Improved Video-Based Vehicle Detection Methodology

Jinman Luo
School of Computer Science & Engineering
South China University of Technology
Guangzhou, China
shouhunizhutou@126.com

Juan Zhu
School of Computer Science & Engineering
South China University of Technology
Guangzhou, China
zjuan.scut@163.com

***Abstract*—Focusing on the problem that the detection accuracy of traffic detection system is sensitive to the changes of complex environments, this paper presents an improved method of vehicle detection. It builds and updates the background adaptively. Additionally, to improve the computation efficiency of shadow elimination, a fast algorithm of neighbor mean based on HSV model is proposed. As the occlusion is inevitable, a new solution is presented to deal with occlusion in this paper. First, a method based on Kalman filter is applied for occlusion identification. And then a search algorithm of template matching based on hierarchical pyramid is utilized for real-time segmentation. Experimental results have shown that the proposed method is effective and high real-time, and it can effectively improve the detection rate of video-based traffic detection system.**

Keywords-vehicle detection; shadow elimination; HSV model; Kalman filter; template matching

I. INTRODUCTION

Intelligent Transportation Systems (ITS) is the development trend of traffic monitoring in the 21st century. Traffic incident automatic detection system is an important part of ITS. It processes the video signal using image recognition technology to simulate artificial methods of identification of traffic anomalies in complex weather conditions such as rain, snow, fog, cloudy, day and night-time. In addition, incident detection system can improve the efficiency of traffic relief and reduce the losses caused by traffic accidents.

Vehicle detection technology based on video is the core of traffic incident automatic detection system. But in the real environments, the vehicle detection approaches, including difference approach and background modeling approach, are vulnerable to the changes, shadows and occlusion, resulting in lower detection rate. Difference approach is faster, but is vulnerable to the impact of background noise and cannot adapt to the slow changes in background. Background modeling method is robust for background noise and able to adapt to the slow changes of background, but it takes a lot of computing resources so that it cannot be applied for real-time system. This paper presents an improved method of vehicles detection based on adaptive background model. It extracts moving vehicles using background subtraction, and uses a method of local-mean based on HSV model to eliminate shadow. In order to identify overlapped vehicles, a method of occlusion identification based on Kalman filter is proposed. In order to segment overlapped vehicles, this paper utilizes a search algorithm of template matching based on hierarchical pyramid for real-time segmentation.

II. IMPROVED VEHICLE DETECTION METHOD

A. Background Model

The effectiveness of vehicle detection is affected by background model directly. A realistic background model is needed to improve detection accuracy. Gaussian mixture background model is widely use because of its robustness and stability, and it can adapt to environmental changes. However, its high computational load is not suitable for high real-time detection system. In order to meet the special and real-time processing requirements of detection system, this paper uses a sequence mean method of statistical model to build background. Taking into account the height of installation of highway camera is high and speeds of vehicles are fast, so the sequence mean method can apply for highway vehicle detection. As this paper is based on the RGB image for detecting, so it needs to compute RGB separately as follows:

$$\begin{aligned} B_{c,t}(x,y) &= Mean(F_{c,t}(x,y), F_{c,t-1}(x,y), \ldots, F_{c,1}(x,y)) \\ &= \frac{1}{t}\left(\sum_{t=1}^{t} F_{c,t}(x,y)\right) \end{aligned} \tag{1}$$

where c is one of three primary colors, t is the current time, $B_{c,t}(x,y)$ is a sequence mean of c at pixel (x,y), $F_{c,t}\ (x,y)$ is a value of c at pixel (x,y) .The value of n can be adjusted by traffic flow of road. If traffic flow is heavy, then n should be increased. And n should be decreased to save the time of building background for the other hand.

The environment of actual road changes constantly. In order to adapt to a different environment, background model needs to update regularly. This paper proposes a method of adaptive update of background to decrease the time of the background update and avoid considering stopped vehicles as background. The information of detected objects is used to adjust the update of background by this method. If $F_{c,n-1}(x,y)$ is foreground, then the value of $B_{c,n-1}(x,y)$ is calculated in background update to replace this point so that background update can avoid mixing pixels of foreground. This adaptive method of background update as follows:

$$F_{c,n}^{'}(x,y)=\begin{cases}B_{c,n-1}(x,y) & if(F_{c,n}(x,y)\in foreground)\\ F_{c,n}(x,y) & if(F_{c,n}(x,y)\in background)\end{cases} \quad (2)$$

This is only the background pixels are used to update the background. Therefore, it can decrease the time of background update and avoid affecting the detection by the ghost. Experimental results show that adaptive method of background update can be completed within 8 seconds, and does not affect real-time running.

Image information is loss when the color image is converted to grayscale image, so background subtraction on gray level often causes errors by those gray values of vehicles are close to the road. Taking into account the distinction of three primary colors between foreground and background, then can utilize the threshold segmentation method for foreground detection. First, subtract the three primary colors of foreground and background respectively and get the maximum difference value among them. And then this maximum value is compared with the user-defined threshold. The value is classed as foreground when it's greater than the threshold, and classed as background on the other hand. It can be calculated by (3) and (4).

$$\begin{aligned}D(x,y)=\max(&|F_{R,t}(x,y)-B_{R,t}(x,y)|,\\ &|F_{G,t}(x,y)-B_{G,t}(x,y)|,|F_{B,t}(x,y)-B_{B,t}(x,y)|)\end{aligned} \quad (3)$$

$$C(x,y)=\begin{cases}1 & D(x,y)\geq T\\ 0 & D(x,y)<T\end{cases} \quad (4)$$

Thereinto, *D(x,y)* is a maximum difference value at pixel *(x,y)*, *C(x,y)* is a result of segmentation at pixel *(x,y)*, *T* is a user-defined threshold.

B. Denoting of Vehicles

In extraction of vehicles, often use the smallest enclosing rectangles of connected regions to denote vehicles. However, if the direction of roadside camera angle is not parallel to the road and there is an angle between them, then the image obtained is biased in certain angle. So the path of vehicles moving on the road is not parallel to the direction of the camera angle. Then using traditional method of smallest enclosing rectangles to denote vehicles isn't applicable for this situation. This paper puts forward the hexagon based on the direction of drive line bias to denote vehicles.

In order to denote vehicles more fitness that need to recursively search for the object matrix to find out four points (the top, the bottom, the most left and the most right) of vehicle as the bias matrix's four-point. But the calculation of this bias matrix involves a number of rectangular recursive algorithms. To simplify operation here, use a simple method of approximate simulation to denote vehicles. First of all, obtain the smallest enclosing rectangle of connected region of vehicle, and then find out six points of rectangle based on the direction of drive line bias to get the hexagonal of vehicle approximately. If drive line is tend to left, then take the upper left corner point of enclosing rectangle as the first point of hexagonal, the midpoint on top as the second point of hexagonal, the midpoint of right side as the third point of hexagonal, the lower left corner point as the fourth point of hexagonal, the midpoint of the bottom as the fifth point of hexagonal and the midpoint of left side as the sixth point of hexagonal, eventually formed a special hexagonal shape of the vehicle and it can adapt to the cases of drive line bias. The result is shown in Fig. 1.

Figure 1. Denoting of Vehicle.

C. Shadow Elimination

The types of vehicles are easily misjudged by system because of the existence of shadow. Not only that, detection rate can be reduced when multiple vehicles are considered as a vehicle affected by shadow. Because the HSV (Hue, Saturation, Value) color model reflect the color consistency, and it retains the accurate information of grayscale and color that can highlight the difference between shadow and moving foreground, so this paper uses HSV model to eliminate shadow as follows:

$$Shadow(x,y)=\begin{cases}1 & \alpha\leq\Delta V\leq\beta\ \text{且}\ \Delta S\leq T_s\ \text{且}\ \Delta H\leq T_H\\ 0 & other\end{cases} \quad (5)$$

thereinto: $$\Delta V=\frac{I_v(x,y)}{B_v(x,y)} \quad (6)$$

$$\Delta S=I_S(x,y)-B_S(x,y) \quad (7)$$

$$\begin{aligned}\Delta H=\min(&(360-(|I_H(x,y)-B_H(x,y)|)),\\ &(|I_H(x,y)-B_H(x,y)|))\end{aligned} \quad (8)$$

In the formula, α and β are thresholds of V, T_S is the threshold of S, T_H is the threshold of H, $I_{H,S,V}(x,y)$ is the value of HSV of current frame at pixel *(x,y)*, $B_{H,S,V}(x,y)$ is the value of HSV of background at pixel *(x,y)*. α is inversely proportional to the intensity of light of current environment and it should be greater than zero in order to avoid considering the foreground as shadow which color is dark. Because the brightness of pixel of the shadow is less than the corresponding point of non-shadow, β is less than one. As shadow will reduce the saturation of pixels, so the difference of saturation of shadow and background is often negative, then T_S is less than zero. And consider the difference of Hue of shadow and background is relatively small, T_H is a small value.

Shadow elimination algorithm based on HSV space needs to convert RGB space to HSV space. The cost of conversion is too much, and then affects the real-time of detection. Therefore, this paper presents a fast elimination method of neighbor mean based on HSV model to decrease the processing time of shadow elimination. Its basic unit of calculation is a small window such as 3×3 unit. As the pixels in their neighborhood space are similarity, calculate the mean of 8-neighborhood of a pixel and as an input of elimination algorithm. For each move of this small window, the algorithm calculates only once. This method can eliminate the shadow effectively and does not lose priori premise.

Taking into account the existence of half-shadow, the edge of shadow is impossible to eliminate completely. Therefore, this paper start the shadow elimination algorithm at the center of detected object and the direction of processing is from center to edge. When get close to the edge, and then scan the points processed before. If these points are classed as shadow then the edge should be classed as shadow too.

D. Occlusion Identification

Vehicles blocking is inevitable because of the camera's installation of angle and height. If there is no corresponding process for it, the system will consider blocking vehicles as an object, and then the detection rate is reduced. Aim at this problem, a method of overlapped vehicles identification based on Kalman filter is proposed to identify overlapped vehicles. In order to segment overlapped vehicles, this paper proposes a search algorithm of template matching based on hierarchical pyramid.

Kalman filter is a recursive filter for linear minimum variance estimation of dynamic system of state sequence. It estimates the state of dynamic system based on the received measurement data containing noise. As the movement of vehicles on the highway can be approximated as straight line uniform motion, Kalman filter can apply to this system. Kalman filter contains two equations:

$$\text{State equation: } Z(t) = P(t)Z(t-1) + \omega(t-1) \quad (9)$$

$$\text{Observation equation: } g(t) = C(t)x(t) + \nu(t) \quad (10)$$

where $P(t)$ is a state transition matrix of n×n dimension at time t, $C(t)$ is a measurement matrix of m×n dimension at time t, $Z(t)$ is a n-dimensional state vector at time t, $g(t)$ is a m-dimensional observation vector at time t, $\omega(t)$ is a n-dimensional vector of process noise at time t, $v(t)$ is a m-dimensional vector of measurement noise at time t. State vector and observation vector of Kalman filter used in this paper are defined as follows:

$$Z(t) = (LT(t), LB(t), SLT(t), SLB(t), H(t), \\ V(t), SH(t), SV(t))^T \quad (11)$$

$$g(t) = (LT(t), LB(t), H(t), V(t))^T \quad (12)$$

Thereinto, $LT(t)$ and $LB(t)$ are left vertex and right vertex of hexagon at time t, $H(t)$ and $V(t)$ are horizontal edge and vertical edge of hexagon at time t, $SLT(t)$ and $LB(t)$ are moving speeds of each direction of corresponding vectors, $SH(d)$ and $SV(d)$ are speeds of changes of each direction of corresponding vectors. The Observation matrix as follows:

$$C(t) = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Bmatrix} \quad (13)$$

As this system processes 25 frames per second, so the state transition matrix for this system is

$$P(t) = \begin{Bmatrix} 1 & 0 & 0.04 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0.04 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0.04 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0.04 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{Bmatrix} \quad (14)$$

The method of overlapped vehicles identification we proposed: First of all, Kalman filter is utilized to predict the next frame position of vehicles and get the center points of them. Then, compare these predicted center points to the regions of current frame of vehicles. We believe that blocking is happen if there are two or more predicted center points falls on the same region.

E. Real-time Segmentation

The traditional algorithm of template matching is widely used in vehicles match, but it is vulnerable to noise and needs large computation. Therefore, this paper presents a search algorithm of template matching based on hierarchical pyramid as follows:

[1]The initial position for matching is the predicted position of vehicles get from Kalman filter.

[2]The regions of template and object are represented by pyramid image: For each 2×2 neighborhood of the regions of template and object, calculate the average of them to get two low-resolution images. And then do the same to the low-resolution images recursively. Finally, two set of pyramid images of regions of template and matching can be got.

[3]First, perform matching operation on the minimum resolution images. Define the similar measure of two images as follows:

$$E = \sum_{i=1}^{P} (D_R(f_{tem}(x,y), f_{obj}(x+\Delta x, y+\Delta y)) \,\&\, D_G(f_{tem}(x,y), f_{obj}(x+\Delta x, y+\Delta y)) \,\&\, D_B(f_{tem}(x,y), f_{obj}(x+\Delta x, y+\Delta y))) \quad (15)$$

$$D_{C=R.G.B}(f_{tem}(x,y), f_{obj}(x+\Delta x, y+\Delta y)) = \begin{cases} 1 & abs(f_{tem}(x,y) - f_{obj}(x+\Delta x, y+\Delta y)) \le T_{match} \\ 0 & abs(f_{tem}(x,y) - f_{obj}(x+\Delta x, y+\Delta y)) > T_{match} \end{cases} \quad (16)$$

where E is similar measure, $f_{tem}(x,y)$ is a pixel of template, $f_{obj}(x,y)$ is a pixel of object, T_{match} is a user-define threshold. Increase E by 1 if all the differences between RGB are less than or equal T_{match}. Finally, if E is greater than a certain threshold then go to the higher resolution images and do the same operation. If math completely in the highest-resolution images then we can get the segmentation of the vehicles. Otherwise, go to [4].

[4]The region of object should be moved a few pixels to get a new region, and then go to [2].

III. Experiment Results

To verify the effectiveness of the algorithm, we use four videos to test vehicle detection method our proposed on the same computer platform (CPU: P4 2.2G, memory: 512M). The first two videos were recorded in Guangyuan highway of Guangzhou, and the others were recorded in Guangqing expressway. The performances of different videos are displayed as follow:

Figure 2. The test on video 1.

Figure 3. The test on video 2.

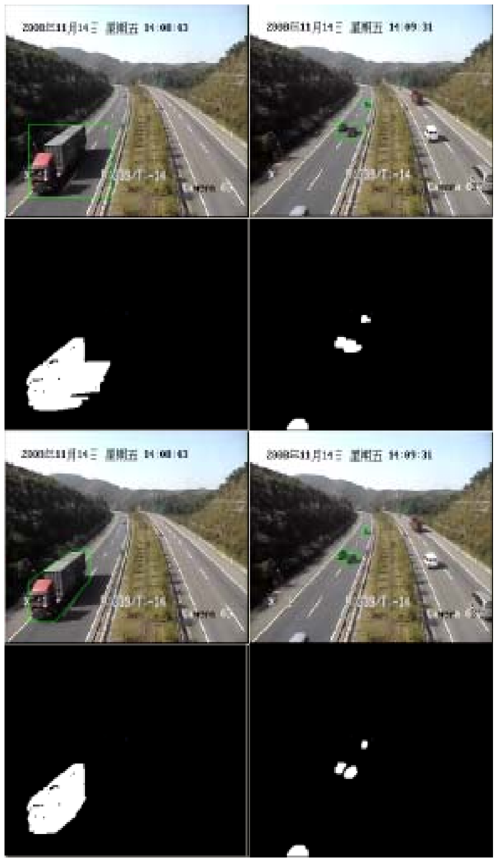

Figure 4. The test on video 3.

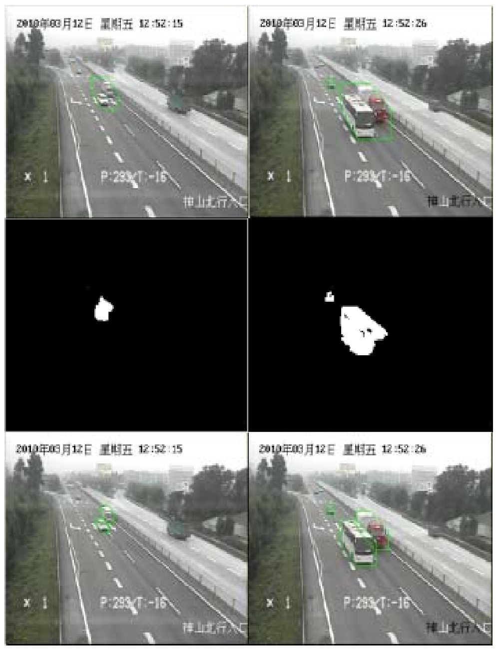

Figure 5. The test on video 4

As shown above, the method our proposed is able to eliminate shadow effectively and has good noise immunity. It can identify and handle the occurrence of occlusion.

IV. CONCLUSION

In this paper, an adaptive methods based on statistical model is used to establish and update the background. According to the direction of drive line bias, make use of hexagon to denote detected vehicles. In order to improve the efficiency and achieve better results, a fast elimination method of neighbor mean based on HSV model is proposed to eliminate shadow. As the occlusion is inevitable, this paper use a method of identification based on Kalman filter for occlusion identification, and segment the blocking vehicles using a search algorithm of template matching based on hierarchical pyramid. Experimental results have shown that the algorithm our proposed can execute fast and improve the detection rate of vehicles. The traffic incident detection system based on this paper has applied for Guangqing expressway in Guangdong Province.

REFERENCES

[1] Chen Y T, Chen C S, Huang C R, Hung Y P, "Efficient hierarchical method for background subtraction," Pattern Recognition, 2007, 40(10): 2706-2715.

[2] S. Zehang, G. Bebis, and R. Miller, "On-road vehicle detection: a review, " Pattern Analysis and Machine Intelligence, IEEE Transactions, 2006, vol. 28, pp. 694-711.

[3] A. Francois and G. Medioni, "Adaptive Color Background Modeling for Real-Time Segmentation of Video Streams," Proc. of International on Imaging Science, System, and Technology, 1999: 227-232.

[4] R. Cucchiara, C. Grana, M. Piccardi, and A.Prati, "Statistical and knowledge-based moving object detection in traffic scene," in Proceedings of IEEE Int'L Conference on Intelligent Transportation Systems, Oct. 2000, pp. 27-32.

[5] W. Liu, X. Wen, B. Duan, H. Yuan, and N. Wang, "Rear Vehicle Detection and Tracking for Lane Change Assist, " in Intelligent Vehicles Symposium, 2007, pp. 252-257.

[6] W. Lu, H. F. Wang, and Q. Z. Wang, "A Synchronous Detection of the Road Boundary and Lane Marking for Intelligent Vehicles, " in Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2007. SNPD 2007. Eighth ACIS International Conference, 2007, pp. 741-745.

[7] Zivkovic Z, van der Heijden F, "Efficient adaptive density estimation per image pixel for the task of background subtraction, " Pattern Recognition Letters, 2006, 27(7): 773-780.

[8] Weiming Hu,Xuejuan Xiao,Dan Xie.Traffic, "accident prediction using 3-D model-based vehicle tracking," in IEEE Transactions on Vehicular Technology, 2004, 53(3): 677-694.

[9] J. Stauder, R. Mech, and J. Ostermann, "Detection of moving cast shadows for object segmentation," IEEE Transactionson Multimedia, Mar. 1999, vol. 1, no. 1, pp. 65–76.

[10] C.Jiang and M.O.Ward, "Shadow identification," Proceedings of IEEE Int'l Conference on Computer Vision and Pattern Recognition, 1992, pp. 606-612.

[11] Sohail Nadimi, Bir Bhanu, "Physical Models for Moving Shadow and Object Detection in Video," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(8): 1079-1087.

[12] I. Mikic, P. Cosman, G. Kogut, and M.M. Trivedi, "Moving shadow and object detection in traffic scenes," in Proceedings of Int'l Conference on Pattern Recognition, Sept. 2000.

[13] J I Xiaopeng, WEI Zhiqiang, "Effective Vehicle Detection Technique for Traffic Surveillance Systems," in Journal of Visual Communication and Image Rep resentation, 2006(17): 647-658.

[14] L I Peihua, ZHANG Tianwen, MA Bo, "Unscented Kalman Filter for Visual Cure Tracking," Realtime Image and Vision Computing, 2004, 22(2): 157.

[15] RATH I Y, VASWAN IN, TANNENBAUM A, et al, "Tracking Deforming Objects using Particle Filtering for Geometric Active Contours," in IEEE Trans on Pattern Analysis and Machine Intelligence, 2007, 29(8): 1470-1475.