論文2001-38SC-5-6

영상을 기반 교통 파라미터 추출에 관한 연구
(An Approach to Video Based Traffic Parameter Extraction)

郁 梅 *, 金 容 得 **

(Mei Yu and Yong-Deak Kim)

요 약

차량검출은 교통량 관측을 위해서 필요한 가장 기본적인 요소이다. 영상을 기반으로 한 교통정보 추출 시스템은 다른 방식을 이용하는 시스템들과 비교했을 때 몇 가지 두드러진 장점을 가지고 있다. 그러나, 영상기반 시스템에서는 영상에 포함된 그림자가 차량검출의 정확도를 저해하는 요소로 작용하는 데, 특히 이동중인 차량에 의해서 발생하는 활성 그림자는 심각한 성능저하를 야기할 수 있다. 본 논문에서는 차량검출과 그림자 영향 제거를 위해서 배경 빼기와 에지 검출을 결합한 새로운 접근방법을 제안하였다. 제안한 방법은 노변의 지형지물에 의해서 발생하는 비활성 그림자가 크게 증가하는 상황에서도, 98[%]이상의 차량검출 정확도를 나타내었다. 본 논문에서 제안한 차량검출 방법을 기반으로 하여, 차량 추적, 차량 계수, 차종 분류, 그리고 속도 측정을 수행하여 각 차선의 부하를 나타내는 데 사용되는 차량 흐름과 관련된 여러 가지 교통정보를 추출하였다.

Abstract

Vehicle detection is the basic of traffic monitoring. Video based systems have several apparent advantages compared with other kinds of systems. However, In video based systems, shadows make troubles for vehicle detection, especially active shadows resulted from moving vehicles. In this paper, a new method that combines background subtraction and edge detection is proposed for vehicle detection and shadow rejection. The method is effective and the correct rate of vehicle detection is higher than 98% in experiments, during which the passive shadows resulted from roadside buildings grew considerably. Based on the proposed vehicle detection method, vehicle tracking, counting, classification and speed estimation are achieved so that traffic parameters concerning traffic flow is obtained to describe the load of each lane.

I. Introduction

ITS (Intelligent transportation systems) apply modern computer and communications technologies in transportation systems, resulting in improved mobility, safety, air quality, and productivity. ATMS (Advanced traffic management systems), as one of five functional areas of ITS, collects, uses, and distributes real-time information on congestion of arterial streets and expressways. Traffic parameters concerning traffic flow can be obtained by means of various sensors, such as buried loop sensors, radar, infrared detectors and other sensors.[1] However, most of the signals acquired from these sensors have to be interpreted. By contrast, video image processing systems employ machine vision technology to analyze data collected through Closed Circuit Television systems. Compared with other sensors,

* 正會員, 宁波大學
(The Institute of Circuits and Systems, Ningbo University)
** 正會員, 亞洲大學校 電子工學部
(Ajou University of Department of Electronics Engineering)
接受日字 : 2000年 9月4日, 수정완료일 : 2001年 6月19日

video based systems are easily intervened by humans because images from video surveillance cameras can be viewed directly by operators. Additionally, single camera and processor can serve multiple lanes, thus, video based system spends relatively lower cost than some of the other systems. Video based techniques are able to detect, track, classify, and identify vehicles, they are widely used in intersection and freeway monitoring and control. In video based traffic monitoring systems, shadows make troubles for vehicle detection, especially active shadows resulted from moving vehicles. Work in[2] analyzes types and properties of shadows and tries to extract features that can distinguish vehicles from shadows. However, these features are very complex.

In this paper, a combinative method is proposed for vehicle detection and shadow rejection, based on which vehicle tracking, counting, classification, and speed estimation are accomplished. In section 2, some video based methods for vehicle detection are first analyzed, and then the combinative method is described. Traffic parameter extraction based on vehicle tracking is presented in section 3. Experimental results and analysis of vehicle detection and traffic parameter extraction are given in section 4.

II. Video Based Vehicle Detection

1. Analysis of video based vehicle detection methods

Vehicle detection is the basic of traffic monitoring, and has been implemented via many different approaches. For video based vehicle detection, the commonly used approaches are gray-level comparison, inter-frame subtraction, background subtraction, edge detection based method, etc. Gray-level comparison utilizes statistical variation of gray-level features for road surface and vehicles, but is sensitive to environmental change. Moreover, it is almost impossible to determine the range of gray-levels of vehicles due to widely varying vehicle colors. Inter-frame subtraction takes difference between two successive frames so as to remove stationary part and get moving part within the image.[3] It is robust to environmental change, but unable to detect stationary vehicle. Moreover, the results of inter-frame subtraction are also influenced by speed of vehicles. Too low or too high speed may cause errors in vehicle detection.

Background subtraction takes difference between background image and input image.[4] Let $\{B^t_{i,j}\}$ and $\{C^t_{i,j}\}$ be the current estimated background image and the input image, respectively, where $1 \le i \le K$, $\le j \le L$. For each pixel of the input image, calculate the difference from the estimated background image

$$D^t_{ij} = |C^t_{i,j} - B)i, j^t|, \tag{1}$$

and the corresponding binary difference image can be obtained by

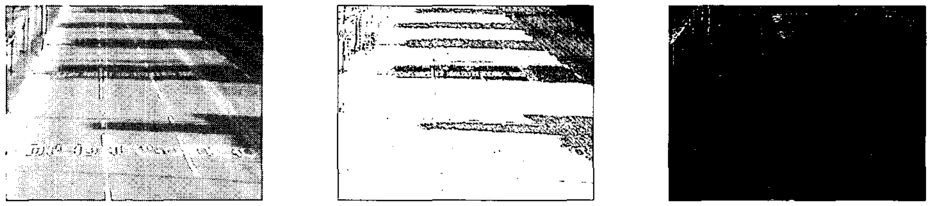
$$DB_{i,j} = \begin{cases} 1, & \text{if } D^t_{i,j} \ge T, \\ 0, & otherwise \end{cases}, \tag{2}$$

where T is a threshold. $DB_{i,j} = 1$ indicates that the pixel $C^t_{i,j}$ belongs to objects, otherwise it is a pixel of background. In background subtraction method, the result of vehicle detection strongly depends on the quality of estimated background image. Background generation consists of two steps: background extraction which generates initial background, and background update which is needed due to the change of ambient lighting, shadow, weather, etc. It is noticed that selective update (SU) scheme is better for background update, which is represented by

$$B^{t+1}_{i,j} = \begin{cases} jB^t_{i,j} + (1-k)C^t_{i,j}, & \text{if } DB_{,j} = 0 \\ B^t_{i,j}, & otherwise \end{cases}, \tag{3}$$

where $k(0 \le k \le k1)$ determines the update rate of the background and it is influenced by the sampling rate of image sequence. Problems with background subtraction include accumulation of update error and sensitivity to ambient lighting conditions and

(a) Frame C^{t-1} *ALum*=89.95 (b) Background subtraction with SU (c) Background subtraction with improved SU
(a) 프레임 C^{t-1} 평균조도=89.95 (b) 배경빼기(BS) (c) 개선된 배경 빼기



(d) Frame C^t *ALum*=117.85 (e) Background subtraction with SU (f) Background subtraction with improved SU
(d) 프레임 C^t 평균조도=117.85 (e) 배경 빼기(BS) (f) 개선된 배경 빼기

그림 1. 연속적인 두 프레임간의 급격한 평균 조도 변화에 의해 발생하는 오류 검출
Fig. 1. Error detection caused by sudden change in average luminance of two continuous frames.

shadows.

Edge based method is another used approach to vehicle detection, as the edge information still remains significant despite the variation of ambient lighting. Various surfaces and different parts and colors of a vehicle create significant edges. Even the vehicles, which have the same color as the surface of the road, reflect more light and can be detected in this way.[5] Moving edge detection can be obtained by using spatial and temporal image gradients. The spatial edges can be identified by any kind of edge detector, while the temporal gradient can be approximated using the difference image between successive frames. Alternative moving edge detection can be accomplished by subtracting edge image of background image from edge image of current frame. However, edge detection based method will fail in detecting vehicle, whose edges are unclear, especially when a vehicle with dark color is within a shadow.

2. Video based combinative method for vehicle detection

In background subtraction, the key step is to obtain a reliable background image. It is noticed that the average luminance of successive frames sometimes may have sudden change, which influences the results of background subtraction because the background image may not meet with the current input image. Fig. 1 gives such an example, in which the average luminance of the two successive frames are 89.95 and 117.85, respectively. Fig. 1(b) and Fig. 1(e) show the binary difference images between background images and current input images. It is seen that Fig. 1(e) is incorrect due to the sudden change of luminance of the input image. To solve this problem, the average luminance *ALum* of current input frame C^t is first calculated, and compared with that of the previous frame C^{t-1}. If the difference *DALum* between the two average luminance is larger than a threshold, each pixel value of the background image is added with the difference, so as to meet with the changed luminance of current input image. This is called background adjusting. Fig. 1(c) and Fig. 1(f) give the corresponding difference images obtained by using this improved background subtraction with selective updating scheme, and Fig. 1(f) is more effective for vehicle detection than Fig. 1(e).

Background subtraction and edge detection method for vehicle detection have advantages and also disadvantages. Background subtraction suffers from

(a) Input image (b) Background subtraction (c) Edge image

(a) 입력 영상 (b) 배경 빼기 (c) 에지 검출

그림 2. 배경 빼기와 에지 검출 방법에서의 그림자 제거 결과(에지 검출 방법이 효과적)

Fig. 2. Shadow rejection is easier for edge detection method than background subtraction.

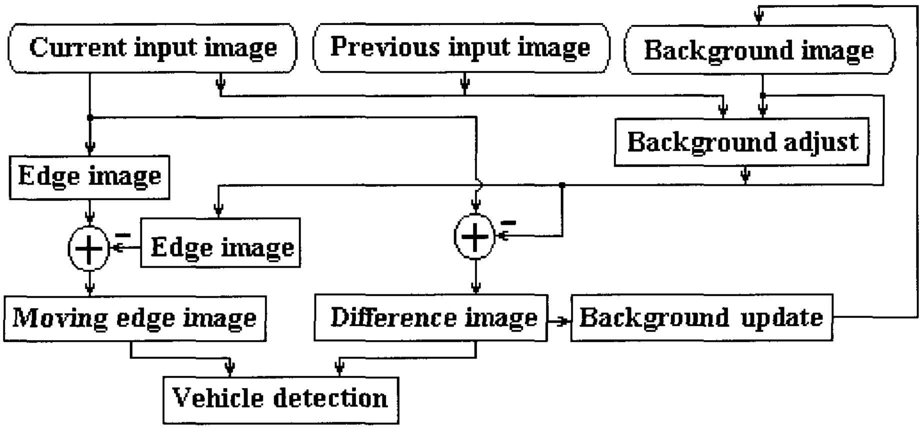


그림 3. 차량 검출과 그림자 제거를 위한 결합 방법의 처리 과정

Fig. 3. Flow of combinative method for vehicle detection and shadow rejection.

(a) Vehicle location (b) Partial difference image $\{DB_{i,j}\}$ (c) Partial moving edge image $\{E_{i,j}\}$

(a) 차량 위치 (b) 차이 영상 (c) 이동 에지 영상

그림 4. 차량 검출 결과

Fig. 4. Vehicle detection results.

big problems caused by active shadows. On the contrary, edge detection method is easier for shadow rejection than background subtraction, because shadows usually possess less edge pixels than vehicles, which can be used to distinguish shadows from vehicles. In Fig. 2, there is a big active shadow of the first vehicle projected on the left lane. For background subtraction, the shadow is too big that may be mistaken as a vehicle. By contrast, the edge pixels of the shadow are significantly less than that of the vehicle, as shown in Fig. 2(c). The windows in Fig. 2 on each lane are "virtual detectors", which emulate inductive loops. These virtual loops are used to judge whether vehicles are present on the loop or not.

Based on the above analysis, a combinative method, which takes advantage of background subtraction and edge detection, is proposed in this paper. Fig. 3 gives flow of the combinative method, and it is described as follows.

1) Calculate the average luminance of current input image $\{C_{i,j}^t\}$ and previous image $\{C_{i,j}^{t-1}\}$, and the difference *DALum* between the two average luminance, if $|DALum|>th$, in which th is a threshold (empirically th=5), then adjust the background image $\{B_{i,j}^t\}$ by

$$C_{i,j}^t = B_{i,j}^t + \mathrm{DALum}. \tag{4}$$

2) Calculate edge images of current input image and background image, respectively. Sobel operators $\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$ and $\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$ are used for vertical and horizontal edge detection, which is represented as

$$d_{i,j} = \max \left\{ \frac{1}{4} \sum_{y=j-1}^{j+1} w_y |I_{i-1,y} - I_{i+1,y}|, \frac{1}{4} \sum_{x=i-1}^{i+1} w_x |I_{x,j+1} - I_{x,j-1}| \right\} \tag{5}$$

where I is current input image $\{C_{i,j}^t\}$ or background image $\{B_{i,j}^t\}$, w_y and w_x correspond to the weights in the two Sobel operators. The corresponding binary edge image is obtained by

$$EdgeI_{i,j} = \begin{cases} I, & \text{if } d_{i,j} \geq T_E \\ 0, & otherwise \end{cases}, \quad T_E = \alpha + \log 2 d_{i,j} \tag{6}$$

where α is a constant 20, as the lowest limit of the threshold. Then obtain the moving edge image $\{E_{i,j}\}$ by subtracting edge image of $\{B_{i,j}\}$ from edge image of $\{C_{i,j}\}$.

3) Calculate difference image $\{DB_{i,j}\}$ between background image and current input image according to Eqs.(1)-(2). The threshold T in Eq.(2) can be adopted as $\beta + \log_2 D_{i,j}^t$, and is a constant (empirically β=23).

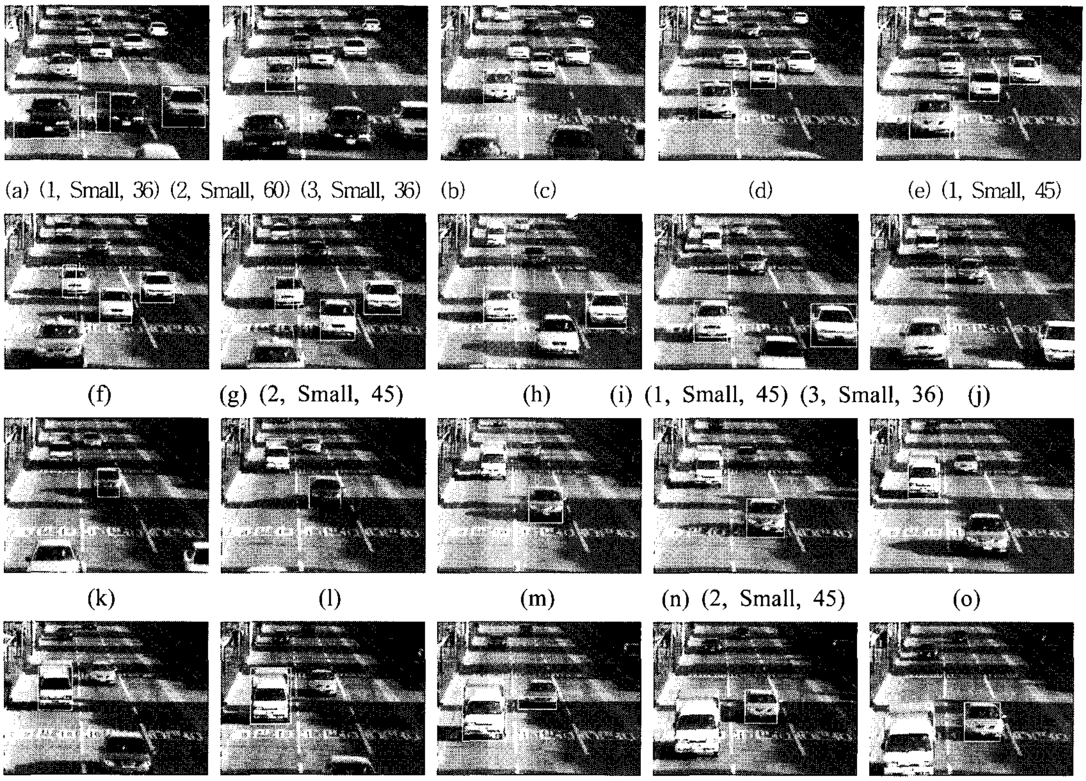
4) Check each eight connected white area of $\{DB_{i,j}\}$ in a window to determine whether it is an object, and also check the number of edge pixels at the corresponding area $\{E_{i,j}\}$ in for shadow rejection, so as to judge whether vehicles are present in the window or not. If there is a vehicle in the window, its information is recorded for post-processing. Fig. 4 gives such an example for vehicle location.

5) Selectively update the background image according to Eq.(3).

Ⅲ. Traffic Parameter Extraction Based Vehicle Tracking

In video based loop emulation, difference traffic parameters requires different "virtual detectors", that is, different size of windows. To measure the speed, study[5] places two parallel windows on each lane, and the time between the passage of a vehicle at the two separate parallel windows is detected. The information about the number of frames taken for the passage of each vehicle between the two windows is recorded. This information is used to compute the speed of vehicles because the distance between the windows is fixed. To categorize the vehicles into small, medium, and large groups, the length of vehicles is computed by using the speed and the number of frames taken to detect vehicle in the two parallel windows.

Speed estimation can also be accomplished via vehicle tracking. In the proposed combinative method, a large enough eight connected area in difference image $\{DB_{i,j}\}$ is regarded as an object, that is, a vehicle. Individual vehicle is tracked within a fixed long window on a lane. The windows are indicated by boxes in Fig. 2(a). When a vehicle is detected to enter the window, its position is recorded and updated as it moving within the window. When it is about to exit the window, its speed is computed by



(a) (1, Small, 36) (2, Small, 60) (3, Small, 36) (b) (c) (d) (e) (1, Small, 45)
(f) (g) (2, Small, 45) (h) (i) (1, Small, 45) (3, Small, 36) (j)
(k) (l) (m) (n) (2, Small, 45) (o)
(p) (q) (r) (1, Medium, 45) (s) (t) (2, Small, 60)

그림 5. 차량 추적, 차종 분류, 그리고 속도 측정
Fig. 5. Vehicle tracking, classification and speed estimation.

$$Speed = \frac{D \times F}{F_t} (Km/s), \tag{7}$$

where D is the physical distance between the entrance and exit of the window, F and F_t are frame rate and number of frames taken for the vehicle to pass the window, respectively. Usually high frame rate results in more accurate speed estimation. The type (small, medium and large) of individual vehicle can be classified according to length and width of the vehicle, and additionally, the number of vehicles passed through each lane can also be counted.

Fig. 5 gives an example of vehicle tracking by means of the combinative method. The boxes in Fig. 5 indicate vehicle locations calculated by computer automatically. The vehicle type and speed identified by computer are also listed as (lane no., vehicle type, speed) below the corresponding frames.

Ⅳ. Experiments and Analysis

1. Results of vehicle detection

To compare the accuracy of some video based methods for vehicle detection, some experiments are implemented, and Table 1 gives the vehicle detection results of these different methods. These methods include background subtraction method with selective updating scheme (BS_SU), the improved background subtraction method with background adjusting (Improved BS_SU), edge detection based method (Edge detection), and the combinative method (Combination). In the table, "missed detection" means that there was a vehicle present in the window, but the algorithm failed to detect it, while "false detection" corresponds to the opposite condition. The test video lasted about 86 minutes and the passive

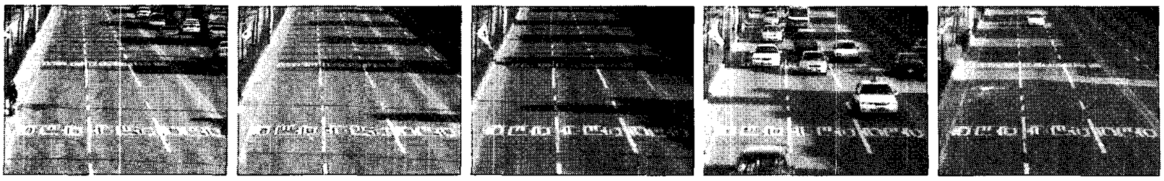


그림 6. 실험용 영상에서 비 활성 그림자의 심각한 증가
Fig. 6. Passive shadows grew considerably in the test image sequence.

표 1. 86분 분량의 실험용 입력영상에 대한 차량 검출 결과
Table 1. Results of vehicle detection in 86 minutes.

Lanes / Methods	Left lane			Middle lane			Right lane		
	Missed detection rate	False detection rate	Correct rate	Missed detection rate	False detection rate	Correct rate	Missed detection rate	False detection rate	Correct rate
BS_SU	2.64%	2.47%	94.89%	4.57%	2.14%	93.29%	2.33%	1.85%	95.82%
Improved BS_SU	2.73%	1.27%	96.00%	4.54%	0.94%	94.52%	2.35%	0.86%	96.79%
Edge detection	1.52%	0.40%	98.08%	2.26%	0.95%	96.79%	2.59%	1.19%	96.23%
Combination	0.52%	0.14%	99.34%	1.12%	0.37%	98.51%	0.72%	0.69%	98.59%

표 2. 86분 분량의 입력영상에 대한 차량 계수 측정 결과
Table 2. Total results of vehicle counting in 86 minutes.

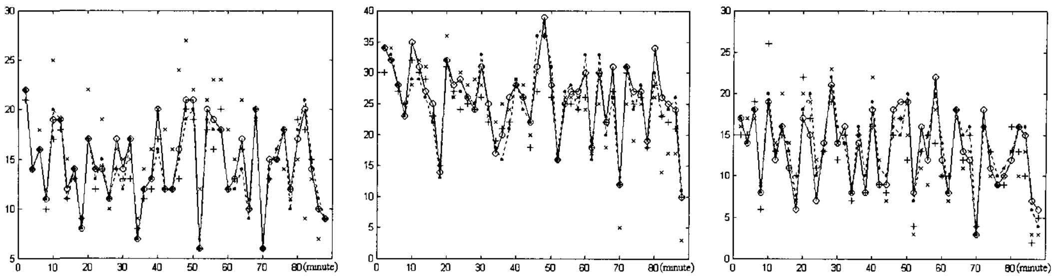
Lanes / Methods	Left lane		Middle lane		Right lane	
	Vehicles	Deviation rate	Vehicles	Deviation rate	Vehicles	Deviation rate
Manual counting	644		1142		575	
Improved BS_SU	699	+8.54%	1067	- 6.57%	561	- 2.43%
Edge detection	613	- 4.81%	1073	- 6.04%	554	- 3.65%
BS & Edge	626	- 2.80%	1116	- 2.28%	598	+4.00%

shadows resulted from roadside buildings grew considerably during the period, as shown in Fig. 6.

It is noticed that in background subtraction and edge detection method, missed detection mainly comes from vehicles with dark colors, especially when such a vehicle is within shadows. In this case, the luminance of the vehicle is similar to the luminance of the background, and the edges of the vehicle are also not clear enough. False detection, on the other hand, is mainly caused by active shadows resulted from vehicles passing through the neighbor lane in background subtraction. Edge detection based method is a good solution for shadow rejection, because shadows have less edge pixels than vehicles. Moreover, edge detection method is robust to luminance change compared with background subtraction. However, it also easily misses vehicles when the edges of vehicles are not clear. BS_SU is quite sensitive to luminance change of the frames. The improved BS_SU with background adjusting performs better than BS_SU.

In the experiments, compared with background subtraction and edge detection method, the combinative method achieves better results, since it takes advantage of the two methods. The accuracy of vehicle detection of the combinative method is higher than 98%, as shown in Table 1.

It is seen that vehicles that are changing lane within a monitoring window may cause both missed detection and false detection, because they affect two



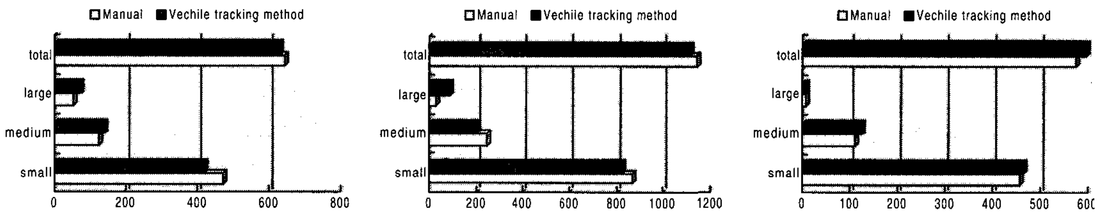
(a) Left lane (b) Middle lane (c) Right lane

(a) 좌측 차선 (b) 중앙 차선 (c) 우측 차선

Note: 'x' Improved BS_SU; '+' Edge detection; —○— Manual; --●--Vehicle tracking based on combinative method

그림 7. 차량 계수 결과(2분 간격)

Fig. 7. Vehicle counting results in every 2 minutes.



(a) Left lane (b) Middle lane (c) Right lane

(a) 좌측 차선 (b) 중앙 차선 (c) 우측 차선

그림 8. 차량 추적에 의한 차종 분류 결과

Fig. 8. Vehicle classification by means of vehicle tracking.

lanes. Large vehicles such as buses and trucks may be detected as two vehicles, while small vehicles are possibly missed in this case.

The location of the video camera is critical to effective operation of the video based vehicle detector. The best possible place for the camera is directly over the center lane. The farther a lane is from the camera, the less reliable the vehicle detection of the lane is. However, it should be noticed that the errors arisen from some vehicle overlap the others in image, is the inherent problem of image processing based vehicle detection methods.

No matter in which kinds of video based vehicle detection methods, the selection of threshold is also significant to the accuracy of vehicle detection. The optimal threshold can be empirically obtained through numerous experiments, in which the weather conditions and the time in a day should be taken into account.

2. Results of traffic parameter extraction

Vehicle counting is one of the basic functions for real-time traffic monitoring. It can be used to describe the load of each lane. Fig. 7 gives the results of vehicle counting by using the combinative method based vehicle tracking, compared with the results of manual counting and background subtraction with background adjusting (Improved BS_SU) and edge detection (Edge detection) methods, respectively. The three figures correspond to the three lanes as shown in Fig. 1~2 and Fig. 4~6. The horizontal axes of the figures represent the time, the unit of which is minute, while the vertical axes are the number of passing vehicles within the time intervals. The test video lasts about 86 minutes. The counting results in Fig. 7 are refreshed every 2 minutes. The solid and dotted line in each figure correspond to the result of manual counting and the result of the proposed method, respectively. The meanings of the symbols are given in the figure. It is also shown that the results of the proposed method approach to the results of manual counting.

The results of vehicle counting within total 86 minutes are given in Table 2, in which the columns of "Vehicles" are the number of vehicles detected via each method, while the percentages in "Deviation rate" give the deviation of each automatic counting from the manual counting. In fact, during a long time interval, the effects of missed and false detection in vehicle counting will partly counteract, thus the evaluation of approaches to vehicle counting within a short time is more cogent than in a long time.

Fig. 8 gives the results of vehicle classification for the three lanes. In the figure, the "small", "medium", and "large" correspond to the three types of vehicles, while the "total" is the sum of them. The horizontal axes are the numbers of vehicles with each type. The lower bar of each group is manual classification, while the upper bar corresponds to result made by computer.

The experiments show that the method based on vehicle tracking works well under light traffic condition. By contrast, under heavy traffic condition, the method may mistake a vehicle as a bigger one in vehicle classification when the vehicle is connected with other following vehicle in input images. The farther the vehicle is from the camera, the worse the problem is. It is clear that high location of camera will reduce such kind of mistake. In single frame analysis, active shadow projected on other lane sometimes may still be recognized as a vehicle, but analysis of successive frames can reduce such errors, because a possible vehicle region may be removed as false detection if the corresponding region is not detected in next frame. However, the vehicle that changes lane within the window may cause trouble because it crosses two lanes.

V. Conclusion

In this paper, a new method that combines background subtraction and edge detection is proposed for vehicle detection and shadow rejection, since shadows usually possess less edge pixels than vehicles, which can be used to distinguish shadows from vehicles. The method is effective and the correct rate of vehicle detection is higher than 98% in experiments, during which the passive shadows resulted from roadside buildings grew considerably. Based on the proposed vehicle detection method, vehicle tracking, counting, classification and speed estimation are achieved so that traffic parameters concerning traffic flow is obtained to describe the load of each lane.

Reference

[1] Klein Lawrence A., Kelley Michael R., and Mills Milton K., "Evaluation of overhead and in-ground vehicle detector tchnologies for traffic flow measurement", *Journal of Testing & Evaluation*, vol.25, no.2, pp.205-214, 1997.

[2] Jiang Zifeng, "A shadow rejection algorithm for vehicle presence detection", *1998 IEEE International Conference on Intelligent Vehicles*, pp.182-188, 1998.

[3] Hongjiang Zhang, Yihong Gong, Dan Patterson, and Atreyi Kankanhalli, "Moving object detection, tracking and recognition", *The Third International Conference on Automation, Robotics and Computer Vision*, pp.1990-1994, 1994

[4] P.G. Michalopoulos, "Vehicle detection video through image processing: The Autoscope system", *IEEE Trans. on Vehicular Technology*, vol.40, no.1, pp.21-29, 1991

[5] M. Fathy, and M. Y. Siyal, "A window-based image processing technique for quantitative and

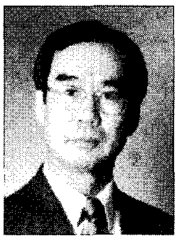
qualitative analysis of road traffic parameters", IEEE Trans. on Vehicular Technology, vol.47, no.4, pp.1342-1349, 1998.

저 자 소 개

郁 梅(正會員)

received the B.S. and M. S. in Computer Science from Hangzhou Institute of Electronics Engineering, China, in 1990 and 1993; the Ph. D. in Electronics Engineering from Ajou University, Korea, in 2000. She in now with the Institute of Circuits and Systems, Ningbo University, China, as an associate professor. Her current research interests include digital image processing, computer vision, intelligent transportation system.

1990 중국 항주대학교 전자공학과 학사. 1993 중국 항주대학교 전자공학과 석사. 2000 아주대학교 전자공학과 박사. 2001 현재 중국 닝보대학교 부교수. 연구분야 : 디지털영상처리, 컴퓨터비전, 지능형교통체계(ITS)



金 容 得(正會員)

received the B.S., M.S. and Ph. D. from Department of Electronics Engineering at Yonsei University, Korea, in 1971, 1973, and 1976, respectively. From 1973 to 1974, he was a research member at Department of Electronics Ecole Superieur Delectricite in France. From 1979 to 1980, he was a visiting professor of Department of Electronics Engineering at Standford University. Since 1974, he has been a professor of Department of Electronics Engineering at Ajou University, Korea. His current research interests include ATM Network, multimedia system, and intelligent transportation system.

1971 연세대학교 전자공학과 학사. 1973 연세대학교 전자공학과 석사. 1973~1974프랑스 ESE 전자공학연구실 연구원. 1978 연세대학교 전자공학과 박사. 1979 미국 Standford 대학교 연구교수. 1974~현재 아주대학교 교수. 연구분야 : ATM 통신망, 멀티미디어 시스템, 디지털영상처리, 지능형교통체계(ITS)