Estimating Traffic Flow by Image Processing and the Usage of an Adaptive Traffic Signal Control System

Sheng-Fuu Lin[*1] Shih-Che Chien[*2] Kuo-Yu Chiu[*3]
Dep. Electrical and Control Eng., National Chiao Tung University,[*1*2*3]
1001 Ta Hsueh Road, Hsinchu, Taiwan 300, ROC
(TEL: 886-3-571-2121 Ext: 54365, E-mail: sfline@mail.nctu.edu.tw)[*1]
(TEL: 886-3-571-2121 Ext: 54411, E-mail:{ shihche.chien[*2],*kuoyu*[*3]*}@gmil.com)*

The Robust and reliable traffic information detection from images acquired by traffic surveillance system is an important problem with numerous applications including adaptive traffic signal control system and traffic analysis system. This paper describes an application of computer vision techniques to traffic surveillance. The objective in this paper is to improve the efficiency of traffic flow estimation that include pedestrian flow and vehicular flow estimation. Specifically, some vision based methodologies that including perspective transformation, object segmentation and classification have been employed to traffic flow estimation. Traffic congestion in urban area is still an important issue to be solved. One of the reasons of congestion in urban area is imperfect operation of traffic signal control. A traffic signal control problem can be formulated as a decision making problem for a stochastic dynamic system. To minimize the queue length and vehicle delay time, an adaptive traffic signal control system that used signification traffic information to find the optimal phase sequencing of traffic signal has been proposed in this paper. The experiment results of intersection traffic simulation indicate that our approach has better performance than other two control systems.

Keywords: *Traffic flow, vehicular flow, pedestrian, adaptive traffic signal control, fuzzy inference system.*

1. Introduction

As a result of the growing rapidly of urbanization, many traffic problems take place in the modern traffic systems. The intelligent transportation system (ITS) is a special issue to solve above-mentioned problems. Broadly speaking, the topics treated in ITS can be divided into three groups: 1) modeling, analysis, and control of traffic systems; 2) vision-based techniques for intelligent vehicles and traffic systems; and 3) emerging and other topics in ITS [1]. The traffic congestion occurs when too many vehicles attempt to use a common urban road with limited capacity. The efficient, safe, and less polluting transportation of persons and goods calls for optimal utilization of the available infrastructure via suitable application of a variety of traffic control measures and mechanism. So the adaptive traffic signal system become a development trend of ITS.

Since the vision based surveillance system has advantages of flexible, large search area and low cost, therefore, it becomes popular in applications to detect traffic information. The purpose of traffic flow estimation is to incorporate the data gathered by sensors into the adaptive traffic signal control system to make traffic flow smoother and more unhindered. Previous research on real-time traffic adaptive signal control systems has primarily focused on vehicle traffic for adaptive signal control. Although pedestrians play an important role in an intersection, very little research has been conducted on incorporating real-time pedestrian traffic data into adaptive signal control systems. The objective of this paper is to present an approach to estimate the number of pedestrians approximately in a crosswalk and the vehicular flow of an approaching road using fixed cameras. The collected data can then be incorporated into an adaptive traffic signal controller system.

The literature on pedestrian detection and vehicle detection is vast, and a comprehensive survey is beyond the scope of this paper. Methods of detecting pedestrians and vehicles can generally be divided into two groups: 1) microscopic measurement [2]-[10]; 2) macroscopic measurement [11]-[13]. Methods of microscopic measurement can identify each pedestrian or vehicle in an image, and their movement is tracked based on pattern matching. However, these methods demand heavy computational loads, and their performance degrades severely when scenes get complex. Methods of macroscopic measurement are to extract global features of an image, then classify scenes into several classes. These methods can simplify the issue of estimating traffic information, while these systems have to be trained with hundreds sample images or data sets. Once the scene is changed, the machine learning process is forced to restart from the very beginning. Previous developments of adaptive traffic signal control [14]-[18] used dynamically adjustable phase length to get better control performance, but they rarely considered the topic of optimal cycle and relative saturation degree, thus the traffic resources was wasted in some occasions. There

has still space to improve the adaptive traffic signal control system.

For applications to adaptive traffic signal control, there is no need to recognize each pedestrian or vehicle. What we need is the traffic flow data, such as vehicular flow on the road and pedestrian flow in the crosswalk. In this paper, an efficient algorithm is proposed to estimate the traffic flow. The proposed algorithm uses pre-computed global features, so it needs not to be trained and the estimation can be done efficiently. In order to incorporate the proposed traffic information to signal controller, a fuzzy adaptive traffic signal control system is presented in this paper. The system also takes the optimal cycle and relative saturation degree into consideration.

The remainder of this paper is organized as follows. In Section 2, the vision based measurement system including pedestrian evaluation and vehicular flow estimation are described. The fuzzy adaptive traffic signal controller system is presented in Section 3. Several experimental results are shown to validate the proposed system in section 4. Finally, conclusions and future work are given in Section 5.

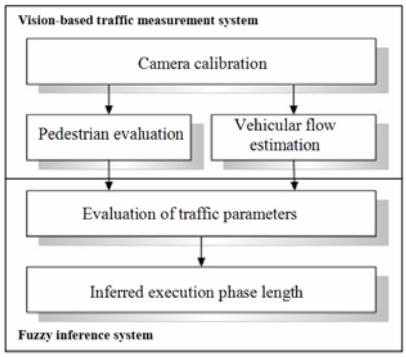


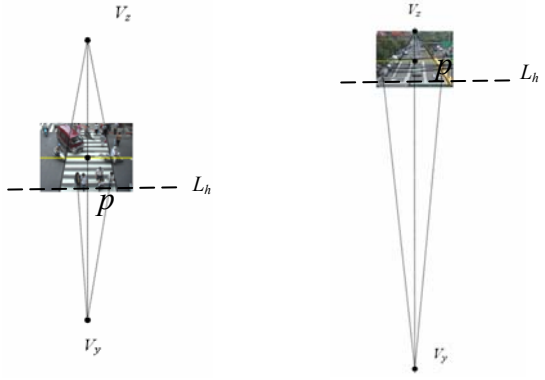
Fig. 1 Simplified block diagram of the proposed traffic signal control system.

2. The vision-based traffic measurement system

The structure of the vision-based traffic measurement system is shown as half up of block diagram in Fig. 1. There are two stages in the system. The first stage is camera calibration that can perform the perspective transformation. Traffic flow estimation is performed in the second stage that can be divided into two traffic flow measurements: 1) pedestrian evaluation, and 2) vehicular flow estimation. These two traffic information are the major impact on traffic.

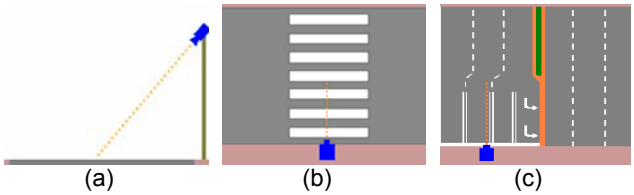
2.1. Camera calibration

The goal of camera calibration is to compute the perspective transformation matrix. Thus, recovering 3D structure from images becomes a simpler problem when the images are taken with calibrated cameras. The camera calibration can be done if its intrinsic parameters and its position and orientation with respect to some reference in the scene are given. Camera parameters can be obtained by using standard methods if a calibration object or measurements of enough 3-D points in the scene are available [19]. However, such measurements are not always available. Vanishing points of parallel lines have been proven to be useful features for this task. Furthermore, intrinsic and extrinsic parameters can be obtained by using the method that just only used one vertical vanishing point and a horizontal vanishing line to provide an approximate solution with some assumptions [20]. There are two assumptions, the angle from horizontal line to the horizon line is zero and the principle point locates on the center of image. By using this method, how to estimate the vertical vanishing point and the horizontal line has become an interesting problem. In this approach, vertical vanishing point (V_y) and the horizon line can be obtained by observing the scene of traffic surveillance video sequence. Some traffic signs such as lanes and zebra crossing and the objects which are perpendicular to ground can provide some useful information to estimate vertical vanishing point and the horizontal vanishing line in the scene of intersection traffic surveillance. For examples, in the scene of crosswalk, Fig. 2 (a), people on the crosswalk can be used to get the vertical vanishing point (V_y) and the crosswalk sign can be used to find out the horizon line (L_h) and vanishing point of z axis (V_z). The principle



(a) (b)

Fig. 2. Illustrate the vanishing points (V_y and V_z), the horizon line (L_h) and the principle point (p) in (a) the scene of crosswalk; (b) the scene of road.



(a) (b) (c)

Fig. 3. The camera setup. (a) camera side view; (b) camera top view in the scene of crosswalk; and (c) camera top view in the scene of road.

point (p) here is located on the center of scene. Similarly, in the scene of road, Fig. 2 (b), vertical structure like traffic sign pole can be used to get the vertical vanishing point and the lines of lane can be used to get the horizon line and vanishing point of z axis.

In this paper, the camera is installed at a high place such that it can take images with a bird's-eye view, as shown in Fig. 3 (a). This allows a larger coverage and less occlusion, especially avoiding the situation that the entire scene is occluded by one object. In addition, the direction of the camera lens is set to be parallel to the crosswalk and road, as shown in Fig. 3 (b), (c). This special arrangement is helpful to observe and find out the vertical vanishing point and horizon line. By using the camera calibration method and camera arrangement mentioned above, the perspective transform can be obtained more easily. Then, 3D structure can be recovered from 2D image if the height of transformed point is given.

2.2. Pedestrian estimation

This section demonstrates an efficient method to calculate the approximate numbers of pedestrians in the crosswalk.

2.2.1. Foreground detection. For saving much processing time, background subtraction is generally adopted to extract out foreground. It is not reliable for direct subtraction with arbitrary background. Therefore, how to build a strong background is very important task. In this paper, median filter is adopted to build up strong background. The background is separated into three sets with red, green and blue color space:

$$Ib_k(u,v) = \text{median}(I_k^{(i)}(u,v)),\ i = 0,1,2,...,n-1, \quad (1)$$

where k represents the red (r), green (g), or blue (b) color space of RGB color model and n is the necessary amount of input image to generate background; $I^{(i)}$ is the input image i. For instance, the input image of crosswalk is shown in Fig. 4(a) and the background image is shown in Fig. 4(b).

The foreground image can be derived from comparing the background and the current image. Foreground image is given by

$$Im(u,v) = \begin{cases} 1, & \text{if } \|I(u,v) - Ib(u,v)\| > \tau_f, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

$$\|I(u,v) - Ib(u,v)\| = \Big[(I_r(u,v) - Ib_r(u,v))^2 + (I_g(u,v) - Ib_g(u,v))^2 + (I_b(u,v) - Ib_b(u,v))^2 \Big]^{1/2}, \quad (3)$$

where $I(u,v)$ is the input image and $Ib(u,v)$ is the background image; τ_f is a suitable threshold value of separation of foreground from background. In order to deal with the illumination change problem, the background image should be updated in each period of time which depends on the variation of illumination. After the foreground extraction, the foreground image might still contain some noise and incomplete area. Therefore, image enhancement including median filter and morphological operation is necessary to deal with these tasks. Median filtering is performed to remove isolated noise, and the 3 pixel diameter disc morphologic closing operation is used to fill gaps, eliminates small holes and smooth contour in foreground blob. The foreground image is more concrete after these image processes. An example of image enhancement process of Fig. 4 (c) is shown in Fig. 4 (d).

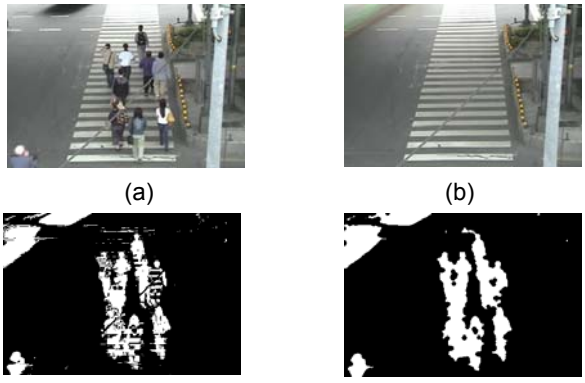


Fig. 4. The process of foreground detection. (a) input image; (b) background image; (c) foreground image; (d) result of image enhancement process of (c).

2.2.2. Eliminate vehicles in the crosswalk. After foreground extracting, foreground blobs may contain pixels which belong to the objects, such as sedans, trunks, and pedestrians. Those foreground pixels that belong to vehicles must be discarded because it will do a great effect upon the accuracy of pedestrian evaluation.

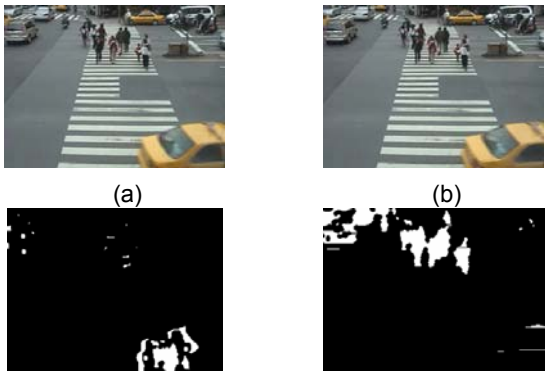


Fig. 5. The process of eliminating vehicles on the crosswalk. (a)-(b) two sequential frames; (c) vehicular blob; (d) the foreground image without vehicle.

The information about the moving direction of objects could provide a robust estimation, since the pedestrians are moving along with crosswalk and the vehicles are opposite. By labeling the foreground blobs and performing inter-frame differencing between sequential frames, it is easy to find the moving direction of each blob. If the component along with crosswalk of moving direction is larger than its orthogonal component, the blob is defined as pedestrian. On the contrary, the blob is defined as vehicle, thus most of the pixels belong to vehicles in foreground image can be eliminated. Unavoidably, few pixels which belong to pedestrians are eliminated in this step, but it just has a little effect upon the final result of the pedestrian evaluation. An example of eliminating vehicles on the crosswalk is shown in Fig. 5.

2.2.3. Estimate the number of pedestrians. Crosswalk area is the interested area in the pedestrian estimation part, and therefore, the area that out of interested area will not be concerned. A simple ROI mask can clean up the foreground image and preserve the crosswalk area only. The *ROI* mask can be denoted as follow:

$$c(u,v) = ROI\{I_{EV}(u,v)\} \tag{4}$$

where I_{EV} (u,v) is the result of image enhancement processing and vehicle elimination.

The contour of pedestrian is similar to ellipse. Hence, the contour of pedestrian can be replaced by ellipse to simplify. In camera projection model, the projection image (ellipse with center at (u, v)) of a pedestrian who stands far away form the camera is much smaller than that (ellipse with center at (u_0, v_0)) of the pedestrian standing near to the camera, as show in fig. 6(a). In order to normalize the effect of camera projection image, a normalized gain (*ng*) is adopted and it designed to equal the inverse ratio of representative ellipse size. Here, we assume that the point $P^{-1}(u, v, L)$ which is derived by inverse perspective transformation of camera image point (u, v) with given height L (half major axis of ellipse; half height of pedestrian) is the center of ellipse with major axis length $(2L)$. Then, the two vertex points, $(P^{-1}(u, v, L), 2L)$ and $(P^{-1}(u, v, L), 0)$, of major axis can be obtained. Similarly, the two vertex points of major axis of projection ellipse in camera image, $P(P^{-1}(u, v, L), 2L)$ and $P(P^{-1}(u, v, L), 0)$, can be derived by perspective transformation. The length of major axis of ellipse in camera image can be obtained. (Length of major axis = $P(P^{-1}(u, v, L), 2L) - P(P^{-1}(u, v, L), 0)$) Therefore, the *ng* (u, v) at camera coordinate (u, v) can be derived as follow:

$$ng(u,v) = \frac{\left[P\left(P^{-1}(u_0,v_0,L),2L\right) - P\left(P^{-1}(u_0,v_0,L),0\right)\right]^2}{\left[P\left(P^{-1}(u,v,L),2L\right) - P\left(P^{-1}(u,v,L),0\right)\right]^2}, \tag{5}$$

where (u_0, v_0) is the reference point and denoted at the center of image, that is the point p in Fig. 2(a) or Fig. 2(b). Subsequently, the number of normalized pixels (*NP*) is counted in stead of the number of original pixels and expressed as follow:

$$NP = \sum_{u,v} c(u,v) \cdot ng(u,v), \tag{6}$$

where $c(u, v)$ is the crosswalk region separated from foreground image by using ROI mask and $ng(u, v)$ is the normalized gain at coordinates (u, v).

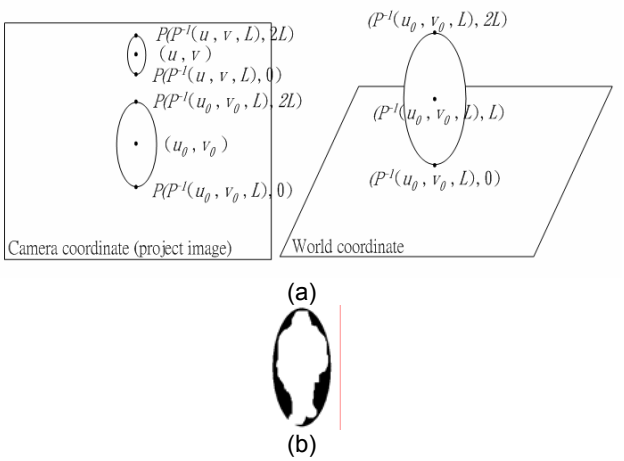


Fig. 6. Illustrate the (a) *ng* and relation of these two coordinates and (b) an ellipse template of pedestrian.

An ellipse pedestrian template of height 180cm is constructed, as show in Fig. 6(b) According to experimental observation, about 70% to 80% of template pixels are also pedestrian foreground pixels. Here, 75% is adopted as the ratio of pedestrian pixel occupying in ellipse template. Therefore, the number of pedestrians *N* on a crosswalk can be estimated approximately by the following equation:

$$N = round\left(\frac{NP}{75\% \cdot AB\pi}\right), \tag{7}$$

where operation *round*(·) is to round a fraction toward nearest integer; *A* and *B* are respectively the major axis and the minor axis of an ellipse which is mapped onto the image plane from the pedestrian template; *B* is set to be 0.4 times *A* in the algorithm.

2.3. Vehicular flow estimation

This section demonstrates a fuzzy inference method to estimate vehicular flow. Here, two features are used for fuzzy inference system. One is the amount of vehicular pixel and the other one is the vehicle texture on the road.

2.3.1. Amount of vehicular pixel. The calculation of vehicular pixel on the road is similar to estimation the pedestrian that mention above. First, background establishment is necessary and then background subtraction is adopted for foreground extraction. After foreground extraction, ROI mask is taken to use for road region, $c(u, v)$, separation from whole foreground image. A normalize gain, $ng(u, v)$, is performed to aid the calculation of vehicular pixel on the road. Subsequently, the amount of vehicular pixel (NV) can be derived from

$$NV = \sum_{u,v} c(u,v) \cdot ng(u,v). \tag{8}$$

2.3.2. Texture feature extraction. Although, the amount of foreground objects can be calculated in the ROI region, it is not enough to replace vehicular flow. Thus, more useful features are needed for traffic control system, such as vehicular position. For example, if traffic light turns to red, the holding time for the vehicle depends on the distance between it and the stop line. Therefore, how to extract more feature of vehicle on the road will be a challenge and important task here.

The texture parameters are extracted via *gray level co-occurrence matrix* (GLCM) [21] which is a well-know statistic technique for feature extraction, and four directions, 0°, 45°, 90° and 135°, are chosen to run position operator. The five attributes, contrast *Cts*, correlation *Cor*, energy *Eng*, homogeneity *Hom*, and entropy *Ept*, are used to extract the texture features from preceding matrix, and the definitions of these attributes can be found in [21] and [22]. Here, four sets, $\mathbf{I}_0$, $\mathbf{I}_{45}$, $\mathbf{I}_{90}$ and $\mathbf{I}_{135}$, of feature vector from input gray scale image I can be obtained; each of them has five components. The feature sets of image I are denoted as:

$$\begin{aligned}
\mathbf{I}_0 &= \begin{bmatrix} ICts_0 & ICor_0 & IEng_0 & IHom_0 & IEpt_0 \end{bmatrix}^T, \\
\mathbf{I}_{45} &= \begin{bmatrix} ICts_{45} & ICor_{45} & IEng_{45} & IHom_{45} & IEpt_{45} \end{bmatrix}^T, \\
\mathbf{I}_{90} &= \begin{bmatrix} ICts_{90} & ICor_{90} & IEng_{90} & IHom_{90} & IEpt_{90} \end{bmatrix}^T, \\
\mathbf{I}_{135} &= \begin{bmatrix} ICts_{135} & ICor_{135} & IEng_{135} & IHom_{135} & IEpt_{135} \end{bmatrix}^T.
\end{aligned} \tag{9}$$

where $ICts_0$, $ICor_0$, $IEng_0$, $IHom_0$, and $IEpt_0$ represent the value of contrast, correlation, energy, homogeneity, and entropy respectively when the position operator is running at 0^o. The other features are described in the same way. Similarly, corresponding sets, $\mathbf{B}_0$, $\mathbf{B}_{45}$, $\mathbf{B}_{90}$ and $\mathbf{B}_{135}$, of feature vectors from gray scale background image B can be obtained as follow:

$$\begin{aligned}
\mathbf{B}_0 &= \begin{bmatrix} BCts_0 & BCor_0 & BEng_0 & BHom_0 & BEpt_0 \end{bmatrix}^T, \\
\mathbf{B}_{45} &= \begin{bmatrix} BCts_{45} & BCor_{45} & BEng_{45} & BHom_{45} & BEpt_{45} \end{bmatrix}^T, \\
\mathbf{B}_{90} &= \begin{bmatrix} BCts_{90} & BCor_{90} & BEng_{90} & BHom_{90} & BEpt_{90} \end{bmatrix}^T, \\
\mathbf{B}_{135} &= \begin{bmatrix} BCts_{135} & BCor_{135} & BEng_{135} & BHom_{135} & BEpt_{135} \end{bmatrix}^T.
\end{aligned} \tag{10}$$

In order to capture the difference between two textures, the absolute difference for each pair features is applied to compute the variations. The absolute difference function is defined as follow:

$$f_{i\alpha} = \left| If_{i\alpha} - Bf_{i\alpha} \right|, \tag{11}$$

where f_i is the i-th component of feature set, α is the angel of position operator running at. The sequent combination of each feature set can generate a new feature vector $\mathbf{F}$:

$$\begin{aligned}
\mathbf{F} = [&Cts_0 \; Cor_0 \; Eng_0 \; Hom_0 \; Ept_0; \\
&Cts_{45} \; Cor_{45} \; Eng_{45} \; Hom_{45} \; Ept_{45}; \\
&Cts_{90} \; Cor_{90} \; Eng_{90} \; Hom_{90} \; Ept_{90}; \\
&Cts_{135} \; Cor_{135} \; Eng_{135} \; Hom_{135} \; Ept_{135}]^T.
\end{aligned} \tag{12}$$

Generally, the features are not equally important. Hence, calculating different weights for each component in $\mathbf{F}$ is necessary. In this paper, ten successive images are arbitrarily chosen to calculate the weights, and the corresponding weighting matrix $\mathbf{W}$ can be calculated [23]. The weighting matrix is defined as follow:

$$\mathbf{W} = \begin{bmatrix} w_{11} & 0 & \cdots & 0 \\ 0 & w_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{nn} \end{bmatrix} \tag{13}$$

$$w_{ii} = \frac{1}{\sigma_i^2 \sum_i^n \sigma_i^{-2}}, \tag{14}$$

where σ_i is the standard deviation value of i-th feature. A scalar *Diff* that represents the texture difference of an input image to the background image can be obtained by computing the one norm of $\mathbf{WF}$:

$$Diff = \left\| \mathbf{WF} \right\|_1. \tag{15}$$

2.4. Traffic parameters

2.4.1. Scene calibration. A procedure of scene calibration is necessary because the configuration of each intersection is different. Subjectively, an image I_{mp} with most pedestrian on the crosswalk and a severely jammed image I_{jam} have been chosen to calibrate the scene. Then, the scalar $N_{\max}$, NV_{jam} and $Diff_{\text{jam}}$, can be obtained from 2.2 and 2.3.For any incoming image of crosswalk, the scalar of pedestrian evaluation PL is defined as follow:

$$PL = \begin{cases} 1, & \text{if } N \geq N_{\max}, \\ \dfrac{N}{N_{\max}}, & \text{if } N < N_{\max}. \end{cases} \quad (16)$$

where N is the amount of pedestrian on the crosswalk.

For incoming input image of road , the two features of vehicular flow estimation are defined as follow:

$$NV_{ratio} = \frac{NV}{NV_{jam}}, \qquad Diff_{ratio} = \frac{Diff}{Diff_{\text{jam}}}. \quad (17)$$

2.4.2. Fuzzy vehicular flow estimation system. The fuzzy vehicular flow estimation system (FVFES) uses fuzzy inference system to infer a value between 0 and 1, which represents slightest and heaviest vehicular flow respectively. The inputs of the fuzzy vehicular flow estimation system are NV_{ratio} and $Diff_{ratio,}$ and the output is IV (QL) of range from 0 to 1. Scalar IV means the heavy degree of incoming vehicular flow in green signal approach. Similarly, scalar QL means the heavy degree of queuing length in red signal approach. The membership functions of NV_{ratio}, $Diff_{ratio}$, and IV (QL) are derived by tuning to the best performance of vehicular flow estimation, as shown in Fig. 7. These are trapezoid and triangular membership functions. The fuzzy rules are listed in Table 1. This fuzzy inference system is accomplished using stander *Mamdani* –type method with COA-defuzzifier (center of area).

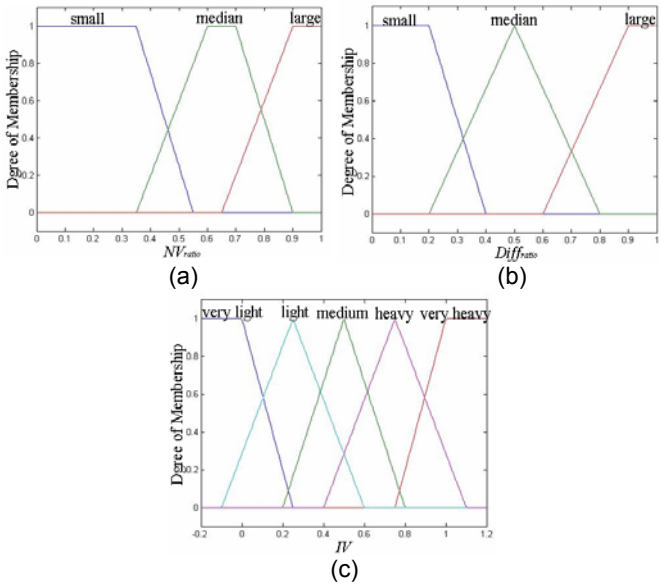


Fig. 7. Memberships of fuzzy vehicular flow estimation system. (a) NV_{ratio}; (b) $Diff_{\text{ratio}}$; (c) IV(QL).

TABLE 1
FUZZY RULES OF FVFES

$Diff_{ratio}$ \ NV_{ratio}	small	median	large
small	very light	light	medium
median	light	medium	heavy
large	medium	heavy	very heavy

3. Fuzzy adaptive traffic signal control

3.1. Control strategy

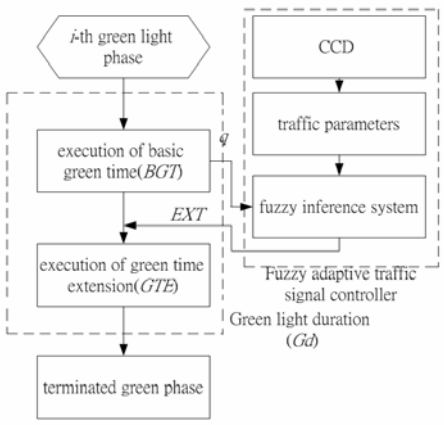


Fig. 8. The flowchart of fuzzy adaptive traffic signal control

In traditional traffic signal control system, a fixed signal switching time is used to manipulate the whole intersection. However, the traditional method will spend a lot time to wait and queue. An adaptive traffic signal control system will be more efficiently than traditional system. Hence, the fuzzy adaptive traffic signal control system (FATSCS) is designed here to gather the traffic information and then to optimize the traffic control system. In this paper, the fixed time of *Webster optimal cycle* [24] is adopted to determine the optimal cycle beforehand and traffic information is used to adjust the green time extension dynamically. The flowchart of FATSCS is shown in Fig. 8. The green light duration can be divided into two parts: basic green time (*BGT*) and green time extension (*GTE*). *BGT* is the minimum green time that can ensure the essential green time in green phase. When the time comes to five seconds before the end of basic green time, control system will make a request to fuzzy inference system to infer a green time extension (*EXT*). Subsequently, control system will continue green phase until the end of green time extension. The green duration of phase i, denoted as Gd_i, can be represented by the following equation:

$$Gd_i = BGT_i + GTE_i, \quad BGT_i = 0.8 \cdot G_{opt}, \quad GTE_i = 1.7 \cdot G_{opt} \cdot EXT_i, \quad (18)$$

where G_{opt} is the Webster optimal green length of phase *i*.

3.2. Fuzzy inference system

The adaptive traffic signal controller uses fuzzy inference system to gather traffic information to infer the green length extension degree. The system demands three inputs: *IV* (incoming vehicular flow in green signal approach), *QL* (vehicular queuing length in red signal approach), and *PL* (pedestrian number) and yields one output, *EXT* (the extension degree of current phase), of range from 0 to 1. Inputs *IV* and *QL* can be obtained from fuzzy vehicular flow estimation system and *PL* can be obtained from the pedestrian estimation system. For deriving the best performance of traffic signal control, we have tried and tuned several combinations of three kinds of most common used membership function type; triangular, trapezoidal, and Gaussian. Then, the best arrangement is that the input components; *IV*, *QL*, and *PL*; adopt the Gaussian type and the output component; *EXT*; adopts triangular type, as show in Fig. 9. The fuzzy rules of fuzzy adaptive traffic signal controller are listed in Table 2. For example, if *PL* is small, *IV* is light, and *QL* is light, then *EXT* is short. In this fuzzy inference system, *Mamdani*-type method is adopted for fuzzy implication and COA (center of area) method is adopted for defuzzification.

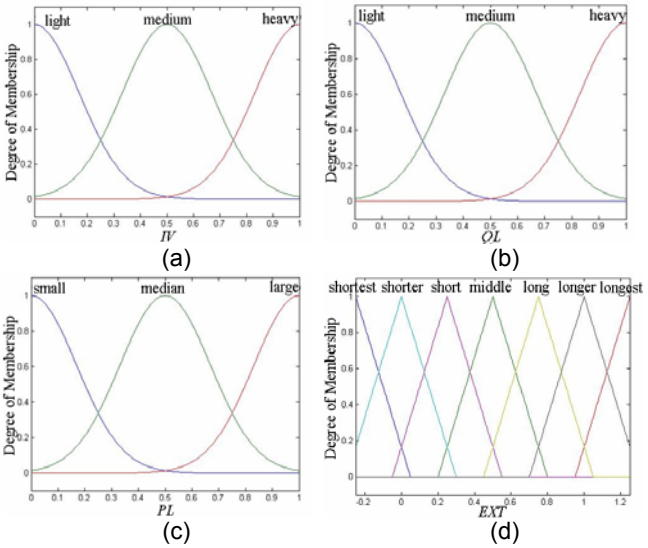


Fig. 9. Memberships of fuzzy adaptive traffic signal controller. (a) *IV*; (b) *QL*; (c) *PL*; (d) *EXT*.

TABLE 2
FUZZY RULES OF FUZZY ADAPTIVE TRAFFIC SIGNAL CONTROLLER

		QL \ IV	light	medium	heavy
PL	small	light	short	middle	long
		medium	shorter	short	middle
		heavy	shortest	shorter	shorter
	median	QL \ IV	light	medium	heavy
		light	middle	long	longer
		medium	short	middle	long
		heavy	shorter	short	middle
	large	QL \ IV	light	medium	heavy
		light	long	longer	longest
		medium	middle	long	longer
		heavy	short	middle	long

4. Experiment results

The experiment in this paper is separated into three parts. The first part is pedestrian estimation. To verify the performance of the proposed system without controversy, the correct number of pedestrian on crosswalk should be counted manually in order to be compared with the result estimated by the system. The second part is vehicular flow estimation. The vehicular flow rate is obtained by fuzzy inference system. Several kinds of road situation are taken to test. To prove the performance of fuzzy adaptive traffic signal control, an intersection traffic simulation model must be established in last part. After all of the parameters are determined off-line, twenty images can be processed per second by the proposed algorithm (where a personal computer with Pentium 4, 2.6 GHz processor is used). All of these experiments and analyses will be detailed in the following subsections.

4.1. Pedestrian estimation system

The pedestrian estimation system is tested by three different crosswalk scenes: 2-lane road, 4-lane road, and 8-lane road. For each scene, 100 frames were randomly chosen for the experiment. All experimental images are 24-bit color image of size 320 by 240. Several test images and their results are shown in Fig. 10. For Fig. 10(a), the number of foreground pixels *p* is 6192, the number of normalized foreground pixels *NP* is 8904, and the estimated pedestrian *N* is 11. For Fig. 10(b), *p* is 3359, *NP* is 8111, and *N* is 10. For Fig. 10(c), *p* is 8827, *NP* is 8237, and *N* is 10. For Fig. 10(d), *p* is 216, *NP* is 352, and *N* is 0. By the way, the figures Fig. 10(a) and Fig. 10(b) are the frame 300 and frame 420 captured by the same camera at the 10^{th} second and the 14^{th} second respectively. Although *p* of Fig. 10(a) is quite different from *p* of Fig. 10(b), the results of *N* are close. Hence the pedestrian positions have changed in this 4-second interval, the results of pedestrian estimation in these two images are closed. This indicates the effects of pedestrian evaluation by using the procedure of normalization. The experimental results show that the pedestrian on crosswalk can be estimated efficiently.

For evaluating the performance of the system, the following equation is taken as the performance index:

$$\sigma_N = \sqrt{\frac{\sum_{i=1}^{n_t} (N_i - N_{MCi})^2}{n_t^2}}, \quad (19)$$

where N_i is the estimated number of pedestrians of ith test frame, N_{MCi} is the manually counted number of pedestrians of ith test frame, n_t is the total number of test images. For 2-lane road, 4-lane road and 8-lane road experiments, σ_N are 1.13, 1.29 and 1.05 respectively. For total test images, σ_N is 1.15. This means that the algorithm will averagely overestimate or underestimate the number of pedestrians by one or two for each scene.

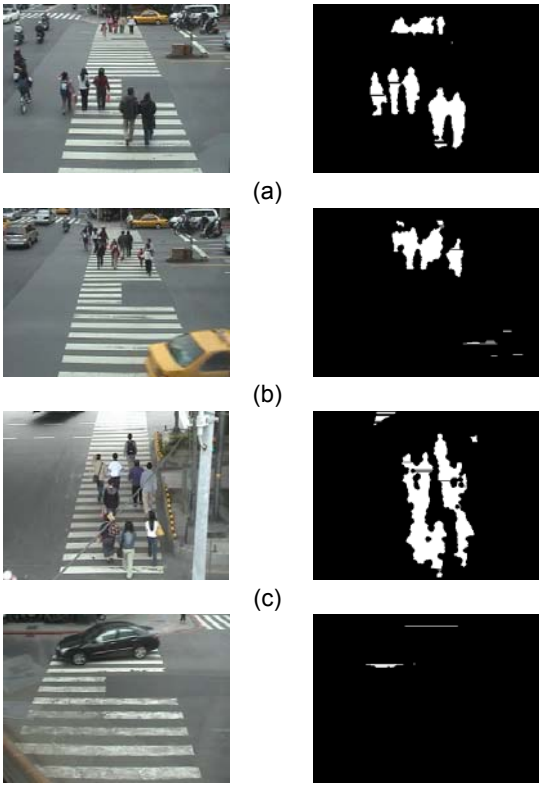


Fig. 10. Some experimental results of pedestrian estimation system with (a) 11 (b) 10 (c) 9 (d) 0 pedestrian in the crosswalk.

4.2. Vehicular flow estimation system

The vehicular flow estimation system is tested by some image sequences taken by road camera and the result of one whole traffic signal cycle is shown in Fig. 11. As it can be seen, NV_{ratio} and $Diff_{ratio}$ vary with the number of vehicle in ROI. Besides, $Diff_{ratio}$ also varies with the distance between vehicle and stop line. The estimated vehicular flow IV is then inferred by using these two features. At the beginning, the traffic signal turned green and the vehicular queue was discharging from the beginning to the 20th second (frame 1 to frame 600); therefore IV decreases in this time interval. From the 20th second to the 45th second (frame 600 to frame 1350), the vehicular queue is totally discharged and a little amount of vehicle enter ROI region from the far side. In this time interval, the estimated IV indicated the vehicular flow was very light. At the 45th second (frame 1350), the traffic signal turned red; therefore the vehicular queues and IV increase simultaneously. From the 70th second up to the end (frame 2100 to frame 2500), the queue of the approach is very long and the estimated IV indicated the vehicular flow was very heavy. From

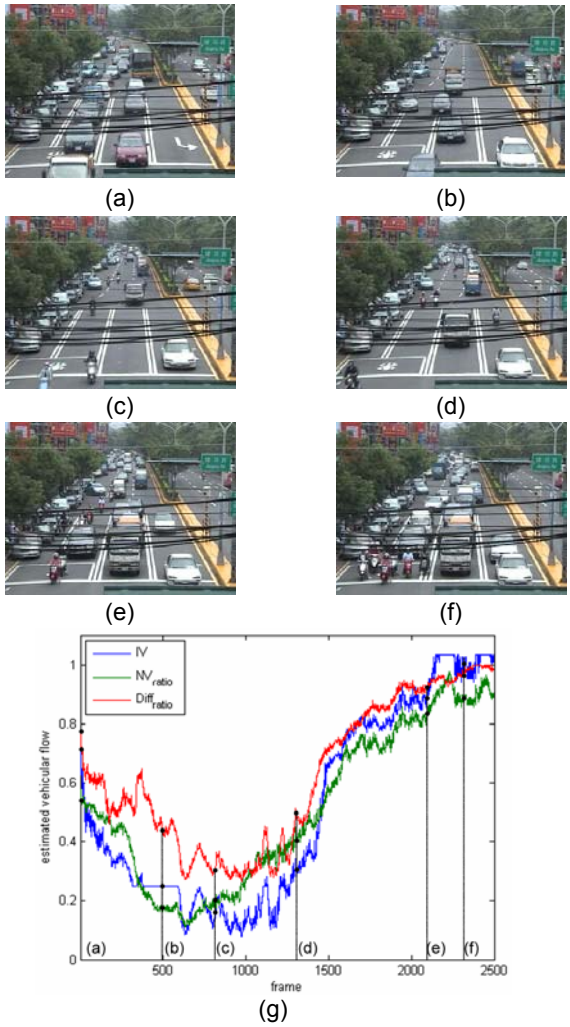


Fig. 11. Illustrate image sequences (a) frame 1, (b) frame 500, (c) frame 800, (d) frame 1350, (e) frame 2100, and(f) frame 2300 and (g) image sequence features NV_{ratio}, $Diff_{ratio}$, and the estimated vehicular flow IV .

comparing with Fig. 11(b) and Fig. 11(c), it is easy to observe that the number of vehicle is almost the same, but the distribution of vehicle is quite different. By conjugation with Fig. 11(g), the value of vehicular pixel (NV_{ratio}) is closed, but the value of texture analysis ($Diff_{ratio}$) is quite different. In Fig. 11(b), the distances

between vehicles and stop line are short than Fig 11(c) so the estimated result *IV* of Fig. 11(b) is larger than Fig. 11(c). Hence, the experiment result indicates that the vehicular estimation system is dependent on not only the number of vehicle, but also each vehicle position.

4.3. Simulation result of fuzzy adaptive traffic signal control

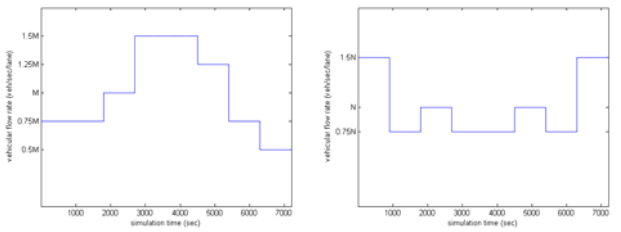
To validate the effectiveness and robustness of the proposed fuzzy adaptive traffic signal controller, six experimental examples for an isolated four-leg intersection are demonstrated. The simulation time is set to be two hours. The traffic signal is controlled in two phases with 3 seconds all red time. Assuming that the intersection has three lanes in each approach (l_{num}=3), the saturation flow rate of each lane(v_{sat_flow}) and crosswalk(p_{sat_flow}) are defined as 1 pcu/sec-lane(pcu: passenger car unit) and 10 ped/sec respectively, the maximum amount of vehicle on each lane(vl_{max}) and pedestrian on crosswalk(pc_{max})are defined as 20 pcu/lane and 30 pedestrian respectively, the vehicle departure rate μ is modeled by an s-type function as follow:

$$\mu(t,a)=\begin{cases}0, & \text{if } t<a,\\ 2\left(\dfrac{t-a}{l_s}\right)^2, & \text{if } a\le t\le a+\dfrac{l_s}{2},\\ 1-2\left(\dfrac{t-a-l_s}{l_s}\right)^2, & \text{if } a+\dfrac{l_s}{2}\le t\le a+l_s,\\ 1, & \text{if } t>a+l_s,\end{cases} \tag{20}$$

where t represents time, a represents the time when the signal turned green, and l_s here is set to be 10 seconds. The simulation traffic flow data are generated by the computer according to average arrival rates and specific probability distribution. The vehicular arrival rate (vr) obeys Poisson distribution and the pedestrian arrival rate (pr) obeys exponential distribution with fixed mean equals 0.5ped/sec. For example, the vehicular flow patterns with the average arrival rates M and N, which represent the north and west directions respectively, are shown in Fig. 12. For simplicity, the traffic pattern in the north and south directions are assumed the same. Similarly, the patterns in the east and west directions are assumed the same. The average arrival rates of six environments for experiments are shown in Table 3.

TABLE 3
THE SIX ENVIRONMENTS FOR THE EXPERIMENTS

scene	1	2	3	4	5	6
M	0.12	0.2	0.25	0.2	0.25	0.25
N	0.1	0.1	0.1	0.15	0.15	0.2



(a) (b)

Fig. 12. The vehicle flow patterns in the (a) north and (b) west directions.

In our signal control system simulation, the transfer function of each input vector is defined as follow:

$$IV\begin{cases}q_t=q_{t-1}+vr_t-\mu(t,a)\cdot v_{sat\_flow}\cdot l_{num},\\ IV_t=\begin{cases}1, & \text{if } q_t\ge vl_{\max}\cdot l_{num},\\ \dfrac{q_t}{vl_{\max}\cdot l_{num}}, & \text{if } q_t<vl_{\max}\cdot l_{num},\end{cases}\end{cases} \tag{21}$$

$$QL\begin{cases}q_t=q_{t-1}+vr_t,\\ QL_t=\begin{cases}1, & \text{if } q_t\ge vl_{\max}\cdot l_{num},\\ \dfrac{q_t}{vl_{\max}\cdot l_{num}}, & \text{if } q_t<vl_{\max}\cdot l_{num},\end{cases}\end{cases} \tag{22}$$

$$PL\begin{cases}p_t=p_{t-1}+pr_t-p_{sat\_flow},\\ PL_t=\begin{cases}1, & \text{if } p_t\ge pc_{\max},\\ \dfrac{p_t}{pc_{\max}}, & \text{if } p_t<pc_{\max},\end{cases}\end{cases} \tag{23}$$

where q_t and p_t are the amount of vehicle on the road and pedestrian on crosswalk respectively.

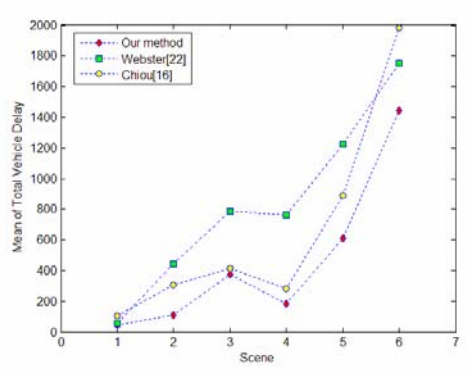
The control performance of the proposed method is compared with two famous adaptive signal control methods: Webster optimal cycle fixed timing [24] and generic fuzzy logic traffic signal controller [18]. Webster optimal cycle method is a well-know to deriver optimal cycle and split green allocation, which major concerns saturation flow rate and incoming vehicle flow rate. Generic fuzzy logic traffic signal controller is an iterative evolution algorithm for selecting the logic rules and tuning the membership functions for traffic signal control with genetic fuzzy logic controller. The GFLC model employs incoming vehicle flow rate and queue length as state variables and extension of green time as control variable. The input value; incoming vehicle flow and queue length; of these comparison methods can be derived from Eq.(21) and Eq.(22), and the minimum green time of GFLC is set at 0.8times of Webster optimal green length. For each experimental environment, experiments have been done by five times repeatedly to reduce the experimental uncertainties. Three kinds of delay: total vehicle delay (TVD), total

pedestrian delay (TPD), and total delay (TD), are computed respectively. The performances of mean and standard deviation of each kind of delay are compared.

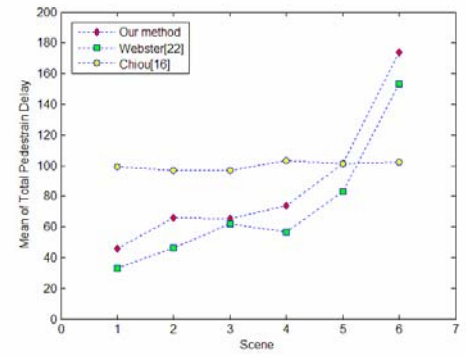
The experimental results are summarized in Table 4 and illustrated in Fig. 13. Under different experimental conditions, the best performance of each evaluation item is marked in boldface. As it can be seen, the proposed method has the best control performance in vehicle delay, but generally the control performance of pedestrian delay is second to Webster's method. The reason of this phenomenon is due to the pedestrians' discharging mode. Unlike vehicular queues, pedestrian queues will be totally discharged in a cycle. Therefore, the total pedestrian delay is bound up with the cycle length; the shorter thecycle length is, the smaller the total pedestrian delay is. When considering the vehicles and pedestrians simultaneously, the proposed method has the best control performance. Speaking of the control stability, the standard deviation of TPD of each method is relatively low, thus the control stability is dominated by the standard deviation of TVD. As a result, the proposed method has lowest standard deviation of TVD and TD in each experimental case. Hence, the proposed fuzzy adaptive traffic signal controller has the best control stability.

TABLE 4
EXPERIMENTAL RESULTS OF THE SIMULATION

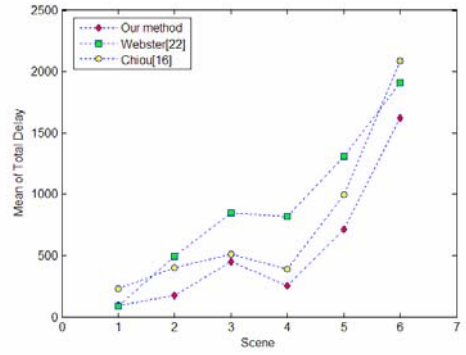
case	Method	Mean of TVD	Mean of TPD	Mean of TD	Std. of TVD	Std. of TPD	Std. of TD
1	Proposed method	**48.3**	45.9	94	**1.5**	0.35	**1.74**
	Webster[24]	55.8	**33.3**	**89**	3.6	**0.26**	3.81
	Chiou[18]	105	99.2	229	22.7	0.81	22.5
2	Proposed method	**111**	65.7	**176**	**26.6**	1.49	**27.5**
	Webster[24]	444	**46.3**	491	53.6	**0.49**	53.4
	Chiou[18]	304	96.9	401	54.7	1.17	54.8
3	Proposed method	**375**	65.2	**450**	**39.9**	0.57	**40.3**
	Webster[24]	788	**62.0**	850	69.9	**0.28**	69.6
	Chiou[18]	414	96.6	514	51.6	1.01	51.2
4	Proposed method	**182**	73.7	**256**	**27.9**	1.54	**26.7**
	Webster[24]	760	**56.6**	815	61.5	**0.31**	63.8
	Chiou[18]	283	103	387	33.6	1.49	34.2
5	Proposed method	**612**	101	**713**	**72.7**	1.76	**73.2**
	Webster[24]	1225	**82.9**	1308	78.2	**0.58**	78.2
	Chiou[18]	888	101	995	125	0.94	133
6	Proposed method	**1445**	174	**1619**	**84.1**	2.04	**84.6**
	Webster[24]	1753	153	1907	98.1	**1.70**	99.2
	Chiou[18]	1984	**102**	2088	129.6	2.04	130



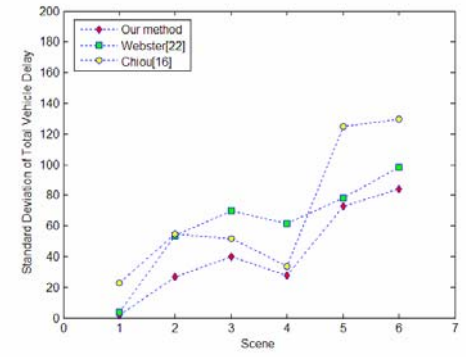
(a)



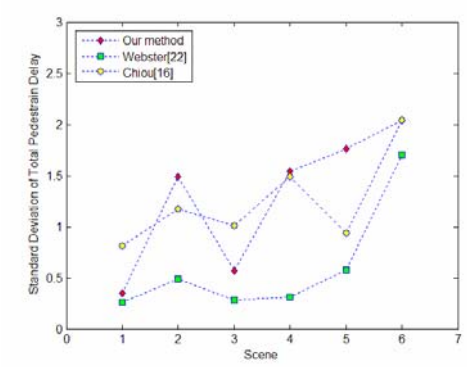
(b)



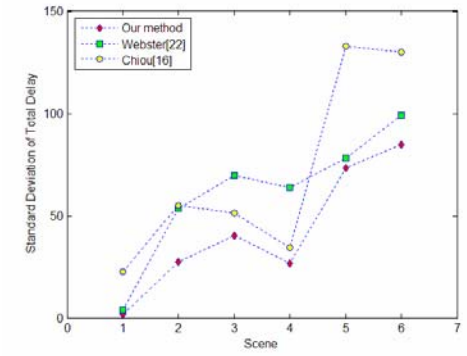
(c)



(d)



(e)



(f)

Fig. 13. Illustrate the performance comparison of three adaptive signal control system. (a) mean of total vehicle delay (b) mean of total pedestrian delay (c) mean of total delay (d) standard deviation of total vehicle delay (e) standard deviation of total pedestrian delay (f) standard deviation of total delay.

5. Conclusion

The main objectives of traffic signal controller design method are fast response to traffic flow change and efficiently traffic management. The vision based traffic follow estimation and adaptive controller are under this goal. A vision-based traffic flow estimation system has been adopted in order to achieve a wide observation field, and the optimal traffic signal phase cycle is obtained by using an adaptive fuzzy controller. The experimental results demonstrate that the proposed image processing algorithms is able to estimate the traffic flow efficiently and the proposed fuzzy adaptive traffic signal controller has good control performance and stability. In the future, the control system will be considered the neighbor intersections traffic flow and then to involve the presented system into a block area adaptive traffic signal control system. Even a local area adaptive traffic signal control system can be looking forward to base on presented approach. Besides, applying this system to the continuous signalized intersections in a corridor or network is worthy of exploring.

6. Acknowledgments

This work was supported in part by the National Science Council, Taiwan, R.O.C., under Grant NSC 96-2221-E-009-238.

7. References

[1] F. Y. Wang, P. B. Mirchandani, and N. Zheng, “Advances and Trends in Research and Development of Intelligent Transportation Systems: An Introduction to the Special Issue,” IEEE Trans. ITS, vol. 5, no. 4, pp. 222-223, Dec. 2004.

[2] B. Maurin, O. Masoud, and N. P. Papanikolopoulos, “Tracking All Traffic: Computer Vision Algorithms for Monitoring Vehicles, Individuals, and Crowds,” IEEE Trans. Robotics & Automation, vol. 12, no. 1, pp. 29-36, March 2005.

[3] C. J. Pai, H. R. Tyan, Y. M. Liang, H.Y. Mark Liao, and S.W. Chen, “Pedestrian Detection and Tracking at Crossroad,” Proc. IEEE int’l Conf. Image Processing, vol. 2, pp. 101-104, Sept. 2004.

[4] I. Haritaoglu, D. Harwood, and L. S. Davis, “W4: Real-Time Surveillance of People and Their Activities,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 809-830, Aug. 2000.

[5] T. Zhao and R. Nevatia, “Tracking Multiple Humans in Complex Situations,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 9, pp. 1208-1221, Sept. 2004.

[6] T. Zhao and R. Nevatia, “Car Detection in Low Resolution Aerial Image,” Proc. IEEE int’l Conf. Computer Vision, vol. 1, pp. 710-717, July 2001.

[7] G. L. Foresti, V. Murino, and C. Regazzoni, “Vehicle Recognition and Tracking from Road Image Sequences,” IEEE Trans. on Vehicular Technology, vol. 48, no. 1, pp. 301-318, Jan. 1999.

[8] Z. Sun, G. Bebis, and R. Miller, “On-road Vehicle Detection Using Evolutionary Gabor Filter Optimization,” IEEE Trans. On ITS, vol. 6, pp. 125-137, June 2005.

[9] S. Kamijo, “Spatio-Temporal MRF model and its Application to Traffic Flow Analyses,” ICDE Workshops 2005, pp. 1203-, April 5-8, 2005.

[10] T. Nishida, S. Kamijo, K. Ikeuchi, and M. Sakauchi, “Automated system of acquiring and visualizing traffic event statistics from traffic images,” ICME 2001, pp. 269-, Aug. 22-25, 2001.

[11] S. F. Lin, J. Y. Chen, and H. X. Chao, “Estimation of Number of People in Crowded Scenes Using Perspective Transformation,” IEEE Trans. on Systems, Man, and Cybernetics, vol. 31, no. 6, pp. 645-654, Nov. 2001.

[12] T. Coianiz, M. Boninsegna, and B. Caprile, “A Fuzzy Classifier for Visual Crowding Estimates,”

IEEE int'l Conf. Neural Networks, vol. 2, pp. 1174-1178, Nagoya, Japan, 1996.
[13] S. Y. Cho, T. W. S. Chow, and C. T. Leung, "A Neural-Based Crowd Estimation by Hybrid Global Learning Algorithm," IEEE Trans. on Systems, Man and Cybernetics, vol. 29, no. 4, pp. 535-541, Aug. 1999.
[14] G. Beauchamp-Baez, E. Rodriguez-Morales, and E.L. Muniz-Marrero, "A Fuzzy Logic Based Phase Controller for Traffic Control," Proc. Sixth IEEE int'l Conf. Fuzzy Systems, vol. 3, pp. 1533-1539, July 1997.
[15] J. Nittymäki, "Installation and experiences of field testing a fuzzy logic controller," European J. of Operational Research, vol. 7C, pp. 273-281, 2001.
[16] J. Nittymäki, T. Kosonen, and R. Nevala, "Fuzzy Traffic Signal Control in Major Arterials," Proc. IEEE int'l Conf. Intelligent Transportation System, pp. 346-351, Aug. 25-29, 2001.
[17] C. H. Chou and J. C. Teng, "A Fuzzy Logic Controller for Traffic Junction Signals," Information Sciences, vol. 143, pp. 73-97, 2002.
[18] Y. C. Chiou and L. W. Lan, "Adaptive Traffic Signal Control with Iterative Genetic Fuzzy Logic Controller," Proc. IEEE int'l Conf. Networking, Sensing, and Control, vol. 1, pp. 287-292, March 2004.
[19] D. A. Forsyth and J. Ponce, Computer Vision: A Modern Approach. Upper Saddle River, NJ: Prentice Hall, 2003.
[20] F. Lv, T. Zhao, and R. Nevatia, "Self-Calibration of a Camera from Video of a Walking Human," Proc. Int'l Conf. Pattern Recognition, vol. 1, pp. 562-567, Quebec, Canada, 2002.
[21] R. M. Haralick, "Statistical and Structural Approaches to Texture," Proceedings of the IEEE, vol. 67, no.5, pp. 786-804, May 1979.
[22] R. C. Gonzalez and R. E. Woods, Digital Image Processing. Upper Saddle River, NJ: Prentice Hall, 2002.
[23] M. Friedman and A. Kandel, Introduction to Pattern Recognition. London, UK: Imperial College, 1999.
[24] F. V. Webster, Traffic Signal Settings. London, UK: Road Research Laboratory, Technical Paper, no. 39, 1958.

Sheng-Fuu Lin (S'84–M'88) was born in Tainan, R.O.C., in 1954. He received the B.S. and M.S. degrees in mathematics from National Taiwan Normal University in 1976 and 1979, respectively, the M.S. degree in computer science from the University of Maryland, College Park, in 1985, and the Ph.D. degree in electrical engineering from the University of Illinois, Champaign, in 1988. Since 1988, he has been on the faculty of the Department of Electrical and Control Engineering at National Chiao Tung University, Hsinchu, Taiwan, where he is currently a Professor. His research interests include image processing, image recognition, fuzzy theory, automatic target recognition, and scheduling.

Shih-Che Chien was born in Chiayi, R.O.C., in 1978. He received the B.E. degree in electronic engineering from the Nation Chung Cheng University, in 2002. He is currently pursuing the M.E. and Ph.D. degree in the Department of Electrical and Control Engineering, the National Chiao Tung University, Hsinchu, Taiwan. His current research interests include image processing, image recognition, fuzzy theory, 3D image processing, intelligent transportation system, and animation.

Kuo-Yu Chiu was born in Hsinchu, R.O.C., in 1981. He received the B.E. degree in electrical and control engineering from National Chiao Tung University, Hsinchu, Taiwan, R.O.C, in 2003. He is currently pursuing the Ph. D. degree in the Department of Electrical and Control Engineering, the National Chiao Tung University, Hsinchu, Taiwan. His current research interests include image processing, face recognition, face replacement, intelligent transportation system, and machine learning.

Received date: November 6, 2007
Received in revised forms: July 24, 2008 and January 20, 2009
Accepted date: January 20, 2009
Editor: Shunsuke Kamijo