World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:4, No:6, 2010

A Video-based Algorithm for Moving Objects Detection at Signalized Intersection

Juan Li, Chunfu Shao, Chunjiao Dong, Dan Zhao, Yinhong Liu

Open Science Index, Computer and Information Engineering Vol:4, No:6, 2010 publications.waset.org/11762.pdf

***Abstract*—**Mixed-traffic (e.g., pedestrians, bicycles, and vehicles) data at an intersection is one of the essential factors for intersection design and traffic control. However, some data such as pedestrian volume cannot be directly collected by common detectors (e.g. inductive loop, sonar and microwave sensors). In this paper, a video based detection algorithm is proposed for mixed-traffic data collection at intersections using surveillance cameras. The algorithm is derived from Gaussian Mixture Model (GMM), and uses a mergence time adjustment scheme to improve the traditional algorithm. Real-world video data were selected to test the algorithm. The results show that the proposed algorithm has the faster processing speed and more accuracy than the traditional algorithm. This indicates that the improved algorithm can be applied to detect mixed-traffic at signalized intersection, even when conflicts occur.

***Keywords*—**detection, intersection, mixed traffic, moving objects.

I. INTRODUCTION

INTERSECTIONS are crucial to road networks, because the efficiency of intersections would strongly influence the performance of the whole traffic system. The data collected from an intersection have important practical significance for urban traffic management and control [1]–[3]. By using such data, traffic engineers can analyze the movements of traffic participants (e.g., pedestrians, bicycles and vehicles), calculate the time duration of various influences, and quantify the impacts of each transport mode has on the capacity of an intersection [4]. Such work is very helpful for better designing intersections and traffic signal control.

For the purpose of measuring traffic data of an intersection, some traffic sensors, such as loop, sonar, and microwave, have been applied in practice. However, since pedestrians and bicycles usually have weak signatures and are less restricted by lane discipline [5]–[7], they cannot be directly measured by these sensors. This problem becomes more serious in developing countries where mixed traffic plays a leading role in urban transportation. To overcome these limitations, two techniques, hand counting and video camera, are frequently used to collect both pedestrian and bicycle data [8]. For reasons of the labor intensity and narrow monitoring capacity of the hand counting method, the video camera technique has been increasingly deployed at signalized intersections [9]. From the data collected by a video camera, information of pedestrians, bicycles and vehicles, such as position, velocity, density and flow can be retrieved. In the case of monitoring the mixed traffic at an intersection, the video method faces the problem of detection failure due to traffic conflicts. Little research has been done concerning this aspect.

This paper focuses on improving the image processing algorithm to clearly capture moving objects at a signalized intersection. A new detection algorithm is proposed by improving traditional GMM with a mergence time adjustment scheme. The previous studies are briefly reviewed before presenting the details of the proposed algorithm. The improved algorithm is composed of five steps, which are model initialization, parameter updating, background estimation, foreground segmentation and shadow removal. To evaluate the validity of the proposed algorithm, real-world traffic videos are used for test. The results indicate that the detection performances are improved comparing to the traditional algorithm. It is expected that the proposed algorithm is able to detect mixed traffic flows at **s**ignalized **i**ntersection effectively.

II. LITERATURE REVIEW

A. Previous Works

Intersection traffic data collection by using the image processing technique has been thoroughly developed in the last decade. Video cameras are used to monitor a scene to acquire moving objects, so a reliable and flexible background model is very important [10]–[12].

At present, many algorithms are used to subtract moving objects, such as background subtraction, optical flow, frame difference, *etc*. Among them background subtraction has become a more common approach, due to its high computational speed and good adaptability. In this method, the moving objects are acquired by building a representation of the scene background and comparing new frames with this representation. Many background subtraction methods are introduced [13]–[15], among which GMM is the most popular one. The reason is that the background subtracts by GMM can adapt to illumination change and background disturbance, using

J. Li is with MOE Key Lab. for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing, 100044 CHINA (phone: 86-10-51683665; fax: 86-10-51688344; e-mail: juanli_bjtu@ yahoo.com.cn)

C. F. Shao, C. J. Dong, D. Zhao and Y. H. Liu are with the MOE Key Lab. for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing, 100044 CHINA. (e-mail: cfshao@center.bjtu.edu.cn, 07114181@bjtu.edu.cn, zhaodans@yahoo.com, liuyinhong2008@hotmail.com).

Open Science Index, Computer and Information Engineering Vol:4, No:6, 2010 publications.waset.org/11762.pdf

priority and probability of each pixel to judge whether it belongs to background or foreground [16]–[22]. In this method, each pixel is modeled by a mixture of *K* Gaussian distributions. The expectation-maximization (EM) algorithm based on recursive filter is used to train GMM. This model is able to adapt to multiple background objects and efficiently estimate the temporal background. One limitation of the model is that it cannot detect objects which stop for a while after moving.

B. Limitations of the Traditional Algorithm

Traffic conflicts are most likely to occur at signalized intersection. In order to safely cross the intersection, traffic participants usually decelerate or even stop to find a sufficient gap. In this case, the traffic participants cannot be detected correctly by traditional algorithm. Fig 1 and Fig 2 show the detection results of traditional GMM algorithm. They describe situations where traffic participants going through a T-shaped signalized intersection are not recognized by the traditional algorithm. The solid and dash lines represent the moving direction of vehicles and pedestrians flow respectively.

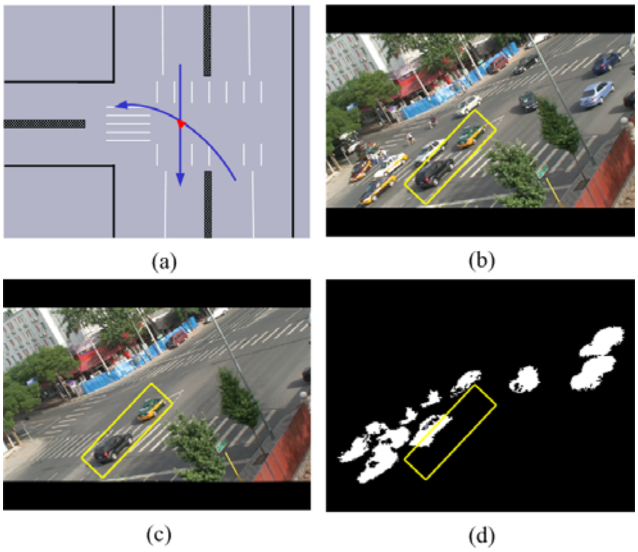


Fig. 1 Vehicle-vehicle conflict detected by traditional algorithm: (a) conceptual drawing; (b) original image; (c) subtracted background; and (d) extracted moving objects

Fig 1 illustrates a case of the vehicle-vehicle conflict. Specifically (a) shows the T-intersection, where a left-turning vehicle was encountering with the opposing vehicles, and the traffic conflict point is marked by the red triangle, (b) indicates the original conflict scene, where left-turning vehicles in yellow rectangular box were waiting to cross the intersection, (c) is the background image which subtracted by traditional algorithm, i.e., stopped vehicles inside the yellow box are merged into the background image, (d) is the detection results by traditional algorithm. It is difficult to find the stopped vehicles in the foreground.

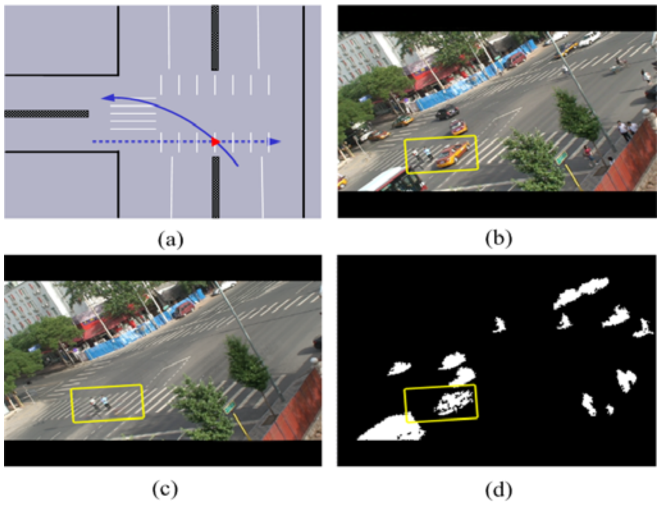


Fig. 2 Pedestrian-vehicle conflict detected by traditional algorithm: (a) conceptual drawing; (b) original image; (c) subtracted background; and (d) extracted moving objects

Fig 2 shows a pedestrian-vehicle conflict. In Fig 2(a), the conceptual drawing is given for the expression of the pedestrian-vehicle conflict on the crosswalk. Pedestrians stopped on the crosswalk and waited for a sufficient gap as shown in (b). The waiting time was so long that the pedestrians completely merged into the background as shown in (c) and could not be extracted as foreground in (d).

So it is difficult to detect the waiting participants stopped by traffic conflicts at signalized intersections. To solve this problem, an improved algorithm is proposed in this paper, the details of this algorithm are described in the following section.

III. An Improved Detection Algorithm

The video-based moving objects detection system is comprised of two modules: video flow capture module and motion detection module. The proposed algorithm is employed in the motion detection module. In order to detect the moving objects at signalized intersection, we have improved the traditional algorithm. Our detection algorithm contains 5 steps. Step 1: the first frame is utilized to initialize the Gaussian distributions. Step 2: we use a moving-based model to update the parameters at region level. Step 3 is to determine which Gaussian distribution is most likely to be incorporated into the background model. Step 4 is foreground segmentation and Step 5 is shadow detection. Morphological reconstruction is used to refine the foreground image after shadow removal. Fig 3 shows a simplified flow chart of the algorithm.

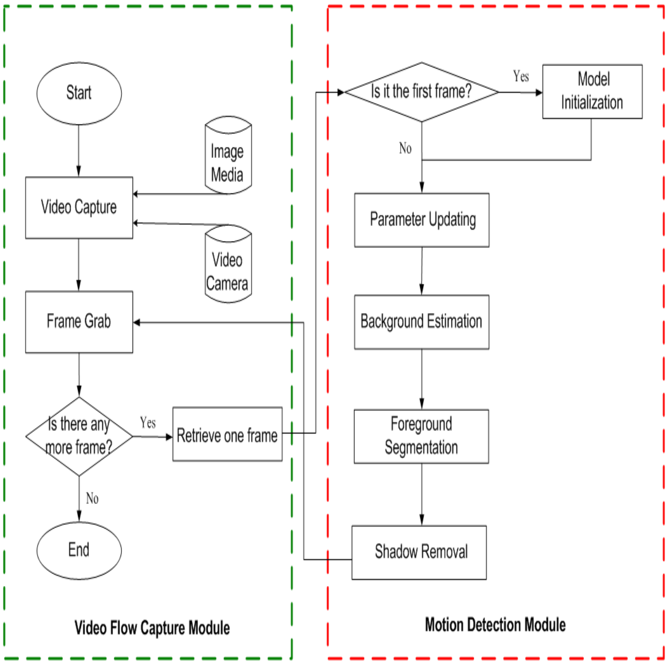


Fig. 3 Flow chart of the system

A. Model Initialization

For the proposed algorithm, it is important to obtain an initial model. If the coming frame is the first frame, the current pixels are used to initialize the model. Each background pixel is modeled by a mixture of K Gaussian distributions. In our study, K is set to 5. Different distributions represent different colors and the weight parameters of the mixture, which indicate the time proportions that those colors stay in the scene. The weights and means are set to 0 in initialization.

B. Parameter Updating

In order to adapt to the changes of illumination, the parameters of the model should be updated with the coming of the new frame. However, the traditional background updating formulations are pixel-based, and the correlation information between pixels is not considered. Pixel-leveled moving object detection is sensitive to noise and unable to monitor stationary objects effectively, whereas region-based background modeling can avoid these problems and is more robust for background subtraction [23]. So the proposed algorithm uses a moving-based model updating approach at region level. This approach contains two stages.

The first stage is to segment the original frame into moving region and static region. Segmentation technology such as morphological method is used to search the moving objects' contours in the foreground. Then, pixels in the original frame are sorted by such contours and the classification results are represented by a motion detection mask M_s, which is defined as:

$$M_S = \begin{cases} 1 & I(x,y) \in Moving \quad Region \\ 0 & I(x,y) \in Static \quad Region \end{cases} \tag{1}$$

The value of this motion detection mask M_s determines whether to update the parameters of the K Gaussian distributions or not. Fig 4 shows an example of graph segmentation, in which pixel $P1$ is in the moving region, so M_s is 1 and the corresponding parameters remain unchanged. On the contrary, pixel $P2$ is out of the moving region and its corresponding model should be updated.

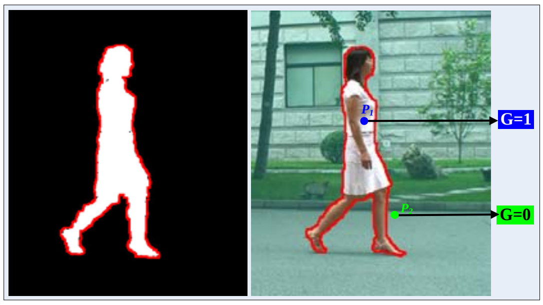


Fig. 4 Graph segmentation of original frame

The second step is updating parameters of the static region. Every new pixel value is checked against the existing K Gaussian distributions until a match is found. The parameters of Gaussian distribution that matches X_t are updated as:

$$\begin{cases} \omega_{k,t} = (1-\alpha)\omega_{k,t-1} + \alpha(M_{k,t}) \\ \mu_{k,t} = (1-\rho)\mu_{k,t-1} + \rho X_t \\ \sum_{k,t} = (1-\rho)\sum_{k,t-1} + \rho(X_t - \mu_{k,t})^T (X_t - \mu_{k,t}) \\ \rho = \alpha\eta(X_t, \mu_{k,t}, \sum_{k,t}) \end{cases} \tag{2}$$

where α is the learning rate and $M_{k,t}$ is 1 for the model which matched and 0 for the remaining model, ρ is the learning rate for parameters. If none of the K distributions matches the current pixel value X_t, the least probable distribution is replaced by a new Gaussian distribution with the current value as its mean value, and an initially large variance, and a small prior weight, which are defined as follows:

$$\begin{cases} j = \arg\min_i \{\omega_{k,t-1}\} \\ \omega_{j,t} = \omega_0 \\ \mu_{j,t} = X_t \\ \sum_{j,t} = \sum_0 \end{cases} \tag{3}$$

C. Background Estimation

It is time to determine which Gaussian distribution is most likely to be incorporated into the background model after model updating. The K distributions are ordered based on the fitness value and the first ***B*** distributions are chosen as a model of the background of the scene. B is estimated as:

$$B = \arg\min_b (\sum_{k=1}^{b} \omega_k > T) \tag{4}$$

where T is the threshold of the background model, which represents the minimum prior probability of the background in

Open Science Index, Computer and Information Engineering Vol:4, No:6, 2010 publications.waset.org/11762.pdf

Open Science Index, Computer and Information Engineering Vol:4, No:6, 2010 publications.waset.org/11762.pdf

the scene. At traffic conflict points, long-term stopped traffic participants would merge into the background, which results in missing targets. Setting a small value for the learning rate will hold the stationary objects as foreground for a longer time. However, updating rate will become slower in this situation. Therefore, we add the foreground counter C into the background estimation model, which records the times of each pixel consecutively classified as foreground. Merging coefficient is determined by the frequency of foreground occurrences in history frames.

$$\beta = \frac{C}{M} \tag{5}$$

where M is a time constant, measured in number of frames, which indicates the expected foreground mergence time. The weight of the candidate Gaussian $\omega_{v,t}$ which is most likely to become background mode is adjusted as follows:

$$\omega_{v,t} = \omega_{v,t} \times \beta \tag{6}$$

where v is the sequence number of the distribution, which is most likely to incorporate into the first B Gaussian distributions.

D. Foreground Segmentation

The procedure of foreground segmentation contains two different judgment criterions. The first stage is to check the pixel value X_t against the first B Gaussian distributions in order of priorities. If the pixel value is no larger than 2.5 times standard deviations of any of the B distributions, a match is obtained.

The second stage is to examine the average weight which is defined as:

$$\overline{\omega_{k,t}} = (1-\lambda)\omega_{k,t-1} + \lambda(M_{k,t}) \tag{7}$$

where $\overline{\omega_{k,t}}$ means the average value calculated from the weight of each Gaussian distribution; λ means learn rate, which is a small fixed value. $\overline{\omega_{k,t}}$ represents the appearance frequency of that Gaussian distribution in the window vision whose length is $L' = 1/\lambda$. Inequality $\overline{\omega_{k,t}} > Th$ (Th is a fixed value, we set 0.6 in this paper) means that Gaussian distributions' appearance frequency in the window vision is high, and the accumulation performance of this model is good.

If the pixel value X_t can not meet either of the two conditions, this point is considered as foreground. Otherwise it belongs to background.

E. Shadow Removal

After foreground segmentation, a binary image of foreground F is obtained. Then, HSV color information is used for shadow detection. As noted by Cucchiara's study [24], a shadow point does not change its hue significantly and its saturation is often lower than background points, so the decision process is reported as follows:

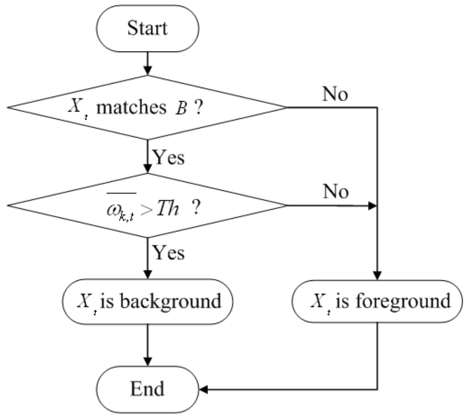


Fig. 5 Steps in foreground segmentation algorithm

$$SP(x,y) = \begin{cases} 1 & if \quad \alpha \le \dfrac{I^V(x,y)}{B^V(x,y)} \le \beta \\ & \wedge (I^S(x,y) - B^S(x,y)) \le \tau_S \\ & \wedge \left| I^H(x,y) - B^H(x,y) \right| \le \tau_H \\ 0 & otherwise \end{cases} \tag{8}$$

where $I(x,y)$ and $B(x,y)$ are the pixel values at coordinate in the original image and in the background model respectively. However, some foreground pixels related to moving objects may be classified as shadow pixels by mistake. Therefore, it is necessary to refine the shadow detection results.

To ensure the integrity of the foreground, the improved algorithm appends shadow refinement after traditional shadow detection. Hence, the shadow removal process contains three stages. The first stage is filter. To remove isolated pixels and fill gaps, an opening operator is employed to the original foreground F, and the processed binary image $F1$ is retrieved. The second stage is shadow detection. As mentioned above, the moving cast shadows are removed by HSV color space conversion and the shadow judging function, and then we get the foreground image $F2$. The third stage is shadow refinement. Morphological reconstruction [25] is utilized to recover the points misclassified as shadow pixels in image $F2$.

$$Fg = F1 \cap (F2 \oplus B) \tag{9}$$

where Fg is the final foreground image without shadow; $F1$ is the mask image; $F2$ is the marker image; B is a binary structuring element; $\oplus$ represents the dilation operation.

This shadow removal algorithm produces an accurate and robust foreground image without shadow. The example of shadow removal process is shown in Fig 6, which demonstrates the performance of this algorithm.

Open Science Index, Computer and Information Engineering Vol:4, No:6, 2010 publications.waset.org/11762.pdf

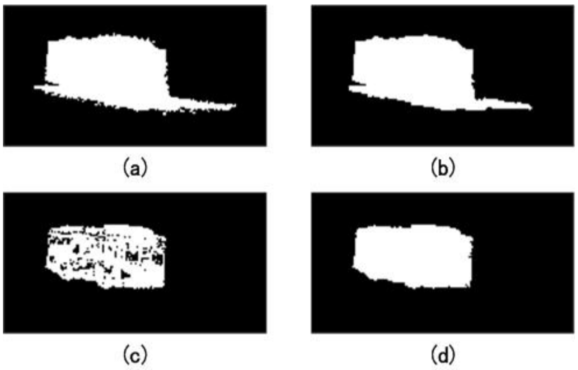
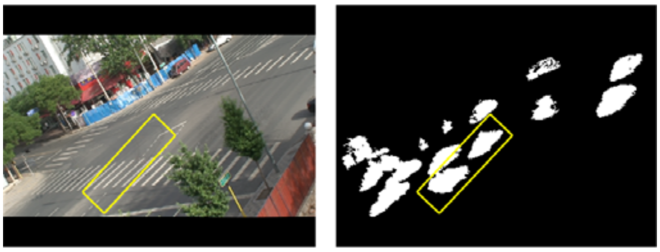


Fig. 6 Step by step illustration of the shadow removal process: (a) original foreground image *F*; (b) foreground image after filter *F1*; (c) foreground image after shadow detection *F2*; and (d) foreground image after shadow refinement *Fg*

IV. Experimental Test and Results

The proposed algorithm was tested by a real-world traffic video, which was taken at a T-intersection, as shown in Fig 1 and Fig 2. This test was set to see whether the algorithm could accurately detect objects at signalized intersection, when conflicts occur.



(a) (b)

Fig. 7 Vehicle-vehicle traffic conflict detected by improved algorithm: (a) subtracted background; and (b) extracted moving objects

(a) (b)

Fig. 8 Pedestrian-vehicle traffic conflict detected by improved algorithm: (a) subtracted background; and (b) extracted moving objects

The detection results are shown in Fig 7 and Fig 8: (a) presents the subtracted background and (b) presents the motion detection results. A traffic conflict point is shown in a yellow rectangle. From the results, we can see that stopped pedestrians/vehicles are clearly detected as foreground compared to those in Fig 1 and Fig 2.

Table 1 shows the computation results of the traditional algorithm and improved algorithm. It's indicated that the false

TABLE I
Comparison of the Test Results

Algorithm l	False Extracted Pixel per Frame (pixel)	Time per Frame(ms)
Traditional Algorithm	64	22
Improved Algorithm	8	23

extracted pixels by using the improved algorithm are 8 times smaller than those by the traditional algorithm, but the computing time is almost equivalent. Therefore, the improved algorithm can produces the better segmentation results for object detection than the traditional algorithm, especially when conflicts occur.

V. Conclusions and Discussions

A video-based algorithm for moving objects detection at signalized intersection is proposed in this paper, which can overcome the difficulties of detecting long term stopped objects at intersections due to traffic conflicts. In this algorithm, the moving-based model updating scheme decreases the calculation and enhances the segmentation quality. The background estimating model controls the foreground mergence time effectively by reducing the weight of the corresponding Gaussian distribution. The foreground segmentation integrated with average weight enhances the foreground segmentation performance, which effectively decreases the error rate of taking moving objects as background. The results of the tests indicate that, the improved algorithm has the better adaptability, and can enhance the stability and accuracy of foreground segmentation. This algorithm can be used to detect moving objects at signalized intersection, even when traffic conflicts occur.

Acknowledgment

This study is supported by the National Nature Science Foundation of the People's Republic of China (No. 50778015) and National Basic Research Program of China (973 Program, 2006CB705500).

References

[1] Y. Malinovskiy, Y. J. Wu, Y. H. Wang, and C. Lim. Video-Based Monitoring of Pedestrian Movements at Signalized Intersections. In Transportation Research Record: Journal of the Transportation Research Board, No. 2073, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 11-17.

[2] H. J. Hwang, U. K. Lee, K. R. Baek, Y. J. Wu, and Y. H. Wang. Measuring Signalized Intersection Performances in Real-Time with Traffic Sensors, The 88th Annual Meeting of the Transportation Research Board (CD-ROM), Washington, D.C., 2009 .

[3] J. Y. Zheng, X. L. Ma; Y. H. Wang, and P. Li. Measuring Signalized Intersection Performances in Real-Time with Traffic Sensors, The 88th Annual Meeting of the Transportation Research Board (CD-ROM), Washington, D.C., 2009.

[4] Y. L. Su, D. Y. Yao, Y. Zhang, L. Li, and Z. Wei. Vehicle/Pedestrian Conflict Analysis and Exclusive Right-Turn Phase Setting Study, In Proceedings. IEEE Intelligent Vehicles Symposium, The Netherlands, 2008.

[5] D. A. Noyce, A. Gajendran, and R. Dharmaraju. Development of Bicycle and Pedestrian Detection and Classification Algorithm for

Active-Infrared Overhead Vehicle Imaging Sensors, In Transportation Research Record: Journal of the Transportation Research Board, No. 1982, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 202-209.

[6] H. Y. Hu, Z. W. Qu, Z. H. Li, and D. H. Wang. Research on Recognition and Classification of Moving Objects in Mixed Traffic Based on Video Detection, The 87th Annual Meeting of the Transportation Research Board (CD-ROM), Washington, D.C., 2008.

[7] J. Heikkila, and O. Silve´n. A Real-time System for Monitoring of Cyclists and Pedestrians, Alberta, Canada. Image and Vision Computing, Vol. 22, No. 7, 2004, pp. 563-570.

[8] K. A. Ismail, T. Sayed, N. Saunier, and C. Lim. Automated Analysis of Pedestrian-Vehicle Conflicts Using Video Data. The 88th Annual Meeting of the Transportation Research Board (CD-ROM), Washington, D.C., 2009.

[9] W. D. Cottrell, and D. Pal. Evaluation of Pedestrian Data Needs and Collection Efforts. In Transportation Research Record: Journal of the Transportation Research Board, No. 1828, Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 12-19.

[10] J. Y. Zheng, Y. H. Wang, N. L. Nihan, and M. E. Hallenbeck. Extracting Roadway Background Image: Mode-Based Approach. In Transportation Research Record: Journal of the Transportation Research Board, No. 1994, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp 82-88.

[11] R. P. Avery, G. H. Zhang, Y. H. Wang, and N. L. Nihan. An Investigation into Shadow Removal from Traffic Images, In Transportation Research Record Journal of the Transportation Research Board, No. 2000, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 70-77.

[12] G. H. Zhang, R. P. Avery, and Y. H. Wang. A Video-Based Vehicle Detection and Classification System for Real-Time Traffic Data Collection Using Uncalibrated Video Cameras, In Transportation Research Record: Journal of the Transportation Research Board, No. 1993, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 138-147.

[13] W. M. Hu, T, N, Tan, L. Wang, and S. Maybank. A Survey on Visual Surveillance of Object Motion and Behaviors, IEEE Trans. on System, Man, and Cybernetics Part C: Applications and Review, Vol. 34, No. 3, 2004, pp. 334–351.

[14] Y. Malinovskiy, Y. J. Wu, Y. H. Wang, and N. Saunier. Video-Based Vehicle Detection and Tracking Using Spatio-Temporal Maps. The 88th Annual Meeting of the Transportation Research Board (CD-ROM), Washington, D.C., 2009.

[15] N. J. Kanhere, S. T. Birchfield, W. A. Sarasua, and T. C. Whitney. Real-Time Detection and Tracking of Vehicle Base Fronts for Measuring Traffic Counts and Speeds on Highways, In Transportation Research Record: Journal of the Transportation Research Board, No. 1982, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 202-209.

[16] W. E. L. Grimson, C. Stauffer, R. Romano, and L. Lee. Using Adaptive Tracking to Classify and Monitor Activities in a Site, In Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Santa Barbara, CA, 1998, pp. 22-29.

[17] C. Stauffer, and W. E. L. Grimson. Adaptive Background Mixture Models for Real-time Tracking, In Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Fort Collins, CO, 1999, pp. 246-252.

[18] C. Stauffer, and W. E. L. Grimson. Learning Patterns of Activity Using Real-time Tracking, In Proceedings. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, pp. 747-757.

[19] P. KaewTraKulPong, and R. Bowden. An Improved Adaptive Background Mixture Model for Real-Time Tracking with Shadow Detection, In Proceedings. 2nd European Workshop on Advanced Video Based Surveillance Systems, 2001, pp. 149-158.

[20] D. S. Lee, J. J. Hull, and B. Erol. A Bayesian Framework for Gaussian Mixture Background Modeling, In Proceedings. International Conference on Image Processing, 2003, pp. 973-976.

[21] D. S. Lee. Effective Gaussian Mixture Learning for Video Background Subtraction, In Proceedings. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, pp. 827-832.

[22] Y. C. Zhang, Z. Z. Liang, Z. G Hou, H. M. Wang, and M. Tan. An Adaptive Mixture Gaussian Background Model with Online Background Reconstruction and Adjustable Foreground Mergence Time for Motion Segmentation, In Proceedings. IEEE International Conference on Industrial Technology, 2005, pp. 23-27.

[23] K. Toyama, J. Krumm, B. Brumitt, and B. Meryers. Wallflower: Principles and Practice of Background Maintenance, In proceedings. International Conference on Computer Vision, Kerkyra, Greece, 1999, pp. 255-261.

[24] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti. Improving Shadow Suppression in Moving Object Detection with HSV Color Information. In Proceedings. IEEE Intelligent Transportation Systems Conference. Oakland, CA, 2001, pp. 334-339.

[25] R. C. Gonzalez, and R. E. Woods. Digital Image Processing. 2nd ed. Prentice-Hall, Inc., Englewood Cliffs, NJ, 2002.

Open Science Index, Computer and Information Engineering Vol:4, No:6, 2010 publications.waset.org/11762.pdf