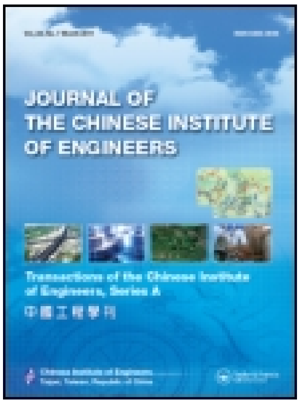

Journal of the Chinese Institute of Engineers
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tcie20

Color image vehicular detection systems with and without fuzzy neural network: A comparison

Lawrence W. Lan [a], April Y. Kuo [b] & Yeh-Chieh Huang [b]
[a] Institute of Traffic and Transportation, National Chiao Tung University, Taipei, Taiwan 100, R.O.C. Phone: 886-2-23494958 Fax: 886-2-23494958 E-mail:
[b] Institute of Traffic and Transportation, National Chiao Tung University, Taipei, Taiwan 100, R.O.C.
Published online: 03 Mar 2011.

To cite this article: Lawrence W. Lan, April Y. Kuo & Yeh-Chieh Huang (2003) Color image vehicular detection systems with and without fuzzy neural network: A comparison, Journal of the Chinese Institute of Engineers, 26:5, 659-670, DOI: 10.1080/02533839.2003.9670819

To link to this article: http://dx.doi.org/10.1080/02533839.2003.9670819

Journal of the Chinese Institute of Engineers, Vol. 26, No. 5, pp. 659-670 (2003)

COLOR IMAGE VEHICULAR DETECTION SYSTEMS WITH AND WITHOUT FUZZY NEURAL NETWORK: A COMPARISON

Lawrence W. Lan*, April Y. Kuo, and Yeh-Chieh Huang

ABSTRACT

This paper develops a color image vehicular detection (CIVD) system in which background differencing technique is employed to detect whether a vehicle passes through the detecting points equally spaced out on a pseudo line detector. Two methods (interval search and regression) are tried to determine the optimal crisp threshold values to cope with various lighting conditions. To compare the detection performance with and without incorporating a fuzzy neural network (FNN), a three-layer FNNCIVD system is further designed with trapezoidal membership function and network parameters trained by back propagation algorithm. Under different environments (freeway and urban street) with various lighting conditions (daytime and nighttime), it is found that the detection success rates for interval-search CIVD and regression CIVD are about the same. However, both perform worse than the FNNCIVD system in which about 90% success rates are reported with seven detection points. Compared with the interval-search CIVD system, the FNNCIVD system can increase the success rates at a range of 14% to 22% on the freeway mainline and 18% to 26% on the urban street. It is also found that daytime detection performance is slightly better than nighttime detection. Possible reasons for missed detection and false detection are discussed.

Key Words: color image vehicular detection (CIVD), fuzzy neural network (FNN), background differencing technique, back propagation algorithm.

I. INTRODUCTION

Advanced traffic control and management relies heavily on the collection of accurate traffic flow data. In recent years, more and more traffic parameters have been automatically collected by video image detectors rather than by conventional techniques such as loops and magnetic detectors. The major shortcomings for the conventional detectors may include their limitations on the accurate assessment of traffic parameters, small detection zones, and placements without flexibility (Michalopulos, 1991). Additionally, data collected by such conventional vehicular detectors cannot be applied to vehicle tracking, incident detection (Washburn and Nihan, 1999), or vehicle movement monitoring within a junction (Fathy and Siyal, 1995).

Traffic detection with video image processing may improve the shortcomings mentioned above. Recently, extensive research and development efforts have been devoted to image processing techniques applied to traffic data collection and analysis. More applications of image processing to vehicle classification, tracking and incident detection have been reported (for instance, Hoose, 1992; Liao, 1993; Coifman *et al*., 1998; Chang, 1999; Shu, 1999).

Methods of applying image processing to vehicular detection, in general, include blob detection, pattern recognition and background differencing (Li *et al*., 2002; Alejandro *et al*., 2002). Blob detection does not perform well under poor weather conditions because not all vehicles are brighter or darker than the background road surface (Blosseville *et al*., 1989). Neither does pattern recognition work properly when vehicles in the detection zone do not fit well into the

*Corresponding author. (Tel: 886-2-23494958; Fax: 886-2-23494953; Email: lawrencelan@mail.nctu.edu.tw)

The authors are with the Institute of Traffic and Transportation, National Chiao Tung University, Taipei, Taiwan 100, R.O.C.

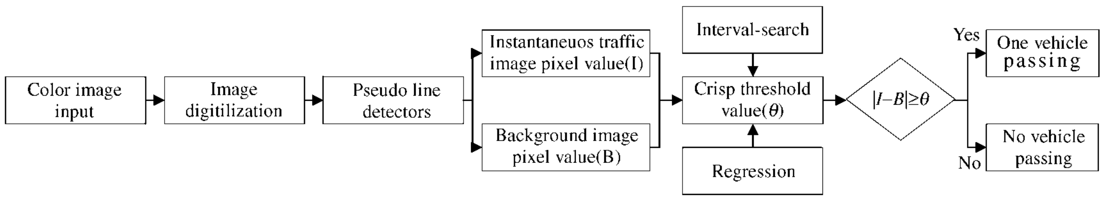

Fig. 1 Detection logic for a CIVD system

defined templates (Dickson and Wan, 1989). Both blob detection and pattern recognition methods need more computation time than the most common and simple approach used in traffic image detection -- the background differencing technique (Fathy and Siyal, 1995). Background differencing technique is based on a pixel-by-pixel comparison between background frames and instantaneous frames of traffic scenes (Dickinson and Waterfall, 1984a, 1984b). This technique tends to generate successful detection in the daytime with fair weather or lighting conditions. However, it can also lose detection accuracy near dusk or dawn or in bad weather because of its sensitivity to ambient lighting.

With learning ability and capabilities of dealing with uncertainties (for instance, Bullock *et al.*, 1993; Buckley and Hayashi, 1994; Dougherty, 1995; Mantri and Bullock, 1995; Chiou and Lan, 1997; Jouseau and Dorizzi, 1999), a fuzzy neural network (FNN) may cope with the variations of ambient lighting. A color-level image is composed of red (R), green (G) and blue (B) pixels; thus it provides more information than a gray-level one (Buluswar and Draper, 1998; Kumar *et al.*, 2002). Consequently, it is presumed that a color-based image processing system incorporating FNN can accommodate more environmental changes than the same system without FNN. The presumption motivates our study.

In this study, a color image vehicular detection (CIVD) system, which simply utilizes a pseudo line detector and background differencing technique, is designed. After that, an FNN structure is further constructed and incorporated into this CIVD system. Traffic scenes on a freeway mainline and an urban street under different lighting conditions are collected and tested by these two systems. Detection “success” rates of these two systems are compared and possible reasons for “missing” and “false” detections are addressed. This paper is organized as follows. Section II elaborates a CIVD system without FNN. Section III further discusses the incorporation of an FNN into this CIVD system. Section IV conducts and compares field experiments on these two systems. Section V summarizes the findings.

II. DESIGN OF THE CIVD SYSTEM

In this paper, the CIVD system mainly contains three modules -- image digitalization, pseudo line detectors allocation, and vehicle detection as depicted in Fig. 1 (Lan and Kuo, 1999). Analog traffic scenes are taken by the video camera and digitized by the image grabber. A pseudo line detector composed of several detection points is then placed across an arbitrary traffic lane on the monitor. The difference of pixel values (R, G, B) at each detection point between instantaneous traffic scenes and background images is calculated in the vehicle detection module. If the differencing value is greater than a designated crisp threshold value, it is identified as a vehicle passing through the line detector. The three modules of this CIVD system are explained in-depth as follows:

1. Image Digitalization Module

This module is to convert analog images into digital ones with an image grabber. After digitalization, video images can then be shown on the monitor and processed by the computer.

2. Pseudo Line Detector Module

This module is to define two-end coordinates of a pseudo line detector on the monitor. When creating a pseudo line detector, the right- and left-end coordinates are directly input on the screen in an interactive way. A pseudo line is in effect composed of several detection points, which actually act as the instrument for vehicular detection. Each detection point automatically reads R, G, B pixel values of the background image and then reads R, G, B pixel values of the instantaneous traffic images at a fixed time interval. The difference of pixel values between background and traffic images is the input to the vehicle detection module.

3. Vehicle Detection Module

This module is to determine whether there is a vehicle passing through the pseudo line detector. The

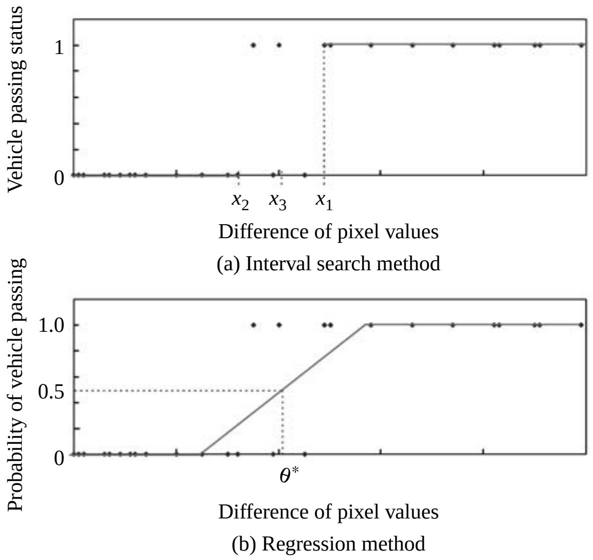

Fig. 2 Determination of optimal crisp threshold values for a CIVD system

difference of pixel values is calculated every one-tenth second. In this paper, different optimal crisp threshold values corresponding to various lighting conditions are determined either by an interval search method or by a regression method, which are depicted in Figs. 2(a) and 2(b). Fig. 2(a) demonstrates the relationship between differencing of pixel values and vehicle passing status, in which X-axis represents pixel values differencing and Y-axis is the status of vehicle passing ("0" represents a status of no vehicle passing; "1" indicates a status of vehicle passing). In Fig. 2(a), those differences of pixel values greater than the upper threshold (x_1) are found to be vehicles passing and those less than the lower threshold (x_2) are no vehicles passing. In between these two thresholds, there could be either a vehicle passing or no vehicle passing.

The interval search method can be explained as the following steps:

Step 1. Identify from the videotape the upper (x_1) and lower (x_2) thresholds as indicated in Fig. 2(a). If the difference between these two thresholds is lower than 5 pixels, go to step 4. Otherwise go to step 2.

Step 2. Find the middle value of these two thresholds, $x_3=(x_1+x_2)/2$.

Step 3. Apply x_1, x_2, x_3 as the crisp thresholds to the detection logic in Fig. 1 and compare their detection results. Choose any two thresholds with higher detection success rates as the new upper and lower thresholds. Go to step 1.

Step 4. Out of the upper and lower thresholds, select the one with higher success rate as the optimal crisp threshold value (θ). Stop.

The regression method can be formulated as follows:

$$y_i=\beta_0+\beta_1 x_i+\mu_i$$

where

$$y_i = \begin{cases} 1 & \text{if a vehicle is passing through the detection point} \\ 0 & \text{if no vehicle is passing through the detection point} \end{cases}$$

x_i= difference of pixel values (ΔR or ΔG or ΔB)

μ_i= independently distributed random variable with 0 mean

Since y_i can take on only binary values, 1 and 0, we can describe the probability distribution of y_i by letting $P_i=Prob(y_i=1)$ and $1-P_i=Prob(y_i=0)$. Then a linear probability model can be expressed in the following form, which allows the dependent variable to be interpreted as a probability. The optimal crisp threshold value (θ^*) is determined by letting P_i=0.5 as shown in Fig. 2(b).

$$P_i = \begin{cases} 1 & \text{when } \beta_0+\beta_1 x_i \geq 1 \\ \beta_0+\beta_1 x_i & \text{when } 0 < \beta_0+\beta_1 x_i < 1 \\ 0 & \text{when } \beta_0+\beta_1 x_i \leq 0 \end{cases}$$

III. DESIGN OF THE FNNCIVD SYSTEM

In this paper, the FNNCIVD system is composed of four modules -- image digitalization, pseudo line detectors allocation, fuzzy neural network and vehicle detection as depicted in Fig. 3. The detection logic of this FNNCIVD system is very similar to that of the CIVD system except that the background differencing technique is further incorporated with a fuzzy neural network module in which the most appropriate threshold values corresponding to various lighting conditions are determined. The difference of pixel values at each detection point between instantaneous traffic and background images is calculated in the vehicle detection module. If the difference value is greater than a trained threshold value, it will be identified as a vehicle passing. Because image digitalization and pseudo line detector modules for both CIVD and FNNCIVD systems are exactly the same, only FNN module and vehicle detection module are narrated in detail as follows.

1. FNN Module

This module is to construct and train a fuzzy

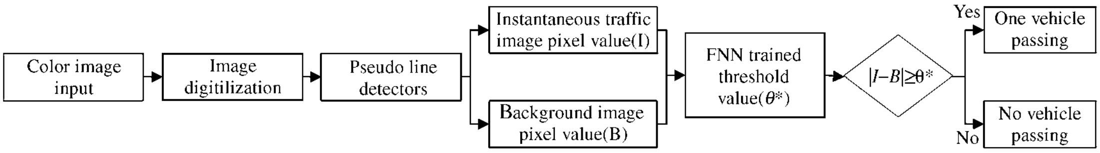

Fig. 3 Detection logic for a FNNCIVD system

Fig. 4 Interface for network training

neural network. Appropriate network parameters can be obtained with back propagation algorithm. Since background pixel values may change over time under different lighting conditions, each lighting condition requires a specific training set. Therefore, different fuzzy neural networks must be trained under different lighting conditions (daytime, nighttime, etc).

In this paper, the training set, collected by the pseudo line detector module, is composed of 1,000 to 1,800 training examples. Each training example contains an input vector (differences of pixel values) and an output vector (binary values representing vehicles passing or not). Fig. 4 demonstrates the designed training interface where one can input desired network parameters. To generate training examples, we simply play the video and press the button "vehicle passing?" at the moment when we see a vehicle-front hitting the pseudo line detector, and press the same button again as we see that vehicle-rear leaving the detector. Repeat such examination until a satisfactory number of training examples are obtained.

A three-layer fuzzy neural network with q detecting points is depicted in Fig. 5. To explain the layers' operation, let the superscript of a notation represent the layer and the subscript indicate the node. The first layer, membership layer, is expressed as:

$$o_j^1 = f_j(u_j^1) = \mu_j(x_j^1)$$

$$= \begin{cases} 0 & \text{for } x_j^1 \le a_j^1 \\ \dfrac{x_j^1 - a_j^1}{b_j^1 - a_j^1} & \text{for } a_j^1 < x_j^1 \le b_j^1 \qquad \text{for } j=1\sim J \\ 1 & \text{for } x_j^1 > b_j^1 \end{cases}$$

where

o_j^1 = the output value of j^{th} node at layer one

x_j^1 = the input value of j^{th} node at layer one

a_j^1 = parameter of trapezoidal membership function

b_j^1 = parameter of trapezoidal membership function

The main function for this layer is to fuzzify the input values by utilizing membership function and then determine membership degrees of input variables. Since each detection point is composed of

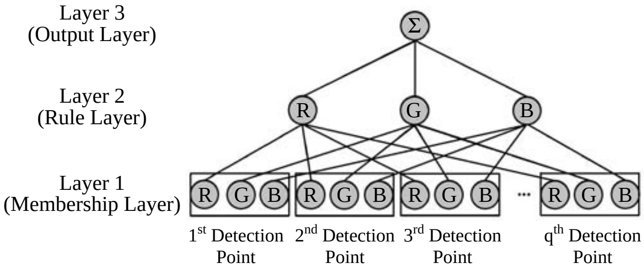

Fig. 5 A fuzzy neural network with q detection points

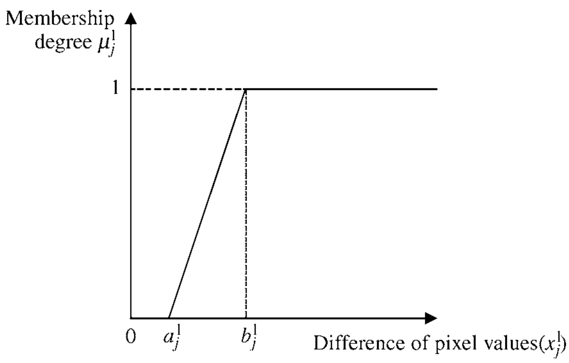

Fig. 6 Trapezoidal membership function

R, G, B nodes, the number of nodes at this layer is $3q$ (q is the number of detection points). Trapezoidal membership function as shown in Fig. 6 is utilized because its shape can correspond to the situations of whether or not one vehicle is passing. A difference of pixel value (x_j^1) less than or equal to the left threshold value (a_j^1) implies that no vehicle is passing through the detector. If x_j^1 is greater than the right threshold value (b_j^1), it is judged as one vehicle passing. If $a_j^1<x_j^1<b_j^1$, the corresponding membership degree must be first calculated and the training algorithm is then employed to identify whether there exists a vehicle.

Rule layer, the second layer, is to establish various rules of fuzzy inference to obtain a reasonable output. The fuzzy rules can be stated as "IF the difference of R or G or B pixel values varies substantially, THEN a signal of vehicle passing is identified at the detection points." The nodes at this layer will perform a summation operation as follows.

$$o_k^2 = f(u_k^2) = \sum_{j=1}^{J} w_{jk}^2 \cdot x_{jk}^{2*} \quad \text{for } j=1\sim J,\ k=1\sim K$$

where

$$x_{jk}^{2*} = x_{jk}^2 z_j$$

$$w_{jk}^2 = 1,\ \forall j \text{ and } k$$

Table 1 An illustration of fixed values distribution

Number of detection points	Number of nodes layer two	Fixed value (z_j)	Nodes (j)
3	9	0.25	1,2,3,7,8,9
		0.5	4,5,6
5	15	0.1	1,2,3,13,14,15
		0.2	4,5,6,10,11,12
		0.5	7,8,9
7	21	0.05	1,2,3,19,20,21
		0.1	4,5,6,16,17,18
		0.2	7,8,9,13,14,15
		0.3	10.11.12

Since each detection point is composed of R, G, B nodes, a pseudo line detector with q detection points will have $3q$ nodes. In this paper, q=3, 5, 7 points are experimented and equally spaced out on an arbitrary traffic lane. To avoid counting any lane-changing vehicle occupying very small portion of the lane, the values for the middle detection nodes are set larger than those for the two-end nodes. Table 1 presents an example of the fixed values distribution over the detection nodes.

For instance, the case of seven detection points is explained as follows. Since nodes 1, 4, 7, 10, 13, 16, 19 at layer one will be aggregated into node 1 (R) at layer two, node R can then be expressed as:

$$o_k^2 = \sum_j w_{jk}^2 \cdot x_{jk}^{2*} \quad \text{for } k=1,\ j=1, 4, 7, 10, 13, 16, 19$$

Similarly, node G and node B can be expressed as:

$$o_k^2 = \sum_j w_{jk}^2 \cdot x_{jk}^{2*} \quad \text{for } k=2,\ j=2, 5, 8, 11, 14, 17, 20$$

$$o_k^2 = \sum_j w_{jk}^2 \cdot x_{jk}^{2*} \quad \text{for } k=3,\ j=3, 6, 9, 12, 15, 18, 21$$

The third layer, output layer, performs the defuzzification to obtain numerical outputs by utilizing the center average defuzzifier. The connection weight w_{km}^3 between k^{th} rule and m^{th} output node represents the consequence fuzzy singleton. Output layer will produce binary output values -- "0" representing no vehicle passing and "1" indicating one vehicle passing. At this layer, the node operation is expressed as follows:

$$o_m^3 = f(u_m^3) = \sum_{k=1}^{v} w_{km}^3 \cdot x_{km}^3 \quad \text{for } m=1\sim M$$

Since the number of nodes in the third layer is

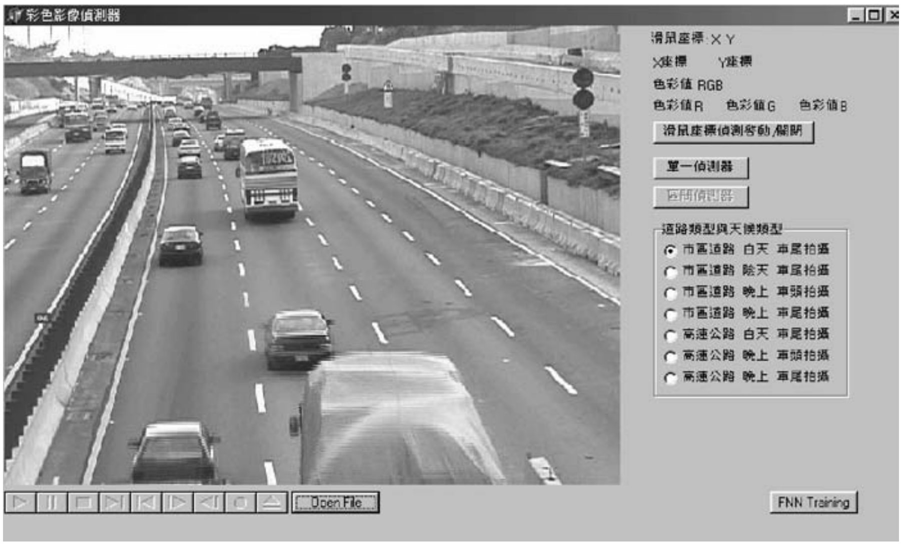

Fig. 7 Interface for vehicular detection module

only one, M is equal to one in this node operation. The connection weights w_{km}^3 are further adjusted by the supervised training algorithm, which is explained in the appendix.

2. Vehicle Detection Module

This module is to determine whether there are vehicles passing through the pseudo line detector. Fig. 7 shows the designed vehicular detection interface where the users can select appropriate environments and lighting conditions. As we move the cursor to any position within the traffic scene, this module will automatically read the exact X and Y coordinates as well as the corresponding instantaneous traffic image pixel (R, G, B) values.

IV. EXPERIMENTS

1. Data Collection and Evaluation Criteria

To take the roadway upstream and downstream traffic flow scenes, a video camera is placed on a grade overpass. The camera's field of view is set vertical to the road surface in order to reduce the vehicle occlusion situations in the daytime and to decrease the error signals (false) caused by headlights or taillights in the nighttime. In this study, a downstream view is taken when recording the traffic scenes in the daytime; while in the nighttime, both upstream and downstream views are taken so that the effects of vehicle headlights and taillights on detection accuracy can be compared.

A pseudo line detector, with three-, five- and seven-detection points equally spaced out on a specific traffic lane, is placed on the monitor. The detection performances for both CIVD and FNNCIVD systems are then compared. The detection outcomes are classified into three situations-- "success," "missing," and "false." "Success" is defined as the situation that a vehicle is detected when it actually passes through the pseudo line detector. "Missing" is the case when no vehicle is detected but a vehicle actually passes. "False" is identified when a vehicle is detected but actually no vehicle does exist.

The detection performance is evaluated by these three criteria. In this paper, traffic on a designated lane is detected. Thus, any lane-changing vehicle, which takes only a small portion of that designated lane width, will not be counted as a success outcome on that lane. Videotapes for both freeway mainline and urban street traffic scenes are recorded from 3 pm to 4 pm (daytime) and from 6 pm to 7 pm (nighttime). Validation experiments are then conducted off-line by sampling from the videotapes. Details of the experimental results are presented as follows.

2. Experiment on Freeway Mainline

The study location is at the Hsin Chu mainline section of Taiwan Freeway No.1 with three lanes southbound and four lanes northbound. Training examples under different lighting conditions associated with various detection points are collected for network training. Table 2 presents the number of training examples and the converged epochs corresponding to the error values (TE). For instance, the number of training examples for nighttime downstream viewing with seven detection points is 1,195. The

Table 2 Network training on freeway mainline

Detection points	Videotape recording time	Training examples	Training epochs	Error value
Three	Daytime	1,837	4,500	84
	Nighttime(upstream viewing)	1,600	2,307	7
	Nighttime(downstream viewing)	1,040	6,000	64
Five	Daytime	1,837	850	105
	Nighttime(upstream viewing)	1,600	800	25
	Nighttime(downstream viewing)	1,040	2,000	47
Seven	Daytime	1,837	700	38
	Nighttime(upstream viewing)	1,600	400	37
	Nighttime(downstream viewing)	1,040	600	119

Table 3 Detection performance on freeway mainline

Lighting conditions	Criteria	Systems	Three points	Five points	Seven points
Daytime (n=220)	Success	CIVD (interval search method)	159 (68.2%)	174 (75.9%)	178 (78.1%)
		CIVD (regression method)	161 (70.0%)	168 (73.0%)	176 (77.5%)
		FNNCIVD	198 (86.8%)	210 (92.5%)	212 (92.9%)
	Missing	CIVD (interval search method)	61 (26.1%)	46 (20.0%)	42 (21.0%)
		CIVD (regression method)	59 (25.6%)	52 (22.6%)	34 (14.9%)
		FNNCIVD	22 (9.6%)	10 (5.2%)	8 (3.5%)
	False	CIVD (interval search method)	13 (5.7%)	9 (4.1%)	8 (2.9%)
		CIVD (regression method)	10 (4.4%)	10 (4.4%)	7 (7.6%)
		FNNCIVD	8 (3.6%)	7 (3.8%)	8 (3.5%)
Nighttime upstream viewing (n=119)	Success	CIVD (interval search method)	72 (56.2%)	82 (64.5%)	86 (67.1%)
		CIVD (regression method)	70 (53.4%)	83 (64.3%)	84 (63.2%)
		FNNCIVD	106 (84.8%)	111 (88.8%)	113 (90.2%)
	Missing	CIVD (interval search method)	47 (36.7%)	37 (29.1%)	33 (25.7%)
		CIVD (regression method)	49 (37.4%)	36 (27.9%)	35 (26.3%)
		FNNCIVD	13 (10.4%)	8 (4.5%)	6 (4.0%)
	False	CIVD (interval search method)	9 (7.1%)	8 (6.4%)	9 (7.2%)
		CIVD (regression method)	12 (9.1%)	10 (7.8%)	14 (10.5%)
		FNNCIVD	6 (4.8%)	5 (2.2%)	4 (1.8%)
Nighttime downstream viewing (n=196)	Success	CIVD (interval search method)	136 (63.5%)	144 (67.9%)	148 (67.8%)
		CIVD (regression method)	135 (63.9%)	146 (69.5%)	149 (69.9%)
		FNNCIVD	186 (85.3%)	190 (90.0%)	188 (89.9%)
	Missing	CIVD (interval search method)	63 (29.4%)	52 (24.5%)	48 (22.0%)
		CIVD (regression method)	61 (28.9%)	50 (23.8%)	47 (22.1%)
		FNNCIVD	10 (4.5%)	6 (2.8%)	8 (3.8%)
	False	CIVD (interval search method)	18 (7.1%)	16 (7.6%)	22 (10.2%)
		CIVD (regression method)	15 (7.2%)	14 (6.7%)	17 (8.0%)
		FNNCIVD	22 (10.9%)	15 (7.1%)	13 (6.2%)

Note: n represents the actual number of vehicles being observed.

energy function is converged at about the 600th epoch with an error value of 119.

Table 3 summarizes the detection performance. Note that with various detection points in both CIVD and FNNCIVD systems, the "success" rate in the daytime is slightly higher than that in the nighttime. Downstream viewing detection performance is slightly better than upstream viewing. Also note that

Table 4 Network training on urban street

Detection points	Videotape recording time	Training examples	Training epochs	Error value
Three	Daytime	860	1,500	54
	Nighttime(upstream viewing)	1,600	2,307	7
	Nighttime(downstream viewing)	1,040	6,000	83
Five	Daytime	860	1,100	53
	Nighttime(upstream viewing)	1,600	800	25
	Nighttime(downstream viewing)	1,040	2,000	47
Seven	Daytime	860	900	58
	Nighttime(upstream viewing)	1,600	850	17
	Nighttime(downstream viewing)	1,040	400	38

the performances for seven and five detection points are much better than that for three points. Five detection points perform almost as well as seven detection points. For the CIVD system, both interval search method and regression method in effect make little difference in the detection performance. However, if we compare the CIVD with FNNCIVD, we find that the FNNCIVD system obviously outperforms. For instance, with seven detection points the "success" rates of FNNCIVD are improved from 78.1% of the interval-search CIVD (77.5% of the regression CIVD) to 92.6% in the daytime, from 67.1% (63.2%) to 90.2% in the nighttime upstream viewing, and from 67.8% (69.9%) to 89.9% in the nighttime downstream viewing. Compared with the interval-search CIVD system, the FNNCIVD system with various detection points can increase success rates in a range of 14% to 22% on the freeway mainline.

We have found that in the two cases of detection failure, "missing" mostly occurs in upstream viewing, while "false" often happens in downstream viewing. The main reason for detection failure in the daytime is due to resemblance of color pixels between the vehicles and the road background. Most of those vehicles are gray or gray-like. The second reason is the lane-changing vehicles that occupy a small portion of the lane width, and are not counted in the detected lane.

The high missing rate for nighttime upstream viewing is due to a fairly high threshold value being trained in the FNN module. In the nighttime, a vehicle is counted only when two headlights are simultaneously detected by the pseudo line detector. If merely one headlight is detected (lane-changing vehicles are the most cases), then that vehicle will not be counted. In contrast, the high false rate for nighttime downstream viewing can be ascribed to a fairly low threshold value being trained. Lane changing vehicles are mostly seen in this case.

3. Experiment on Urban Street

The study location is at section I of Chung Hwa Road, Taipei City, with four lanes northbound and five lanes southbound. Training examples under various lighting conditions with three, five and seven detection points are collected for network training. Table 4 illustrates the training conditions in the daytime and nighttime. For instance, in the nighttime downstream viewing with seven detection points, the number of training examples is 1,040 and the energy function converges at around 400^{th} epoch with an error value of 38.

Table 5 summarizes the detection performance. It is found that the detection performance by the interval search method is about the same as that by the regression method for the CIVD system. The detection "success" rate for the FNNCIVD system is superior to the CIVD system no matter which method is used. With seven detection points, for instance, the "success" rates of FNNCIVD can be enhanced from interval-search CIVD of 75.7% (regression CIVD of 75.3%) to 93.1% in the daytime, from 63.8% (63.7%) to 87.8% in the nighttime upstream viewing, and from 63.5% (64.9%) to 89.4% in the nighttime downstream viewing. Compared with the interval-search CIVD system, the FNNCIVD system with various detection points can increase the success rates in a range of 18% to 26% on the urban street.

The trends of "missing" and "false" on the urban street under various lighting conditions are quite similar to those on the freeway mainline. Reasons for the detection failure on urban street are also the same as those addressed in the freeway mainline case.

V. CONCLUDING REMARKS

In this study, a color image vehicular detection (CIVD) system is developed and validated off-line under different lighting conditions. A fuzzy neural

Table 5 Detection performance on urban street

Lighting conditions	Criteria	Systems	Three points	Five points	Seven points
Daytime (n=153)	Success	CIVD (interval search method)	111 (67.6%)	120 (74.5%)	125 (75.7%)
		CIVD (regression method)	115 (69.6%)	119 (73.9%)	122 (75.3%)
		FNNCIVD	135 (83.8%)	144 (90.5%)	148 (93.1%)
	Missing	CIVD (interval search method)	42 (26.0%)	33 (19.1%)	28 (16.9%)
		CIVD (regression method)	38 (23.0%)	34 (21.1%)	31 (19.1%)
		FNNCIVD	18 (11.1%)	9 (5.6%)	5 (3.1%)
	False	CIVD (interval search method)	8 (6.4%)	8 (6.4%)	12 (7.4%)
		CIVD (regression method)	12 (7.4%)	8 (5.0%)	9 (5.6%)
		FNNCIVD	8 (5.1%)	6 (3.7%)	6 (3.7%)
Nighttime upstream viewing (n=96)	Success	CIVD (interval search method)	58 (55.2%)	63 (61.1%)	67 (63.8%)
		CIVD (regression method)	57 (56.4%)	60 (59.4%)	65 (63.7%)
		FNNCIVD	84 (82.3%)	88 (89.7%)	87 (87.8%)
	Missing	CIVD (interval search method)	38 (36.0%)	33 (32.0%)	29 (27.6%)
		CIVD (regression method)	39 (38.6%)	36 (35.6%)	31 (30.3%)
		FNNCIVD	12 (11.8%)	8 (8.2%)	9 (9.1%)
	False	CIVD (interval search method)	9 (8.8%)	7 (6.9%)	9 (8.6%)
		CIVD (regression method)	5 (5.0%)	5 (5.0%)	6 (6.0%)
		FNNCIVD	6 (5.9%)	2 (2.1%)	3 (3.1%)
Nighttime downstream viewing (n=144)	Success	CIVD (interval search method)	90 (57.0%)	103 (65.1%)	101 (63.5%)
		CIVD (regression method)	91 (59.0%)	105 (68.6%)	100 (64.9%)
		FNNCIVD	130 (83.3%)	137 (90.8%)	135 (89.4%)
	Missing	CIVD (interval search method)	54 (34.6%)	41 (25.9%)	43 (27.0%)
		CIVD (regression method)	53 (34.4%)	39 (25.4%)	44 (28.5%)
		FNNCIVD	14 (8.9%)	7 (4.6%)	9 (5.9%)
	False	CIVD (interval search method)	12 (8.4%)	14 (9.0%)	15 (9.4%)
		CIVD (regression method)	10 (6.6%)	9 (6.0%)	10 (6.6%)
		FNNCIVD	12 (7.8%)	8 (5.1%)	7 (4.6%)

Note: n represents the actual number of vehicles being observed.

network (FNN) is further introduced to this CIVD system. Both freeway mainline and urban street are chosen as the experimental environments in which the traffic images are photographed by a camera such that the field of view is set vertical to the road surface in order to reduce vehicle occlusion situations in the daytime and to decrease the error signals (false) by headlights or taillights in the nighttime. A downstream view is taken when recording the traffic scenes in the daytime; while in the nighttime, both upstream and downstream views are taken so that the effects of vehicle headlights and taillights on detection accuracy can be compared.

We conclude that under different environments (freeway and urban street) and various lighting conditions (daytime and nighttime), the detection success rates for interval search and regression methods of a CIVD system are about the same. No matter which method is used, however, the CIVD system has performed worse than the FNNCIVD system. Compared with the interval-search CIVD, the FNNCIVD system can enhance the detection "success" rates at a range of 14% to 22% on the freeway and 18% to 26% on the urban street. Both CIVD and FNNCIVD systems perform slightly better in the daytime than in the nighttime.

We find that detection failure in the daytime mainly results from the resemblance of the color pixels between vehicles and road backgrounds. The other reason is due to vehicles lane changing, which are not counted in this single-lane vehicular detection system. In the nighttime, "missing" mostly occurs in the upstream viewing while "false" often happens in the downstream viewing. High missing rate in the nighttime upstream viewing is due to a high threshold value. High false rate in the nighttime downstream viewing, on the contrary, is owing to a low threshold value.

This paper deals only with single-lane traffic detection. To apply the FNNCIVD system to multiple-lane traffic detection, one must place a pseudo line detector on the monitor across the entire road section. Besides, one needs to set up rules for assigning any lane-changing vehicle to a specific lane in order to avoid missing or double counting. Future studies may measure more traffic parameters such as vehicle length (classification), headways, and speeds by allocating tandem pseudo line detectors. Moreover, different types of FNN structures can also be explored and tested.

In this paper, the traffic flow images are photographed near 90 degrees (vertical) from the overpass crossing the street or the freeway. It is thought that different photograph angles may influence the detection results and this is worthy of further experiments. In addition, near dawn or dusk or in bad weather, the FNNCIVD system may need real-time training to obtain more appropriate threshold values that can cope with the lighting change. More experiments on such real-time training also deserve to be carried out.

ACKNOWLEDGEMENTS

This work was supported by the National Science Council of ROC under grant number NSC-89-2211-E-009-082. Constructive comments from three referees are highly appreciated.

NOMENCLATURE

a	parameter of trapezoidal membership function
a_j^1	parameter of trapezoidal membership function at layer one
b	parameter of trapezoidal membership function
b_j^1	parameter of trapezoidal membership function at layer one
d	desired output
$d_m^3(t)$	desired output of the t^{th} training example at layer three
j	j^{th} node at layer one
J	number of nodes in the first layer
k	k^{th} node at layer two
K	number of nodes in the second layer
m	m^{th} node at layer three
M	number of nodes in the third layer
n	number of current training epochs
N	number of total training epochs
o	output value at each layer
o_j^1	output value of j^{th} node at layer one
o_k^2	output value of k^{th} node at layer two
o_m^3	output value of m^{th} node at layer three
$o_m^3(t)$	network output of the t^{th} training example at layer three
q	number of detection points
t	t^{th} training example in FNN
TE	energy function
w	connection weight
w_{jk}^2	connection weight between j^{th} node and k^{th} node
w_{km}^3	connection weight between k^{th} node and m^{th} node
x	input value at each layer
x_j^1	input value of j^{th} node at layer one
x_{jk}^2	input value from j^{th} node to k^{th} node at layer two
x_{km}^3	input value from k^{th} node to m^{th} node at layer three
z_j	fixed values
α	momentum parameter
δ	error signal
δ_j^1	j^{th} error signal at layer one
δ_k^2	k^{th} error signal at layer two
δ_m^3	m^{th} error signal at layer three
Δ	differencing value of a specific parameter between t+1th and t^{th} training example
$\Delta a_j^1(t)$	$= a_j^1(t+1) - a_j^1(t)$
$\Delta b_j^1(t)$	$= b_j^1(t+1) - b_j^1(t)$
ΔB_i	differencing of blue pixel values of i^{th} pseudo point detector
ΔG_i	differencing of green pixel values of i^{th} pseudo point detector
ΔR_i	differencing of red pixel values of i^{th} pseudo point detector
$\Delta w_{km}^3(t)$	$= w_{km}^3(t+1) - w_{km}^3(t)$
ε	arbitrary small number
η	learning rate

REFERENCES

Alejandro, A., Hickman, M., Mirchandani, P., and Chandnani, D., 2002, "Methods of Traffic Data Collection Using Aerial Video," *Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems*, pp. 31-36.

Blosseville, J. M., Krafft, C., Lenoir, F., Motyka, V., and Beucher, S., 1989, "TITAN: A Traffic Measurement System Using Image Processing Techniques," *Proceedings of the Second International Conference on Road Traffic Monitoring*, pp. 84-88.

Buckley, J. J., and Hayashi, Y., 1994, "Fuzzy Neural Network: A Survey," *Fuzzy Sets and Systems*, Vol. 66, No.1, pp. 1-13.

Bullock, D., Garrett, J., and Hendrickson, C., 1993, "A Neural Network for Image-Based Vehicle Detection," *Transportation Research*, Part

C, Vol. 1, No. 3, pp. 235-247.

Buluswar, S. D., and Draper, B. A., 1998, "Color Machine Vision for Autonomous Vehicles," *Engineering Applications of Artificial Intelligence*, Vol. 11, No.2, pp. 245-256.

Chang, W. Y., 1999, "Using Image Processing Technique for the Vehicles Detection and Classification," *Master Thesis*, National Cheng Kung University, Taiwan.

Chiou, Y. C., and Lan, L. W., 1997, "Development and Application of Learning Algorithm for Fuzzy Neural Networks: A Case of Pavement Diagnosis," *Transportation Planning Journal*, Vol. 26, No. 2, pp. 233-252.

Coifman, B., Beymer, D., McLauchlan, P., and Malik, J., 1998, "A Real-Time Computer System for Vehicle Tracking and Traffic Surveillance," *Transportation Research*, Part C, Vol. 6, No. 4, pp. 271-288.

Dickinson, K. W., and Wan, C. L., 1989, "Road Traffic Monitoring Using the TRIP II System," *Proceedings of the Second International Conference on Road Traffic Monitoring*, pp. 56-60.

Dickinson, K. W., and Waterfall, R. C., 1984a, "Image Processing Applied to Traffic: (1) A General Review," *Traffic Engineering & Control*, Vol. 25, No. 1, pp. 6-13.

Dickinson, K. W., and Waterfall, R. C., 1984b, "Image Processing Applied to Traffic: (2) Practical Experiences," *Traffic Engineering & Control*, Vol. 25, No. 1, pp. 60-67.

Dougherty, M., 1995, "A Review of Neural Networks Applied to Transport," *Transportation Research*, Part C, Vol. 3, No. 4, pp. 247-260.

Fathy, M., and Siyal, Y. M., 1995, "An Image Detection Technique Based on Morphological Edge," *Pattern Recognition Letters*, Vol. 16, No. 12, pp. 1321-1330.

Hoose, N., 1992, "IMPACT: An Image Analysis Tool for Motorway Analysis and Surveillance," *Traffic Engineering & Control*, Vol. 33, No.3, pp. 140-149.

Jouseau, E., and Dorizzi, B., 1999, "Neural Networks and Fuzzy Data Fusion: Application to an On-Line and Real Time Vehicle Detection System," *Pattern Recognition Letters*, Vol. 20, Issue 1, pp. 97-107.

Kumar, P., Sengupta, K., Lee, A., and Ranganath, S., 2002, "A Comparative Study of Different Color Spaces for Foreground and Shadow Detection for Traffic Monitoring System," *Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems*, pp. 100-105.

Lan, L. W., and Kuo, C. W., 1999, "Preliminary Development on the Color Image Vehicular Detection System," *Civil Engineering Technology*, No. 16, pp. 89-100.

Liao, M. T., 1993, "An Application of Sampling Point in Real Time Vehicle Classification," *Master Thesis*, National Taiwan University, Taiwan.

Li, X., Liu, Z. Q., and Leung, K. M., 2002, "Detection of Vehicles from Traffic Scenes Using Fuzzy Integrals," *Pattern Recognition*, Vol. 35, Issue 4, pp. 967-980.

Mantri, S., and Bullock, D., 1995, "Analysis of Feedforward-Backpropagation Neural Networks Used in Vehicle Detection," *Transportation Research*, Part C, Vol. 3, No. 3, pp. 161-171.

Michalopoulos, P. G., 1991, "Vehicle Detection Video through Image Processing: the Autoscope System," *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 1, pp. 21-29.

Shu, T. S., 1999, "The Study on Dynamic Vehicle Tracking and Traffic Parameter Collection System," *Master Thesis*, National Cheng Kung University, Taiwan.

Washburn, S. S., and Nihan, N. L., 1999, "Estimating Link Travel Time with the Mobilizer Video Image Tracking System," *Journal of Transportation Engineering*, Vol. 125, No. 1, pp. 15-20.

Manuscript Received: Jun. 21, 2002
Revision Received: Feb. 15, 2003
and Accepted: Apr. 04, 2003

APPENDIX

In this paper, the training algorithm can be decomposed by the following nine steps.

Step 1. Set network parameters(α, a_j^1, b_j^1)

$$\eta = 1 - \frac{n}{N}$$

The term η represents the learning rate which decreases as the number of training cycles n increases. N represents the number of the total training epochs. Initially, the network parameters including momentum parameter α, a_j^1, b_j^1 are set equal to 0.8, 20, 35, respectively.

Step 2. Input a training example and compute the network output

A training example is composed of an input vector (differences of pixel values) and an output vector (binary values indicating vehicles passing information). The output values at each layer are calculated by the equations shown in the section of layers operation.

Step 3. Employ a network output and desired output to get δ_m^3 for the output layer

$$\delta_m^3(t) = d_m^3(t) - o_m^3(t)$$

Step 4. Renew the connection weight w^3_{km} between rule layer and output layer

$$w^3_{km}(t+1) = w^3_{km}(t) + \eta \cdot \delta^3_m(t) \cdot x^3_{km}(t) + \alpha \Delta w^3_{km}(t)$$

where $\Delta w^3_{km}(t) = w^3_{km}(t+1) - w^3_{km}(t)$

Step 5. Compute the propagated error signal δ^2_k for the rule layer

$$\delta^2_k = \delta^3_m \cdot w^3_{km}$$

Step 6. Compute the propagated error signal δ^1_j for the membership layer

$$\delta^1_j = \delta^2_k \cdot z_j$$

Step 7. Renew the adjusted parameters for the membership layer

$$a^1_j(t+1)$$

$$= a^1_j(t) + \eta \cdot \delta^1_j \cdot \frac{x^1_j - b^1_j}{(b^1_j - a^1_j)(b^1_j - a^1_j)} + \alpha \Delta a^1_j(t)$$

where $\Delta a^1_j(t) = a^1_j(t+1) - a^1_j(t)$

$$b^1_j(t+1)$$

$$= b^1_j(t) + \eta \cdot \delta^1_j \cdot \frac{a^1_j - x^1_j}{(b^1_j - a^1_j)(b^1_j - a^1_j)} + \alpha \Delta b^1_j(t)$$

where $\Delta b^1_j(t) = b^1_j(t+1) - b^1_j(t)$

Step 8. Repeat step 2 to step 7
In this step, the sum of error squares is calculated. Repeat step 2 to step 7 until all training examples are finished (called an epoch). The energy function for the n^{th} epoch (TE_n) is calculated by

$$TE_n = \frac{1}{2} \sum_{t=1}^{T} [d^3_m(t) - o^3_m(t)]^2$$

where $d^3_m(t)$ is the desired output for the t^{th} training example and $o^3_m(t)$ is the output of t^{th} training example in FNN.

Step 9. Test if the stop condition satisfies
Training is terminated when a predetermined number of training cycles is reached or the energy function converges; namely, $|TE_n - TE_{n-1}| \le \varepsilon$, where ε is an arbitrary small number. If the TE value decreases smoothly, a stop condition is reached. Otherwise, go to step 2. In this paper, the former condition is used.