A Vision-Based Road Surveillance System Using Improved Background Subtraction and Region Growing Approach

Zhang Jinglei[1,2]
[1]School of Electrical Engineering & Automation, Tianjin University, Tianjin 300072, China
[2]School of Electrical Engineering, Tianjin University of Technology, Tianjin 300191, China
zjlmailbox@126.com

Liu Zhengguang
School of Electrical Engineering & Automation, Tianjin University, Tianjin 300070, China
zh_liu@tju.edu.cn

Abstract

Monitoring traffic intersections in real time is an important part towards Intelligent Transportation System. A road surveillance system using background subtraction and threshold segmentation calculated traffic parameters. First an improved background establish and updating algorithm was applied to setup a robust background filtering illumination disturbance. Then a self-adaptive max variance threshold algorithm was adopted to determine the threshold. Third, a fast region growing algorithm was employed to border vehicles in some rectangles fixed their edges. Finally, algorithms were mentioned to calculate traffic parameters including traffic flux, average speed and duty ratio of road. The experimental results show that the system is adapted to monitor a multi-lane road near an urban intersection.

Keywords: *Image processing, road surveillance system, background subtraction, region growing.*

1. Introduction

Traffic management and vehicle information rely on a suite of sensors for estimating traffic parameters. Road surveillance system is an important part of Intelligent Transportation System. Many works on ITS aim at helping traffic flow management by providing information on how many vehicles are in the scene.

In recent years, the vision-based approach is promising since it requires no pavement adjustments and has more potential advantages such as a cheaper cost, larger detection areas, and it is also more flexible and suitable for vehicles identification. These vision-based systems allow the visualization of vehicles on the road by using a single camera (monocular vision) mounted in perspective view of the road segment that it is monitoring, thus enabling traffic-scene analysis, such as traffic-conditions assessment and travel-speed estimation, as well as queue-length measurement, in which traditional nonvisual surveillance systems could not do.

There are many researches related to the topic of vision-based traffic monitoring system [1]-[5]. The main topics of these researches are how to establish a reliable background and background segmentation. There are various approaches to this, with varying degrees of effectiveness. To be useful, the segmentation method needs to accurately separate vehicles from the background, be fast enough to operate in real time, be insensitive to lighting and weather conditions, and require a minimal amount of supplementary information. In [6], a segmentation approach using adaptive background subtraction is described. Though this method has the advantage that it adapts to changes in lighting and weather conditions, it needs to be initialized with an image of the background without any vehicles present. Another approach is time differencing, (used in [7]) which consists of subtracting consequent frames (or frames a fixed number apart). This method too is insensitive to lighting conditions and has the further advantage of not requiring initialization with a background image. However, this method produces many small blobs that are difficult to separate from noise [8].

Existing commercial image processing systems work well in free-flowing traffic, but these systems have difficulties with congestion, shadows and lighting transitions. We are developing a system for calculating road information under these challenging conditions. Our system uses a single camera mounted on a pole,

0-7695-2909-7/07 $25.00 © 2007 IEEE
DOI 10.1109/SNPD.2007.495

pedestrian bridge or other tall structure near the street, looking down on the traffic scene. It can be used for detecting vehicles and calculating road traffic parameters in multiple lanes. The system requires only the camera calibration parameters and lanes position information for initialization.

The whole system is illustrated in Figure.1, which uses a CCD camera mounted on a pedestrian bridge 300 meters near a cross. A totally general analysis of the system is beyond the scope of some texts, especially in camera calibration and tracking methods. At this point we will make no attempt to describe these workings in detail.

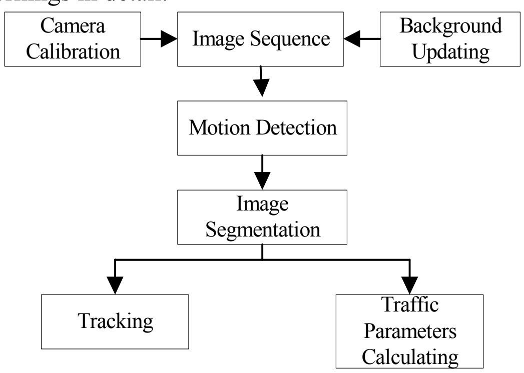


Fig. 1. Block diagram of the System.

2. Motion detection and image segmentation

The first stage of the system applies motion detection on the sequence of images to detect moving regions inside a limited window delimiting the area of interest in the images. Motion detection steps consist of background differentiation, segment filtering and an adaptive background update.

Our system requires inputting some information as initial parameters for improving the system's efficiency and preciseness. These parameters include lane lines, lanes number, top detection line and base line. Thus a multi-lane detection zone can be defined as in Fig.2.

2.1. Background establishment

First some pretreatment of images should be done. A 3 × 3 template based median filtering algorithm is applied to filtering the noise and to smoothen the image. Then the images are sharpened by Laplacian. Second an algorithm based on the integration of Temporal Difference and Background Subtraction is applied to determining the background.

A reliable background means to locate a scene without any vehicles in the detection zone. It need adapt the difference of the scene caused by the illumination changing. In Huang [5] proposed an algorithm to find an initial background. His idea was to find a minimum value from N frames at a position (x, y) in the detection zone, then find the INDEX_BASE and average all pixels belonging to the mode INDEX_BASE. Then an initial background pixel value at pixel (x, y) can be calculated. To circulate this process at every pixel can determine a background. This approach is high cost computation and may take 8.32 second as the author claim; however, the background updating only takes 0.005 second. In this paper, we consider a fact that the detection zone is always empty (no vehicle) or reverse corresponding the direct traffic lights. Thus Temporal Differencing method can be employed.

First we apply a Temporal Differencing between 3 frames, and then calculate the ratio of the changing pixels according to the whole difference image. If the ratio is lower than a preset threshold value, then set the original image as the background. The original background is upgraded periodically in order to filtering disturbance by the change of illumination.

If the ratio is higher than the threshold, we consider that some motion vehicles are passing through. A threshold segmentation of foreground is applied first, and then background subtraction algorithm can be used. This algorithm is adapted to detecting traffic flux near a traffic intersection.

2.2. Image segmentation

A Self-Adaptive Max Variance Threshold Method is employed to segment the foreground [10]. Suppose the image can be segment into two regions by a threshold of gray-level k. The pixels at gray-level between 0~k-1 and k~L-1 can be formed as region A and B. It is obviously that the probabilities of A and B can be defined as follows:

$$\omega_A = \sum_{i=0}^{k\text{-}1} p_i = \omega(k)\,, \tag{1}$$

$$\omega_B = \sum_{i=k}^{L\text{-}1} p_i = 1 - \omega(k)\,. \tag{2}$$

The average gray-level of region A and B is:

$$\mu_A = \frac{1}{\omega_A}\sum_{i=0}^{k\text{-}1} i p_i = \frac{\mu(k)}{\omega(k)}\,, \tag{3}$$

$$\mu_B = \frac{1}{\omega_B}\sum_{i=k}^{L\text{-}1} i p_i = \frac{\mu - \mu(k)}{1 - \omega(k)}\,, \tag{4}$$

where μ is the average grey level of the whole image and can be calculated as follows:

$$\mu = \sum_{i=0}^{L\text{-}1} ip_i = \sum_{i=0}^{k\text{-}1} ip_i + \sum_{i=k}^{L\text{-}1} ip_i = \omega_A \mu_A + \omega_B \mu_B \,. \qquad (5)$$

The total variance of the two regions is

$$\sigma^2(k) = \omega_A(\mu_A - \mu)^2 + \omega_B(\mu_B - \mu)^2 = \frac{[\mu\omega(k) - \mu(k)]^2}{\omega(k)[1-\omega(k)]} \,. \qquad (6)$$

Changing k from 0~L-1, the value that make maximum $\sigma^2(k)$ is the optimum threshold value of segmentation.

2.3. Vehicles location by region growing

After threshold segmentation, a fast region growing approach is used to locate vehicles. The algorithm can be stated in following steps:

Step1: Every lane's midpoint (which site at the middle of the base line) is selected as the reference point. Draw a radial line toward the top line. Define the points as $P\{p_0, p_1, p_2, \ldots p_n\}$ coordinate of every pixel value change from 0 to 1 and reverse.

Step2: Enlarge p_i to a rectangle R_i, growing all sides of the rectangle till the pixels values of every side are 0. Calculate the center of R_i. If the area of R_i is lower than a threshold, it is considered as noise points and will be discarded.

Step3: If the y-coordinate of p_{i+1} is lower than the upside coordinate of R_i, p_{i+1} is discarded, otherwise do Step2. Thus vehicles in the detection lane can be expressed as a rectangle. The center of the rectangle is the centroid of a detected vehicle can be used for tracking.

Step4: Circularly do Step2 and Step3 can locate the vehicles in single lane. By using same process in other lanes can locate the vehicles in detection zone.

The number of available rectangles is the number of vehicles that the system recognition.

3. Traffic parameters calculation

3.1. Road traffic flux

The flux of traffic flow f is defined as the number of vehicles that pass a given region in unit time and can be calculated by follow equation:

$$f = \frac{1}{T}\sum_{i=1}^{k} N \qquad (7)$$

where T is the time interval and k is the setting number of lane. N is the total detected vehicles number of every lane.

3.2. The duty ratio of road

This is a parameter show the density of traffic. In this system, it equals the sum of foreground objects area divided by the detection region area.

$$D = \frac{A}{A_{set}} \qquad (8)$$

where A_{set} is the area of detection region, A is the foreground objects area of every frame. (i.e. the number of pixels).

3.3. The average speed of vehicles

The centroid of a vehicle is used for tracking. For two consecutive frames, centroid of every vehicle is located. Centroid (i, j) on the frame M-1 will be matched with centroid (k, l) on the frame M if their distance is the shortest and $l \leq j$.

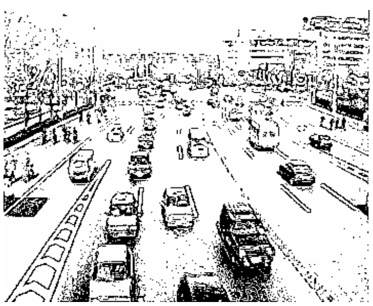
Then calculate the average displacement of centroids of lanes. By means of the ratio of the average displacement and the time between two frames, the average speed can be educed.

4. Experimental results

The experimental results are shown in Fig.2 to Fig.7. A detection zone should be defined first according lanes number and was shown in Fig.2. In the image pretreatment stage, the original image (Fig. 3) is sharpened and was shown in Fig. 4. By using the algorithm mentioned in section 2.2, an optimum threshold of segmentation can be calculated. Next background subtraction is employed. It should be noted that background is defined only within the detection zone. By this means we can easily ignore hash information beyond the detection zone. Fig 5 and 6 show the results of the segmentation. Then vehicles can be located from the foreground image (shown in Fig. 7).

The system was implemented on a PC with an AMD AthlonXP 2600+ CPU. We tested the system on image sequences of urban traffic scenes recorded by a CCD digital video camera. Background updating processed for every ten frames. We have been able to process 10 frames per second that is enough for city traffic scenes, in which vehicles usually move slower than 70Km/h. Our test result shows that over 90% of the vehicles were correctly detected.

Fig. 2. The detection zone. **Fig. 3. The original image.** **Fig. 4. Image pretreatment results.**



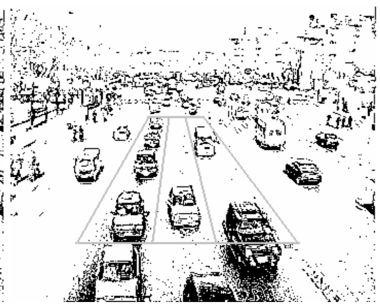


Fig. 5. Threshold segmentation results. Fig. 6. Background subtraction results. Fig. 7. Vehicles location results.

5. Conclusions

In this paper, we proposed a method of integration of Temporal Difference and Background Subtraction to give a better selection in resolving the influence of threshold segmentation caused by the illumination changing.

The method can obtain a nearly perfect dynamic background. Also we proposed a fast region growing approach to locate vehicles. The reliability of moving objects recognition is enhanced. In addition, the algorithms of calculating the traffic parameters give a reference to administration departments to estimate the transport ability of roads. In our future work, we will extend the system to resolve vehicles' overlap and occluding problems.

References

[1] S. Gupte, O. Masoud, "Detection and classification of vehicles", *IEEE Trans. on Intelligent Transportation Systems*, vol.3, no.1, pp.37-47, Mar. 2002.

[2] J. Lou, H. Yang, W. Hu, T. Tan, "An illumination invariant change detection algorithm," in *ACCV2002: The 5th Asian Conf. on Computer Vision*, Jan 2002, Melbourne, Australia.

[3] G. Gordon, T. Darrell, M. Harville, J. Woodfill, "Background estimation and removal based on range and color," in *Int. Conf. on Computer Vision and Pattern Recognition*, pp.459-464, 1999.

[4] C. C. C. Pang, W. W. L. Lam, and N. H. C. Yung, "A novel method for resolving vehicle occlusion in a monocular traffic-image sequence," *IEEE Trans. on Intelligent Transportation Systems*, vol. 5, no. 3, pp.129-141, Sep. 2004.

[5] M.C. Huang and S.H. Yen, "A real-time and color-based computer vision for traffic monitoring system," in *IEEE Int. Conf. on Multimedia and Expo(ICME)*, pp.2119-2122, 2004.

[6] K. Karmann and A. von Brandt, "Moving object recognition using an adaptive background memory," in "*Time-Varying Image Processing and Moving Object Recognition*," 2-edited by V. Capellini, 1990.

[7] A. Lipton, H. Fujiyoshi and R. Patil, "Moving target classification and tracking from real-time video," in *Proc. of the Image Understanding Workshop*, 1998.

[8] S. Gupte, O. Masoud, et al, "Vision-based vehicle classification," in *Proc. IEEE Conference on Intelligent Transportation Systems*, pp.46-51, 2000, Dearborn, USA.

[9] Y. J. Zhang, "*Image Processing and Analysis*," Tsinghua University Press, Beijing, China, 2002, pp.71-87.

[10] W. Y. Liu, "*Photoelectric Image Processing*," Publishing House of Electronics Industry, Beijing, China, 2002, pp.207–208.