A Vision-Based Traffic Light Detection System at Intersections

Yun-Chung Chung\textsuperscript{1}  Jung-Ming Wang\textsuperscript{1}  Sei-Wang Chen\textsuperscript{2}

Department of Information and Computer Education\textsuperscript{1}
Department of Computer Science and Information Engineering\textsuperscript{2}
National Taiwan Normal University, Taipei, Taiwan 106, Republic of China

Abstract

The traffic light detection system is one of the key components of the vision traffic law enforcement system, such as red light runner detecting, turning against traffic light, and stopping at the non-stopping zone. With various conditions of both open outdoor environments and device setups, the traffic light detection must be robust to weather and illumination conditions, and also tolerant to various perspective angles.

An automatic traffic light detection system at intersections is presented in this paper. It performs traffic light detection on traffic videos without any signals from the traffic light controllers. This system is useful to be integrated with another ITS (Intelligent Transportation System) components. Background images are first generated by the system and in the mean time illumination parameters are estimated. The HSI color model is employed, and fuzzy methods together with morphological technique are utilized to acquire the candidate traffic light areas. With the relative spatial and temporal information, the scales, positions, and timing sequences of traffic lights are obtained. Some results from a preliminary trial are reported, and the associated researches are in progress.

Keywords: Traffic Light Detection  Traffic Law Enforcement  ITS Subsystem  HSI Color Image Processing

1. Introduction

A red-light runner, who violates either purposively or non-intentionally the red light signals at an intersection, will easily cause a fatal accident. It is common that in order to pass an intersection as fast as possible, the red-light runner has to accelerate his/her vehicle. Therefore, once a collision occurs it must be a disaster. Traffic accidents do cost a lot almost everywhere. For example, the statistical data reported by The department of statistics of Ministry of Transportation and Communication (MOTC) of Taiwan Government manifested that approximately 3,388 people were killed and more than 1,541 people were critically injured in 2000 due to traffic accidents in Taiwan area. Moreover, in United States, a great loss of 41,800 people’s lives, and 3,219,000 people injured are reported in 2000 by official statistics (U.S. Department of Transportation, U. S. A. Government, 2002).
Recent research, Lai and Yung (1998), has shown that the rate of traffic light violations can be significantly reduced if red-light cameras are tactically incorporated with traffic lights. Currently, most of the law enforcement systems are comprised of a photo film camera and a few buried electric-magnetic loop sensors. Such systems will take in sequence two still snapshots of the scene whenever the loop sensors installed under the ground detect vehicles traveling beyond the stop lines at an intersection when the traffic light is on red. These photos will provide evidence of violations for courts of law.

Traditional traffic law enforcement, based on the concept of increasing the numbers of both policemen and vehicles on patrol, has been shown to be inefficient and expensive. There are clearly never adequate manpower and financial supports for maintaining all roads among 24 hours a day. A better alternative is then the development of automatic traffic law enforcement systems, and clearly an automatic traffic light detection system at intersections is one of the most important components of them (Matsushita, Kamijo, Ikeuchi, and Sakauchi, 2000).

The vision-based system is designed to provide great configuration flexibility, which may be deployed at the roadside, mounted on an overhead structure such as bridges, road sign poles, or gantries, either permanently or temporarily, or operated out of a parked police vehicle. Obviously, it is more mobile, accommodating, and cost effective. There is no more communication facility needed to communicate with traffic light controllers and loop detectors.

The vision-based technology not only makes detectors free to move around, but also cuts the linkage of depending communication to the traffic light controllers. For those systems using radar or laser techniques, the vision system changes the nature of their battle against countermeasure equipments of radar / laser detectors, and the vision system is inherently immune to any known countermeasures. Moreover, people are worry about exposing to these microwaves in a long time may somehow have bad effects or even hurt them.

The vision-based system can be further modified in advance to provide all conventional important traffic data parameters, such as: volume, speed, occupancy, length classification, gap, headway, and concentration, etc. (Beymer, McLauchlan, Coifman, and Malik, 1997; Dickinson, and Wan, 1989). It can also generate traffic citations showing clear digitized photographs of every vehicle committing an offense, as well as detailed statistical reports which may be a histogram of average vehicle counts over a time period, or including the total vehicle count, number of violations, speed profiles, and violation profiles.

Those information can be integrated with other ITS (Intelligent Transportation System) (Institute of Transportation, Taiwan, R. O. C. Government, 2002) components to provide road users on demand information (Jung and Ho, 1999). The real time information is very valuable, and it may propagate via any distribution device such as Internet, in-vehicle satellite navigation devices, wireless media (cellular mobile phones, blue-tooth devices, pagers, etc.), or traditional radio and TV broadcasts.

The architecture of ITS system is defined (Institute of Transportation, Taiwan, R. O. C. Government, 2002) including nine main components, which are traffic signal control, freeway management, transit management, incident management, electronic toll collection, electronic fare payment, railroad grade crossings, emergency management services, and
regional multimodal travel information. Although traffic signal control is one of the components, in recent researches, there are not many titles focus on traffic lights detection. Some of the researches (Dariush, and Fujimura, 1999; Naumann, Rasche, and Tacken, 1998) have assumed that the traffic light signals are provided by the signal control box or given manually; however, this assumption may be avoided by the vision system automatically, and this is the main motivation to drive us to develop this vision-based traffic light detection system.

Some traffic lights detection related researches are briefly discussed here. Lai and Yung in 1998 proposed a traffic light detection algorithm, which utilized the shape and color information to detect the traffic light. Their method defined the disc-shapes, aspect ratios and spatial relationships of the traffic lights, and the traffic light sequence can then be determined. However, since they have assumed the traffic lights have fix color ranges, i.e., red, green and yellow, all have predefined constant color ranges, and thus the system might be affected by weather or shadows. In addition, the nighttime condition is not reported in this paper.

Palmer, Mellerio and Cutler in 1997 utilized a special sunglare protection filters with the concept of Q factors in viewing traffic signal lights. The Q factors can be treated as a measure of color appearance distortion, thus the adoption of Q factors values was obviously arbitrary and not tightly based on experimental data. Their method focused on theoretical discussion and only reported a few experiment results.

In addition, the fuzzy and morphological techniques are employed in our research since the researches related to utilize fuzziness (Klir, and Yuan, 1995) and morphological methods on applications have been proven very powerful. These techniques are again demonstrated very powerful by experiments of our system.

To sum up, an automatic traffic light detection system at intersections is one of the key components of the vision traffic enforcement system, such as red light runner detecting, turning against traffic light, and stopping at the non-stopping zone. With the changeable outdoor environments and different setup conditions of the vision system, the traffic light detection must be robust to weather and illumination conditions (Wixson, Hanna, and Mishra, 1998), and also tolerant to various perspective angles. The detection system is presented in this paper, which is organized as follows.

After first introduction section, the outline of the proposed vision-based system is first addressed in Section 2, and then Section 3 detailed presents the algorithms designed in our system. Section 4 includes a brief report of some experimental examples of trials on real scene video sequences and also some discussions about the variances of the traffic lights. Finally, the concluding remarks and future research topics are contained in Section 5.
2. System Architecture

The system architecture of the proposed automatic traffic light detection system at intersections is described in this section. It detects and outputs the state of the traffic light signals for other ITS applications. Typical scene examples at intersections are illustrated in Fig. 1, where traffic lights are indicated in the rectangles, and Fig. 1(a) is a daytime scene and Fig. 1(b) is a nighttime scene. Some ITS applications could be issued after the traffic light conditions are known, such as taking a photo if somebody runs through the red light.

The block diagram of this system is depicted in Fig. 2, and it is constructed by two main stages: the pre-processing stage and detecting stage. The processing procedures including in these two stages are enclosed by dashed rectangles as indicated in the figure. The input of this system is an outdoor video sequence, $V$, which is composed by a sequence of continuous images. The output of it, $L$, is the detection result of the states of the traffic light. The result $L$ includes traffic light scales, positions, and timing sequence, which is ready for further processing. The thick arrows in the diagram represent the video streams, and the thin arrows are the data streams for the succeeding procedures. It should be addressed that a rectangle box is utilized to represent a processing procedure, and an arced-edge box is used to show the output result after applying a method.
2.1 System setup and pre-processing

The hardware requirement of the proposed system is not restricted, but it is recommended to compose it with a DV-standard digital video camera and a conventional PC based computer in which a large amount of storage space is required when performing video recording continuously. Certain communication subsystems are considered necessary if it is required to integrate with other ITS components. The camera is setup as Fig. 3 shown, where the coordinate of the camera is denoted as \( \{p_x, p_y, p_z\} \), and the coordinate of the world is denoted as \( \{W_x, W_y, W_z\} \). Some physical parameters are measured while installing the system for further processing, such as the height of the camera above the ground, \( h \), the focal length, \( f \), the inclination angle of the camera, \( \alpha \), the pan angle of the camera, \( \beta \), etc. Performing a vision system in an outdoor environment is critical to illumination conditions. Background Image Generation and Illumination Assessment procedures are employed in the pre-processing stage to provide necessary information to further processing.
Fig. 3: The camera setup, the coordinate of the camera is denoted as \( \{p_x, p_y, p_z\} \), and the coordinate of the world is denoted as \( \{W_x, W_y, W_z\} \), (a) top-view of the camera (b) side-view of the camera.

- Background Image Generation

To evaluate the current illumination condition, a background image, \( B \), is necessary to avoid measuring errors. The background image generation procedure (Horprasert, Harwood and Davis, 2000; Long and Yang, 1990) produces the current background image, which provides important reference to current foreground image. Other ITS applications such as vehicles counting, vehicles tracking can also take benefits of the background image if applicable. The background image varies as time passes, thus it requires incessantly updating.

- Illumination Assessment

It is important to evaluate the lighting condition while taking vision-based processing in the outdoor open environments and this process is continuously updating the illumination condition while updating the background image. The illumination assessment procedure is employed to provide the illumination parameters, \( P \), of the environment to those succeeding components. The parameters \( P \) are also recorded with the input video stream for further database processing.

2.2 The detecting stage

The detecting stage deals with the scene and there are many sub-procedures included in this stage, as illustrated in Fig. 2: Color Information, Fuzzification, Fuzzy Morphological Shape Operation, Integration, Traffic Light Extraction, Relative Spatial Information and Relative Temporal Information. Notice that those procedures or results can execute or obtain in parallel if they are depicted as horizontal together. However, it is not implemented in parallel here because the system actually performs very fast. All of them are briefly described in the following paragraphs.

Traffic light detection outputs the red-yellow-green light condition in time sequence. It first exams all possible regions in the scene that may contain traffic lights, and then chooses the best reference traffic light position to construct the traffic light sequence. The
detection sequence of the traffic light is the red first, and then the yellow, finally the green. It is reasonable that detecting the red light first because that the red color is more distinguishable then the others, and due to the low brightness and contrast of the green light, it is detected at last. The position relations of the red-yellow-green light also help to detect the traffic lights. After the sequence has been established, the following ITS application can take advantage with it. Moreover, this stage works continuously and should not stop until shut down the system. The sub-procedures of the traffic light detection stage is briefly described as follows:

♦ Color Information

The input video sequence $V$ is composed with standard RGB color images, thus the first step of the detecting stage is to convert RGB image into HSI color model, i.e., the hue image, $H$, and intensity image, $I$, for further processing.

♦ Fuzzification

The hue image, $H$, and intensity image, $I$, are fuzzifized with illumination parameters, $P$, to be the fuzzy hue map, $\overline{H}$, and fuzzy intensity map, $\overline{I}$.

♦ Fuzzy Morphological Shape Operation

With the disk-like fuzzy morphological template, morphology erosion and dilation operations are performed to eliminate noise and acquire all possible disk-like shape areas for the traffic light. A fuzzy shape map, $\overline{S}$, is obtained after the operations.

♦ Integration and Traffic Light Extraction

The three fuzzy maps; fuzzy hue map, $\overline{H}$; fuzzy intensity map, $\overline{I}$; and the fuzzy shape map, $\overline{S}$; are integrated together. Then the traffic light extraction is operated on the integrated map to have the candidate traffic lights, $C$.

♦ Relative Spatial Information

To eliminate the false detected traffic lights containing in $C$, the relative spatial information is very helpful. The traffic light contains red-yellow-green tricolor lights, which has fix relative spatial relationships.

♦ Relative Temporal Information

Besides relative spatial information, the temporal information is also useful to make confirmation on $C$ to be the true traffic lights. After checking the spatial and temporal information, the most possible of traffic light is obtained, and the scales, positions, and timing sequence of traffic light, $L$, are obtained.

2.3 Applications

The video sequence is employed for providing database of the legal evidence of violations. The system can also transfer information to other ITS applications. The procedures which may contain in the applications are video database recording with detected parameters, a red-light runner detection, alphabet-number recognition of license plates, a zoom in view of the driver’s face, coordinate transformation, and traffic parameters estimation, etc. Note that the extensions of these applications proposed here are our future research topics.

♦ Video database recording with parameters

Video recording with detected parameters offers legal evidence to issue a violation ticket, and the videotape/digital images can be used for other off-line processing or to verify the correctness of the system manually.
Red-light runner detection / Automatic license plate recognition

Red-light runner detection is based on the information of the traffic light condition. Automatic license plate alphabet-number recognition is very useful to access the vehicle’s data from the government’s database via network. It can notify policeman to intercept the vehicle immediately if any stolen car is detected. Furthermore, it can combine with the traffic ticket system to issue the ticket on-line.

3. Methodology

3.1 Color Models

The input of the system is a continuous traffic video sequence, which will first be separated to individual image frames for further processing. The image frames are composed with the RGB tri-color model and then converted into the HSI (Hue, Saturation, and Intensity) color model (Gonzalez, and Woods, 1993). It is well known that the hue value is uniform scaling and shifting invariance and it is also invariant to shadow and shading.

The components of HSI model can be transformed from the general camera RGB model using the following formulas (Gonzalez et al., 1993):

\[ I = \frac{(R + G + B)}{3}, \]
\[ S = 1 - \frac{3}{(R + G + B)}\left[ \min(R, G, B) \right], \]
\[ H = \cos^{-1}\left\{ \frac{1}{2}\left[ \frac{(R - G) + (R - B)}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right] \right\}, \]

where \( R, G, B \) are the spectrums of the red, green, and blue components correspondingly.

Let the current input image at time \( t \) is \( F_t \), which contains the three components as shown in Eq. (2),

\[ F_t = f_t(x, y) = \begin{bmatrix} f_{H,t}(x, y) \\ f_{S,t}(x, y) \\ f_{I,t}(x, y) \end{bmatrix}, \]

where \( f_t \) is the function of pixels, \( (x, y) \) denotes the spatial coordinates of the image, and the \( f_{H,t}, f_{S,t}, \) and \( f_{I,t} \) are the hue, saturation, and intensity components correspondingly. Furthermore, it is important to define red, yellow, and green color models in advance because that the traffic light is composed by these three colors. The specific, red (\( R_t \)), yellow (\( Y_t \)), and green (\( G_t \)), color models are defined as:

\[ R_t = \{ g_t(x, y) \in F_t | g_{H,t}(x, y) < \alpha_H, \text{and } g_{I,t}(x, y) > \alpha_I \}, \]
\[ Y_t = \{ g_t(x, y) \in F_t | g_{H,t}(x, y) - \frac{\pi}{3} < \alpha_H, \text{and } g_{I,t}(x, y) > \alpha_I \}, \]
\[ G_t = \{ g_t(x, y) \in F_t | g_{H,t}(x, y) - \frac{2\pi}{3} < \alpha_H, \text{and } g_{I,t}(x, y) > \alpha_I \}, \]

where \( \alpha_H \) and \( \alpha_I \) are threshold values which may be determined by illumination conditions or manually given.

3.2 Background Image Generation

This procedure creates the current background image, \( B \). Obviously, the background image varies as time passing, thus it requires incessantly updating.
Fig. 4: The background value, \( v \), of \( p \) is obtained from the most frequent values of \( p(x, y) \) among the certain detection period, \( t \). (a) The observation of pixel \( p(x, y) \) during time \( t \), (b) Intensity counts for pixel \( p(x, y) \) during time \( t \).

However, the background image should not be changed radically, and basically it changes with the illumination condition changing (Long, and Yang, 1990). A fast background generation method is employed to generate and update background images. It is assumed that every portion of the background will become visible at some time. For every pixel, the fast background generation method will observe it for a certain period, and then select the most frequent intensity values to be the background value. This method is very quick but it does not create a perfect background image. The background generated by this method may contain some obvious small noises, however, in this paper, these noises are not vital because they will not affect the next procedure, illumination analysis.

The background generation algorithm is depicted as following: given a certain time period, \( t \), for a pixel, \( p \), which is located at \((x, y)\) in image frame \( ft \) of the input video sequence, the background value, \( v \), of \( p \) is obtained from the most frequent values of \( p \) among the certain detection period. In implementation, an intensity histogram is created to record the frequency of each intensity of \( p \). This process is illustrated in Fig. 4, and Fig. 4(a) is the observation of pixel \( p(x, y) \) during time \( t \), and Fig.4 (b) shows intensity counts histogram for pixel \( p(x, y) \) during time \( t \).

3.3 Illumination Assessments

Some parameters in later procedures in the system are determined by the illumination condition. The illumination condition, \( P \), is considered as two different types, daytime and nighttime. The illumination analysis is applied on the background images to avoid the errors caused by vehicles or foreground objects.

Assume that the scene image containing \( N \) pixels, \( g_f(x, y) \) is the intensity value of the pixel located at \((x, y)\) in any frame \( f \) of the input video sequence. The illumination condition, \( p_f \), is determined by the
Table 1: The states of a traffic light

<table>
<thead>
<tr>
<th></th>
<th>Red light</th>
<th>Yellow light</th>
<th>Green light</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO</td>
<td>OFF</td>
<td>OFF</td>
<td>ON</td>
</tr>
<tr>
<td>RTS</td>
<td>OFF</td>
<td>ON</td>
<td>OFF</td>
</tr>
<tr>
<td>STOP</td>
<td>ON</td>
<td>OFF</td>
<td>OFF</td>
</tr>
</tbody>
</table>

Fig. 5: A standard layout of a traffic light (a) Horizontal Layout (b) Vertical Layout

By the design of the traffic light, the red (R), yellow (Y), and green (G) lights all have the same diameter, $d$, and also the same distance, $0.5d$, between the two adjacent lights. The actual size of $d$ is set to 30cm by Taiwan’s traffic law (Ministry of Transportation and Communications and Ministry of the Interior, R. O. C. Government, 1994), which may be used as a scale reference. These position relations of the red-yellow-green light are helpful to detect them. The lighting sequence of the traffic light in detail. Most traffic lights are horizontal layouts in Taiwan, thus the algorithm designed here is specialized to detect horizontal lights. However, if to detect a vertical layout traffic light is desired, the following algorithm can be easily adapted.

3.4 The Traffic Light Detection

The goal of this procedure is to construct a traffic light sequence by input video sequence automatically. A dimension sketch of a regular horizontal layout of the traffic light is illustrated at Fig. 5(a), and a vertical layout is shown in Fig. 5(b). Taiwan’s traffic law (Ministry of Transportation and Communications and Ministry of the Interior, R. O. C. Government, 1994) restricts the standard layout and the lighting sequence of equations (4):

$$p_f = \frac{1}{N} \sum_{x,y} g_f(x,y)$$

It is determined as daytime if $p_f$ is larger than certain threshold, otherwise, it will be nighttime.
although the time period of lighting “ON” is variant at each intersection, the sequence is invariant from GO -> RTS (Ready To Stop) -> STOP and then back to GO, as illustrated in Table 1.

The detection sequence of the traffic light is the red “ON” light first, and then the yellow “ON”, finally the green “ON”. The reason to detect the “ON” red light first is that the red color is more distinguishable then the others, and due to the low brightness and contrast of the green light, it is detected after all.

The system exams all possible regions in the scene that may contain traffic lights, and then chooses the best reference traffic light position to construct the traffic light sequence. The fuzzy membership functions are utilized to help locate the position. A fuzzy intensity function, $\mu_i$, and a fuzzy hue function, $\mu_h$, are employed to decide the possible regions belong to an “ON” red light, as illustrated in Fig. 6, where $l$ is the number of gray level. The fuzzy intensity function, $\mu_i$, is defined as

$$
\mu_i(x) = \begin{cases} 
0 & , x < i_l \\
\frac{x - i_l}{i_r - i_l}, & i_l < x < i_r \\
i_r - i_l & , otherwise
\end{cases}
$$

and the fuzzy hue function $\mu_h$ is a Gaussian function whose mean, $h_m$, is located at “red” color of the hue scale, and its variance is $h_v$. These fuzzy function parameters, $i_l$, $i_r$, and $h_v$, can be determined by referencing the illumination condition. A fuzzy hue map, $\mathbf{H}$, and a fuzzy intensity map, $\mathbf{I}$, are acquired from the fuzzification procedures.

Morphology erosion and dilation operations are then utilized to find out the shape information of the traffic light. The translation of a binary image $A$ by a pixel $k$ is an image defined (Jain, Kasturi and Schunck, 1995) by

$$
A_k = \{a + k \mid a \in A\}.
$$

The binary morphology dilation is then defined (Jain et al., 1995) as a union of a binary image $A$ by

$$
A \oplus B = \bigcup_{h_i \in h} A_{h_i}, \quad (7)
$$

binary image $B = \{b_1, b_2, \ldots, b_n\}$, and is given by

where $A_{h_i}$ is the translation of the binary image $A$ by the “1” pixels of the binary image $B$. Clearly, morphology dilation has both associative and commutative properties by this definition. The morphology erosion (Jain et al., 1995), the opposite of dilation, of a binary image $A$ by a binary image $B$ is “1”
at a pixel $k$ if and only if every “1” pixels in the translation of $B$ to $k$ is also “1” in $A$. Symbolically,
\[ A \otimes B = \{ k \mid B_k \subseteq A \}, \tag{8} \]
and the binary image $B$ is referred to as a structuring element.

The fuzzy morphology dilation and erosion operands are defined as extensions from these binary bases. Given a gray-levels image $A = \{ a_1, a_2, \ldots, a_m \}$, where $a_i$ is an integer and usually ranges from 0 to 255, and also a fuzzy template $B = \{ b_1, b_2, \ldots, b_n \}$, where $b_i$ ranges from 0 to 1 as fuzzy values. The fuzzy dilation unions the translation of $a_i$ of the image $A$ by multiplying the template $B$ with every relative $b_i$, and then normalizes with the pixels number of $B_k$. Symbolically,
\[ \overline{A} \equiv A \circ B = \left\{ k \mid b_i \ast a_i \right\}, \tag{10} \]
where $|B|$ is the pixels number counting.

In our system, a disk like fuzzy morphology mask is applied. For example, a 7x7 mask illustrated in Fig. 7 can be used to extract the shape information. Thus, the most possible “ON” red light is found by its shape information from fuzzy morphology, i.e., a fuzzy shape map, $\overline{S}$, is obtained.

Integrating the results of the three fuzzy maps: $\overline{H}$, $\overline{I}$, and $\overline{S}$, a fuzzy map of belonging to an “ON” red light is obtained. A predefined threshold determined by $P$ will be applied to the fuzzy map to obtain the candidate traffic lights, $C$. If there is no “ON” red light found in the current frame, it will be dropped, and the system will proceed with next frame until at least an “ON” red light is found. The $d$ value is then obtained by this information.

Next, with proper parameters setting, the “ON” yellow light can be found subsequently. Finally, with
illumination information and the spatial relationships, the “ON” green light can be easily detected.

If there is more than one set of traffic lights detected, the one of the largest size is first selected for subsequence analysis. With the relative spatial and temporal information of the traffic lights, the scales, positions, and timing sequences of traffic lights, $L$, will be determined. Note that the traffic light sequence must conform to the spatial and temporal limitations, thus if the largest one does not, the second candidate will be tested again to ensure it is truly a traffic light.

To judge which light is currently “ON”, we need

$$B_R = \frac{1}{A_R} \sum_{(x,y)} g_f(x,y),$$

(11)

the average brightness of it. The average brightness, $B_R$, for the red light first detected as “ON” is defined as:

where $A_R$ is the size of the red light area, and $g_f(x,y)$ is the intensity value of the pixel located at $(x,y)$ in any frame $f$ of the input video sequence.

The definitions of the average brightness of the yellow light, $B_Y$, and the green light, $B_G$, are similar. Symbolically,

$$B_Y = \frac{1}{A_y} \sum_{(x,y)} g_f(x,y),$$

(12)

$$B_G = \frac{1}{A_G} \sum_{(x,y)} g_f(x,y),$$

(13)

where $A_Y$ is the size of the yellow light area, and $A_G$ is the size of the green light area. We can then define the state of the red light as “ON” if the average brightness of the red area in current frame is higher than the half of average brightness, $0.5B_R$. Otherwise, the state of the red light is “OFF”. The states of the yellow and green lights are defined as the same manners.

After the sequence has been established, the pseudo connection between the traffic light controller and the system is ready for use. Moreover, the current traffic light condition can be quickly retrieved and verified at the detected area. The scale, and timing parameters are also useful to other ITS applications, such as a red-light runner detector program.

Fig. 8: Background images of examples in Fig.1, (a) Daytime Scene (b) Nighttime Scene.
4. Experiment Results and Discussion

Some experimental trial is stated in this section. The input video streams are taken by a conventional DV camera, which contain 30 frames per second of 320 x 240 resolutions in NTSC based video system. Although it is possible to handle 30 frames per second, it is not necessary to update the light conditions so fast in the traffic light detection system. The system actually performs detection algorithm every 0.5-1.0 second, i.e., sampling at 1-2Hz rates, and skip all other unused input frames.

The pre-processing stage will first obtain the background image, and the background images of the scene of Fig. 1 are shown in Fig. 8. Although the calculation of the background generation is very fast, it must wait for all the background pixels show up. Thus, it usually takes about several minutes to find the background image, and it will take longer time to initialize if the traffic is crowded in rush hours. The daytime background example image shown in Fig. 8(a) is obtained from about 250 frames, and the nighttime background image shown in Fig. 8(b) is obtained from about 350 frames. It is not critical to obtain the perfect background image in traffic light detection because that the background image is only utilized to determine the illumination value in this system. However, if other applications are developing, such as red-light runner or vehicles counting, etc., the background image should be tuned up. The illumination parameter is then calculated to determine the environment condition, which is daytime or nighttime.

The fuzzy hue map, $\vec{H}$, and fuzzy intensity map, $\vec{I}$, are obtained from image frames with the predefined fuzzy functions. After applying the morphological operations, the three fuzzy maps; fuzzy hue map, $\vec{H}$; fuzzy intensity map, $\vec{I}$; and the fuzzy shape map, $\vec{S}$; are integrated together. The candidate traffic areas are
then obtained.

With the relative spatial and temporal information, the scales, positions, and timing sequences of traffic lights are obtained. This system has been tested on outdoor complex scenes, with different illumination conditions and different camera angles. Fig. 1(a)-daytime and Fig. 1(b)-nighttime are two of the input examples.

An example of the output sequence of the detected traffic light conditions is in Fig. 9, where states 3 is ‘STOP’, state 2 is ‘RTS’, and state 1 is ‘GO’. This output sequence is one of the nighttime sequences of sampling rate of one frame per second. In this example, it shows that the red light lasts about 24 seconds, green light continues about 20 seconds, and the yellow light goes about 2-3 seconds.

In our experiments, total six different intersections are tested, all including daytime and nighttime sequences. The traffic light detection results of them are verified manually and they are all proved correct. These experiments demonstrate that this system performs very good traffic light condition recognition and also robust to outdoor environments. Moreover, the timing of the traffic light is obtained by the input frame timing, and it is useful to traffic control center to check out the timing for this intersection. If there is any problem of the traffic light, such as lights broken or timing error, the system can report this problem to the traffic control center right away.

In some environments, there will be more than one traffic light in the scene. The system will check for all candidate traffic lights, and decide the most reliable one to be the target. To avoid misrecognition of flashing neon advertisements or streetlamps, this decision is actually made after detecting all the sequences of traffic lights, i.e., to select one from the detected traffic light sequences. For example, there are two traffic lights and several small ones of the next intersection exists in Fig. 1(b), and the system will detect at least two traffic light sequences (the small ones may be omitted due to the size limitation), however, the largest shape of the traffic light among the candidates will be chosen by the system as the most reliable one. Note that in this case, it does not matter to choose either one of the two.

On the other hand, the traffic lights have many different combinations and lighting sequences in
Fig. 11: Descriptions of phases cycles, (a) Two phases cycle, (b) Four phases cycle.

practice (Ministry of Transportation and Communications and Ministry of the Interior, R. O. C. Government, 1994), such as combing with right-turn only, left-turn only, or go-straight only, etc. These special cases will be indicated by arrow green lights. The lighting sequence of the traffic light of an intersection can be describe as a phases cycle, and the phase time is determined by the traffic signal control box, or manually by traffic police officers.

A standard two phases cycle (left-turn prohibited), which is controlled by a red-yellow-green tri-color traffic light, can be depicted as Fig. 10(a). The phases cycle of this traffic light is described in Fig. 11(a). Moreover, a four phases cycle traffic light is very common and can easily be seen in a left-turn enabled intersection, and it is shown in Fig. 10(b). The phases cycle of this traffic light is described in Fig. 11(b).

Although the phases cycles of traffic lights may have many combinations, they have common layout: one circular red light and one circular yellow light combing with different arrow-type or circular-type green light(s). Recall that our system detects red lights first (the most reliable feature), and then the yellow lights, and finally the green lights. These arrow-type green lights are also limited by the spatial and temporal relationships of red and yellow lights. Thus these phases cycles can also be detected by modifying the circular green light template with an arrow-like morphological template.

5. Concluding Remarks and Future Researches

In this paper, an automatic traffic light detection system at intersections is presented. From experiment results, it shows that it is validated to perform traffic light detection on traffic videos without any signals from the traffic light controllers. The system architecture is briefly described as follows: it first generates background images and also illumination parameters. The HSI color model is then employed, and fuzzy methods and morphological technique are utilized together to find the candidate traffic light areas. With the relative spatial and temporal information, the scales, positions, and timing sequences of traffic lights
A Vision-Based Traffic Light Detection System at Intersections

are acquired.

An automatic traffic light detection system at intersections is one of the key components of the vision traffic enforcement system, such as red light runner detecting, turning against traffic light, and stopping at the non-stopping zone. Next, we will try to design a vision traffic law enforcement system integrated with the traffic light detection system; moreover, with the changeable outdoor environments and different setup conditions of the vision system, the system must be robust to weather and illumination conditions, and also tolerant to various perspective angles. This new system is currently under developing.

For the further researching, it is suggested that the vision traffic enforcement system can be designed in advanced to provide more traffic data parameters, such as: volume, speed, occupancy and concentration, etc. It can be customized to show the driver’s face or take a zoomed close up photo of his/her license plate if necessary. It can also generate traffic citations showing clear digitized photographs of every vehicle committing an offense, as well as detailed statistical reports which may be a histogram of average vehicle counts over a time period, or including the total vehicle count, speed profiles, and violation profiles. Those information can be integrated with other ITS components to provide road users on demand information. The real time information is very valuable, and it may propagate via any distribution device such as Internet, in-vehicle satellite navigation devices, wireless media, or traditional radio and TV broadcasts.
Reference


Yun-Chung Chung received the B.E. degree in naval architecture and ocean engineering from National Taiwan University, Taipei, Taiwan, R.O.C., in 1993, and the M.E. degree in information and computer education from National Taiwan Normal University, Taipei, Taiwan, in 1995. He is currently pursuing the Ph.D. degree in the department of information and computer education, National Taiwan Normal University. From 1997 to 1998, he worked as an Associate Engineer in Information Science and Technology Exhibition Center, Institute for Information Industry, Taipei, Taiwan. His research interests include traffic monitoring system, visual tracking and surveillance and digital image processing.

Jung-Ming Wang received the B.E. degree in information and computer education from National Taiwan Normal University, Taipei, Taiwan, R.O.C., in 1996. He is currently pursuing the M.E. degree in the department of information and computer education, National Taiwan Normal University. His research interests include digital image processing pattern recognition and computer vision.

Sei-Wang Chen received the B.Sc. degree in atmospheric and space physics, the M.Sc. degree in geophysics from National Central University in 1974 and 1976, respectively, and the M.Sc. and Ph.D. degrees in

About the Authors

Yun-Chung Chung

Jung-Ming Wang

Sei-Wang Chen
computer science from Michigan State University, East Lansing, Michigan, in 1985 and 1989, respectively.

From 1977 to 1983, he worked as a Research Assistant in the Computer Center of Central Weather Bureau, Republic of China. In 1990, he was a researcher in the Advanced Technology Center of the Computer and Communication Laboratories at the Industrial Technology Research Institute. From 1991 to 1994 and from 1995 to 2001, he was an associated professor and full professor of the Department of Information and Computer Education at the National Taiwan Normal University, Taipei, Taiwan, Republic of China, respectively. He is currently a professor of the Graduate Institute of Computer Science and Information Engineering at the same University. His areas of research interest include neural networks, fuzzy systems, pattern recognition, image processing and computer vision.

Manuscript Received March 18, 2002
Revised March 28, 2002
Accepted April 16, 2002

*Acknowledgement
This work is supported in part by the National Science Council, Taiwan, Republic of China under contract NSC90-2213-E-003-001
十字路口紅綠燈自動偵測系統為發展以視覺為基礎之道路交通違規偵測系統之重要基礎，例如闖紅燈違規、紅燈違規左右轉或暫停於禁止臨時停車位置等違規都需要紅綠燈的資訊。這類型的視覺系統由於是架設於室外開放的環境，因此必須要能夠克服天氣、陰影、光線明亮度變化以及架設位置不固定等相關的問題。

我們希望能夠研究出一個以視覺為基礎之十字路口紅綠燈自動偵測系統，可以提供紅綠燈的訊息資料，而不需與紅綠燈控制信號箱有實體的連線，本系統所偵測的紅綠燈訊號即可配合其他智慧型交通運輸系統(ITS)做後續處理。本系統首先產生背景影像，並藉以判斷明亮度值—分辨白天或黑夜。其次採用 HSI彩色模式配合模糊 Morphology 技術處理來取得可能的紅綠燈位置。配合空間與時間上的限制條件，將不合理的候選區域去除，即可將紅綠燈的位置、大小與轉換時間偵測出來。

關鍵字：紅綠燈自動偵測 道路交通違規偵測 智慧型交通運輸系統(ITS) HSI彩色模式處理