.2. Abandon object detection

Although surveillance cameras are already prevalent in banks, stores, and parking lots, video data currently is used only "after the fact" as a forensic tool, thus losing its primary benefit as an active, real-time medium. Intelligent video surveillance system should has the ability to analyze human motion character and thus to prevent the crime. Figure 4 shows the abandon bag detection result of our algorithm. The fist row is captured from the input video. The second row shows the current background layer of our method, in which the abandon bag is filled with green color. The last row represents the detection result without multiple layers. The main advantage of our method is that even the abandon bag has been updated into the background (Figure 4, first column), when the person moves it to another chair (Figure 4, third column), the correct background can be recovered with little time delay in our method (Figure 4, second row, second and third column), while the ghost is also be detected as a abandon object in the detection result without multiple layers (Figure 4, third row, third column).

5. Conclusion

This paper has presented a real-time and accurate algorithm for background maintenance in complex environment. The algorithm has the characteristic of explicitly addressing various difficult situations such as ghosts, abandon object, and illumination changes. Further it avoids problems caused by undesired background modification. Based on this algorithm, an intelligent video surveillance system has been developed and experiment results proved that this system performed robustly in quite different indoor and outdoor environment. Moreover, based on the moving object algorithm of this paper, which is highly computationally cost effective, the system can perform in real-time even on common PC.

6. References

[1] Collins, Lipton, Kanade, Fujiyoshi, Duggins, Tsin, Tolliver, Enomoto, and Hasegawa, "A System for Video Surveillance and Monitoring", VSAM Final Report, Technical report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May, 2000.

[2] Haritaoglu, I. Harwood, D., Davis, L.S., " W4: real-time surveillance of people and their activities", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume: 22 , Issue: 8 , Pages:809 – 830. Aug. 2000.

[3] Pakorn KaewTrakulPong, Richard Bowden, "A real time adaptive visual surveillance system for tracking low resolution colour targets in dynamically chaning scenes", *Image and Vision Computing* 21, Pages:913-929, 2003.

[4] Cucchiara, R., Grana, C., Piccardi, M., Prati, A., "Detecting moving objects, ghosts, and shadows in video streams", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Volume: 25 , Issue: 10 , Pages:1337 – 1342, Oct. 2003.

[5] Xu, H., Younis, A.A., Kabuka, M.R., "Automatic Moving Object Extraction for Content-Based Applications", *IEEE Transactions on Circuits and Systems for Video Technology*, Volume: 14 , Issue: 6 , Pages:796–812, June 2004.

[6] Shu-Ching Chen; Mei-Ling Shyu; Peeta, S.; Chengcui Zhang; " Learning-based spatio-temporal vehicle tracking and indexing for transportation multimedia database systems", *IEEE Transactions on Intelligent Transportation Systems ,* Volume: 4 , Issue: 3 , Pages:154–167, Sept. 2003.

[7] J. Barron, D. Fleet, and S. Beauchemin. "Performance of optical flow techniques", *International Journal of Computer Vision*, 12(1): Pages:42–77, 1994.

[8] Stauffer, C, Grimson, W.E.L., "Learning patterns of activity using real-time tracking", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume: 22 , Issue: 8, Pages:747 – 757, Aug. 2000.

[9] P.KawTraKulPong, R.Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection," *Proceedings of Second European Workshop on Advanced Video-based Surveillance Systems*,2001.

[10] Liyuan Li, Weimin Huang, Gu, I.Y.H., Qi Tian, "Foreground object detection in changing background based on color cooccurrence statistics", *(WACV 2002). Proceedings. Sixth IEEE Workshop on Applications of Computer Vision* , 3-4 , Pages:269 – 274, Dec. 2002.

[11] K. Toyama, J. Krumm, B. Brumitt, and B.Meyers "Wallflower: Principles and practice of background maintenance", *Proceedings of IEEE International Conference on Computer Vision*, Pages:255–261,1999.

[12] T. H. Chalidabhongse, K. Kim, D. Harwood and L. Davis, "A Perturbation Method for Evaluating Background Subtraction Algorithms," *Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS 2003)*, Nice, France, Oct. 11-12, 2003.