

An Effective Video Analysis Method for Detecting Red Light Runners

Nelson H. C. Yung, *Senior Member, IEEE*, and Andrew H. S. Lai, *Member, IEEE*

Abstract—This paper presents a novel method for automatic red light runner detection on a video, which is fundamentally different from the concept of conventional red light camera systems. In principle, it extracts the state of the traffic lights and vehicle motions without any physical or electronic interconnections to the traffic light control system or the buried loop detectors. Purely from the video, the new method first constructs a traffic light sequence and then it estimates vehicle motions beyond the stop line while the light is red. In the former, the spatial and temporal relationships of individual traffic lights are utilized. In the latter, the concept of virtual loop detector has been introduced to emulate the physical loop detectors. A prototype was implemented based on this method and was tested in a number of field trials. The results show that the new method is able to detect multiple red light runners in multiple lanes. It is also capable of tolerating a number of hostile but realistic situation such as: 1) minimum number of traffic light; 2) pseudomotions due to shadows; 3) poor contrast; 4) pedestrian motions; and 5) turning vehicles.

Index Terms—Motion estimation, red light camera, traffic light detection, video analysis.

I. INTRODUCTION

ONE of the major threats to road safety is reckoned to be posted by drivers disobeying the traffic signals at signalized junctions. Statistics from many countries showed that a high percentage of serious road accidents occurred at road junctions, and a high percentage of these are due to drivers disobeying or “running” the red light [1]. According to Baguley [2], not all of these “red light runners” (RLR) run the light deliberately. Disregarding their motives or circumstances, they undoubtedly put the lives of the other road users and pedestrians at risk, including their own, unnecessarily. Aiming to deter red light running, imaging, photographic, and sensor technologies have been developed and integrated for detecting the violation in action and taking photographic records of the incident. This effort resulted in a class of systems called the red light camera (RLC), which operates in conjunction with the traffic light control circuit and inductive loop detectors (ILD) buried under the surface of the road at the junction [3]–[7]. Indeed, all RLC available nowadays share this common methodology. For instance, the Gatsometer RLC 36 series [8], AutoPatrol RL-200 [9], Driver Safety Systems RED EYE 77 [10], Zellweger

MULTAFOT [11] and Control Technologies REDGUARD [12] all stem from the same concept. In essence, such system must be electronically connected to the traffic light control system and the ILD buried under the road surface, usually beyond the stop line. It is typically mounted on a post at some distance from the junction, with a recording camera viewing from behind the traffic direction. The field of view (FOV) of the camera must include at least one traffic light (typically two), the junction and the vehicle’s license number for prosecution. The additional functions of the system are: 1) handling signals from the traffic light and ILD; 2) activating the camera; and 3) recording other relevant information of the offence. When the traffic light is red and the ILD sense the presence of vehicle(s) on either lane, the camera will be triggered to take two photographs of about 1 s apart. The first photo shows the vehicle entering the junction beyond the stop line, and the second photo depicts the vehicle leaving the junction. Detailed information such as date, time, camera identity, lane number, red light duration, and detection identity will also be printed on the photographs for future reference.

At present, many RLC systems are gradually adopting digital technologies for recording, transmission, and pattern recognition. The DSS RED EYE 77 system [10] employs digital cameras, digital tapes and WORM CD to replace the conventional camera and film. Image data may be coded and transmitted over high-speed ISDN lines to a center, where optical character recognition may be carried out to produce citations automatically. On the other hand, AutoPatrol RL-200 [9] supports independent video monitoring and recording of the junction traffic 24 h/day or according to a predetermined schedule. In this case, the functionality of the RLC and the video recording are completely independent. Similarly, the REDGUARD [12] system is able to detect red light violation in a conventional way and record the event with a video camera.

In summary, no matter how the RLC technology evolves, its fundamental principle remains immobile, inflexible and expensive, because of the use of the ILD and traffic light signals. To improve the effectiveness of deterring red light running, the RLC principle must be redefined. One possible approach is to apply video technology and video analysis techniques appropriately, as video contains a large amount of spatial and temporal information that have not been fully exploited so far. Furthermore, video technology and video analysis theories are maturing rapidly. They together present a tremendous opportunity for research and development of a new approach here. In theory, humans can detect RLR easily by visually inspecting the location of the vehicle(s) with respect to the stop line at the junction

Manuscript received September 29, 1998; revised November 13, 2000. This work was supported by the University Grants Committee, Hong Kong, Area of Excellence in Information Technology, Grant AOE98/99.EG01 and by the University of Hong Kong and Li Po Chun Scholarship.

The authors are with the Laboratory for ITS Research, Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong SAR (e-mail: nyung@eee.hku.hk; hslai@ieee.org).

Publisher Item Identifier S 0018-9545(01)03059-6.

TABLE I
FOUR STATES OF A TRAFFIC LIGHT (ON = LIGHT ON; OFF = LIGHT OFF)

	<i>Red</i>	<i>Yellow</i>	<i>Green</i>
Stop	ON	OFF	OFF
Ready to go (RTG)	ON	ON	OFF
Go	OFF	OFF	ON
ready to stop (RTS)	OFF	ON	OFF

and the state of the traffic light. The question is how this can be automated using video analysis techniques. Therefore, it is our intention to study and exploit this possibility in this paper and propose a solution to the above problem using traffic light detection and motion estimation techniques.

In this paper, we present a novel method that performs automatic RLR detection on the traffic video without needing any signals from the traffic light system or the ILD. The new method primarily tackles the problem at two stages: 1) it detects and constructs a traffic light sequence at the junction and 2) it estimates vehicle motions beyond the stop line, while the light is red. The traffic light sequence detection first locates all the traffic lights in the FOV by searching for all the red/yellow/green regions with aspect ratio close to unity and similar sizes. The spatial relationship between the lights is utilized to eliminate other regions, while the light's temporal relationship is used to improve the confidence of the former decision and to construct the light sequence. As there may be more than one traffic light under the FOV, the traffic light with the largest dimension, i.e., closest to the camera, is then chosen for all the subsequent analysis. This method is developed to replace the traffic light control signal. The detection of vehicle motion beyond the stop line is built on a direction-biased motion estimation algorithm. In this case, direction-biased three-step search algorithm is developed upon the concept of virtual loop detectors (VLD). The concept of VLD permits regions to be defined on the video that can emulate the function of the physical ILD. By defining these VLD beyond the stop line, their motions along the road direction may be estimated. In principle, motions due to pedestrians, turning vehicles and shadows may be differentiated according to their motion vectors. Based on this method, a prototype has been implemented and evaluated using a number of traffic junction videos taken under different outdoor conditions, including one with the video camera mounted on an existing RLC post. In all our evaluations, the method identified all the RLR correctly. We also observed that the new method is capable of tolerating a number of hostile but realistic situations: 1) minimum number of traffic light; 2) pseudomotions due to shadows; 3) poor contrast; 4) pedestrian motions; and 5) turning vehicles.

The organization of this paper is as follows. Section II analyzes the problem of red light runner detection. Section III overviews the new methodology by outlining the assumptions and its functional components. Section IV details the algorithms developed for the functional components with illustrative examples. Section V presents the results of a field trial and discusses

the performance of the methodology. This paper is concluded in Section VI.

II. PROBLEM ANALYSIS

In general, a traffic light consists of three separate lights in *red*, *yellow*, and *green* spatially aligned vertically from top to bottom in this order. As it is likely that there will be similar red, yellow, and green objects in the FOV, the first task in detecting the traffic lights is to isolate the lights from a multitude of similar color regions. Of the three colors, it is observed that *green* is the most difficult to isolate because of its variation in shape, and the presence of similar green regions such as trees and bushes around a typical junction. Therefore, it would probably be more efficient if the detection starts with the *red* and *yellow*.

In addition to the spatial relationship as described, there is also a temporal relationship governing the traffic lights. This sequence used in Hong Kong is illustrated in Table I. The duration of the RTG and RTS signals are normally fixed to 3 s in each iteration, while the duration of the stop and go signals may be varied manually or automatically. The argument for utilizing this temporal relationship in the detection is that outdoor environment is hostile. For instance, any sudden change in the ambient light level may affect the brightness of the lights, especially the *green* light and may introduce ambiguity. Therefore, tracking the light sequence would probably provide a more stable and robust solution.

By law, each entrance to a junction or a crossing must have a stop line, which is a thick solid white line across the entrance to the junction. As depicted in Fig. 1(a), each stop line spans across all the lanes in one direction. Vehicles must stop behind this line when the light is *red*. Therefore, the presence of the stop line serves as a good reference. However, stop lines on city roads are often broken, variable width and have poor contrast with the road color.

The prohibited zone can be weakly defined as the area between the entrance and the exit of a junction, bounded by the stop lines (Fig. 1). In general, this zone is only prohibited when the light is *red*. Conventional RLC requires the ILD to be buried in this zone. For detecting RLR on multiple lanes, each lane must be installed a separate set of ILD. In some interactive visual traffic surveillance systems [13], [14], instead of using ILD, one or more detection areas are drawn manually inside the prohibited zone. The questions are thus concerned with the detection area size, number and how the allocation can be achieved automatically.

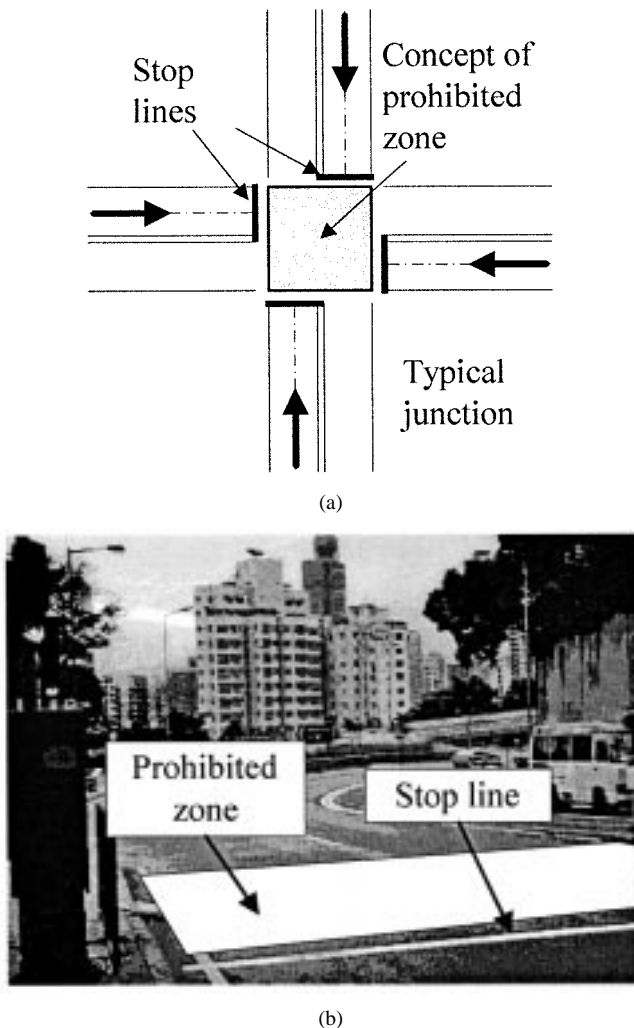


Fig. 1. (a) Stop lines and prohibited zone. (b) Real stop line and prohibited zone.

Vehicle motion estimation is one area that has been attempted by many in the past [15]–[23]. Most approaches perform two functions: vehicle extraction and motion estimation/tracking. To estimate the trajectories of moving vehicles, edge detection [17], straight-line approximation [18], region growing [19], block-based motion estimation [19], background extraction [20] or interactive mask specification [21] have all been reported. They share some common traits. First, most cases including the conventional RLC assumed the camera optical axis is approximately in parallel with the road direction without occlusion. Second, the computation complexity of extraction and motion estimation is usually high.

III. NEW RED LIGHT VIDEO METHODOLOGY

Let us first assume that the video camera is mounted on a 3-m post and located about 10 m from the junction. Its FOV covers multiple lanes in one traveling direction leading to the junction and must have the view of at least one traffic light at the junction. The camera settings are assumed to be fixed, without pan, tilt, or zoom actions. It is also assumed that the camera operates under normal day light and road conditions. The road direction and the number of lanes are assumed known *a priori*.

A functional block diagram of the RLV methodology is depicted in Fig. 2. With reference to Fig. 2, the digital video is first analyzed by the traffic light sequence detection (TLSD), which consists of two subfunctions. The first one is the traffic light detection that locates the spatial coordinates of the centers of the lights. In practice, this detection requires multiple frames as the different colored lights appear at different time. Once the coordinates of the lights are calculated, the set of lights with the largest dimension is chosen, from which the light sequence is constructed in the second subfunction. The traffic light is then continually monitored and the light sequence updated over the entire video. The values output from the light sequence construction are *stop*, *RTG*, *go*, and *RTS*.

Second, the input video is simultaneously analyzed by the vehicle motion estimation (VME), which consists of three subfunctions namely: stop line detection (SLD), virtual loop detectors assignment (VLDA) and direction-biased motion estimation (DME). In theory, SLD only needs one frame for the detection, at the initialization stage. However, if the detected line is too short, this may be repeated for the next few frames until a reasonably long line is detected. From the stop line, the VLDA defines a number of VLD within the prohibited zone and near the stop line. The standard block-based motion estimation used in video coding is performed on the VLD along the road direction. The number of VLD initially assigned is dependent on the size of the prohibited zone and the VLD. However, there may be redundant VLD in this set and computation could be prohibitive if motion estimation is performed on all of them. Therefore, a reduction scheme is adopted based on the mean and standard deviation of the motion vectors (MV) across a small number of frames with vehicle motions. Only those with MV along the road direction and have magnitude larger than the mean and small standard deviation are retained.

Given a set of VLD, the knowledge of the road direction and a stop signal from the TLSD, the DME estimates the motions of the VLD between two consecutive frames. This approach attempts to find the best match in terms of mean-absolute-difference in the forward direction of the road only, for each VLD. The estimated MV are then considered corporately for determining whether they together represent a moving vehicle. The criteria employed here are that if more than half of the MV are larger than half of the mean MV of their corresponding VLD, and are in the direction of the road, then they represent a moving vehicle. This implies a vehicle is moving in the prohibited zone while the light is red or an RLR. The first frame satisfying the criteria marks the beginning of the violation event. The search continues for the next pair of frames until one of the criteria is not satisfied, which marks the end of the event. Fig. 3 illustrates the event flow of the method.

IV. ALGORITHMIC DETAILS

A. Traffic Light Detection

In traffic light detection, a number of possible approaches may be adopted. The common one would be to use edge detectors. The brightness of the traffic light should enable it to be

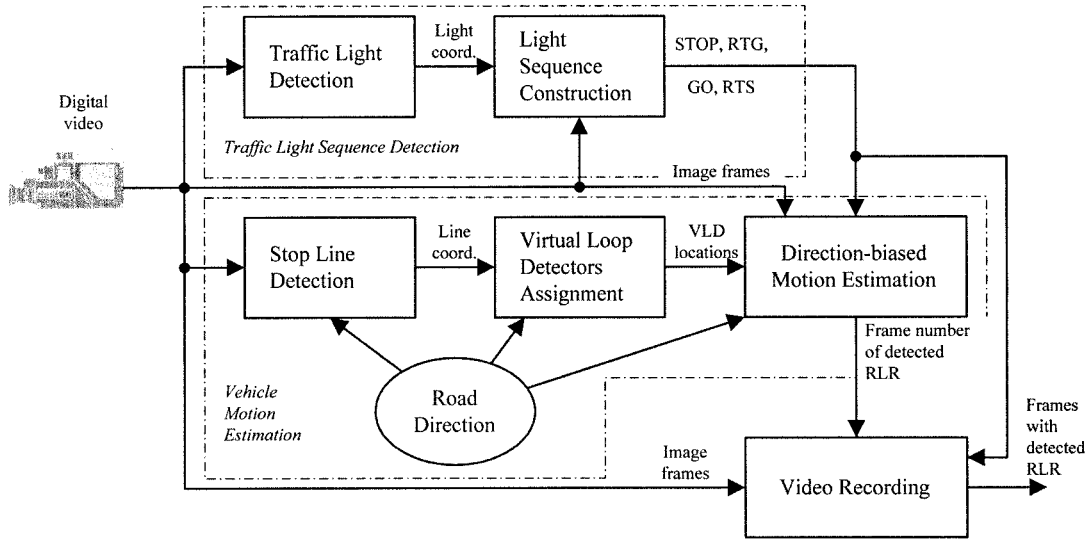


Fig. 2. Functional block diagram of the RLV methodology.

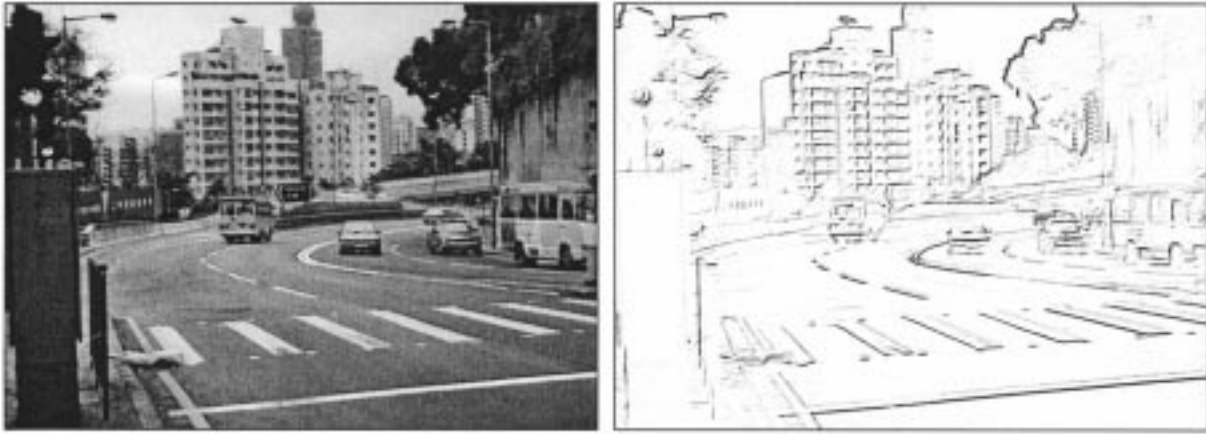


Fig. 3. Edge-detected traffic light.

detected with reasonable ease. However, it would be difficult to first ensure that the edge points representing a traffic light is a close boundary. Second, it is further processing is required to identify the location of the light out of the usually complex edge map (Fig. 3). On the other hand, region detection allows color to be taken into account. Besides, region detection always ends up with closed regions, which require no further processing. For this reason, our approach focuses on the detection of colored regions.

Let us assume that the video consists of color images, I_t , at time t is defined in the hue–saturation–brightness (HSV) space as given by (1)

$$I_t = \left\{ f_t(x, y) = \begin{bmatrix} f_{t,H}(x, y) \\ f_{t,S}(x, y) \\ f_{t,V}(x, y) \end{bmatrix} \right\} \quad (1)$$

where (x, y) denotes the image spatial coordinates and $f_{t,H}(x, y)$, $f_{t,S}(x, y)$, $f_{t,V}(x, y)$ denote the hue, saturation,

and brightness components, respectively, which can be calculated from the RGB components as follows [24]:

$$\begin{aligned} f_{t,H} &= \cos^{-1} \\ &\times \left[\frac{\frac{1}{2}(f_{t,R} - f_{t,G}) + (f_{t,R} - f_{t,B})}{\sqrt{(f_{t,R} - f_{t,G})^2 + (f_{t,R} - f_{t,B})(f_{t,G} - f_{t,B})}} \right] \\ f_{t,V} &= \frac{1}{3}(f_{t,R} + f_{t,G} + f_{t,B}) \\ f_{t,S} &= 1 - \frac{1}{f_{t,V}} \min(f_{t,R}, f_{t,G}, f_{t,B}). \end{aligned} \quad (2)$$

If $f_{t,S} = 0$, then $f_{t,H}$ becomes meaningless. If $(f_{t,B}/f_{t,V}) > (f_{t,G}/f_{t,V})$ then $f_{t,H} = \pi - f_{t,H}$. The white, red, yellow, and green regions are defined as follows:

$$\begin{aligned} \text{White region : } R_{W,k} &= \{ f_t(x, y) : f_{t,S}(x, y) < t_s \\ &\quad \text{and } f_{t,V}(x, y) > t_v \} \end{aligned} \quad (3a)$$

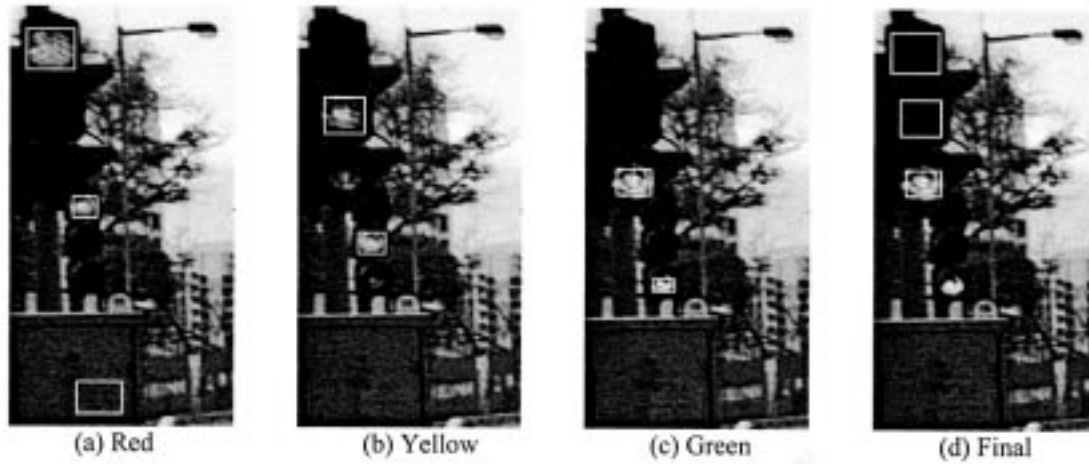


Fig. 4. Traffic light detection. (a) Red. (b) Yellow. (c) Green. (d) Final.

$$\text{Red region : } R_{R,k} = \left\{ f_t(x, y) : \begin{aligned} &|f_{t,H}(x, y)| < t_h \\ &\text{and } f_{t,V}(x, y) > t_v \end{aligned} \right\} \quad (3b)$$

$$\text{Yellow region : } R_{Y,k} = \left\{ f_t(x, y) : \begin{aligned} &\left| f_{t,H}(x, y) - \frac{\pi}{3} \right| < t_h \\ &\text{and } f_{t,V}(x, y) > t_v \end{aligned} \right\} \quad (3c)$$

$$\text{Green region : } R_{G,k} = \left\{ f_t(x, y) : \begin{aligned} &\left| f_{t,H}(x, y) - \frac{2\pi}{3} \right| < t_h \\ &\text{and } f_{t,V}(x, y) > t_v \end{aligned} \right\} \quad (3d)$$

where t_s , t_v and t_h are threshold values; $R_{\psi,k}$ are connected regions that $R_{\psi,k} \cap R_{\psi,j} = \phi$, for all k and j that $k \neq j$ and $\psi = W, R, Y, G$. For instance, hue for red, yellow and green are 0 , $\pi/3$ and $2\pi/3$, respectively. The reason for defining the color range using values such as t_s , t_v or t_h in (3) is that because in outdoor environment, color may appear slightly different from its predefined value, and therefore, allowing for small variations helps to accommodate the nonideal situation. The use of t_v and t_h also ensures that their effect on the size of the detected regions is similar. In the actual implementation, we have $t_s = t_v = 0.7$ for $0 \leq f_{t,v}(x, y) \leq 1$ and $t_h = \pi/6$.

B. Traffic Light Sequence Detection

There are three steps in the traffic light detection. The first step is to locate the position of the *red* lights. This is achieved by using (3b) over a number of image frames. Assume that in each frame, n_R number of *red* regions $R_{R,k}$ for $k = 1, \dots, n_R$ are detected from which the aspect ratio of $R_{R,k}$ is given by

$$\alpha(R_{R,k}) = \frac{\text{width}(R_{R,k})}{\text{height}(R_{R,k})} \quad (4)$$

where $\text{width}(R_{R,k})$ and $\text{height}(R_{R,k})$ denote the width and height of $R_{R,k}$, respectively. As some of these *red* regions are due to other objects in the FOV, this aspect ratio can be used to discriminate them from the red lights by considering only those

$R_{R,k}$ with $\alpha(R_{R,k}) \approx 1$. $R_{R,k}$ is not exactly one as the images are taken in perspective view.

The second step is to locate the *yellow* regions by (3c) in a similar manner, where $R_{Y,j}$ with $0.9 < \alpha(R_{Y,j}) < 1.1$ are retained for $j = 1, \dots, n_Y$ and n_Y is the number of *yellow* regions detected. As the *red* and *yellow* lights are spatially related $R_{R,k}$ and $R_{Y,j}$ represent a potential traffic light pair, if and only if: 1) $R_{Y,j}$ is vertically and immediately below $R_{R,k}$; 2) $\text{width}(R_{R,k}) \approx \text{width}(R_{Y,j})$; and 3) $\text{height}(R_{R,k}) \approx \text{height}(R_{Y,j})$.

The third step is to locate the *green* lights using a different approach because: 1) their brightness and contrast are usually low; 2) they possibly come in different size and shape; and 3) many roadside objects are green too. For each potential traffic light pair obtained from the second step, a region is estimated vertically below the *yellow* light with center-to-center distance being the same as the *red*-*yellow* pair, and of the same size. Within this region, *green* subregions are searched using (3d). If *green* subregions are found, a complete traffic light $TL_m = \{R_{R,m}, R_{Y,m}, R_{G,m}\}$ is detected, where TL_m denotes the m th traffic light in the frame. In the rare circumstances that the green light cannot be found in the estimated region, the traffic light pair will be discarded. From all the traffic lights found, the one with the largest size, i.e., closest to the camera, is selected for subsequent analysis. Fig. 4 depicts an example of traffic light detection in the real world.

To construct the light sequence from the detected traffic light, we compute the average brightness of the lights when they are first detected

$$\bar{B}_R = \frac{1}{S_R} \sum f_{t,V}(x, y) \Big|_{f(x,y) \in R_R} \quad (5a)$$

$$\bar{B}_Y = \frac{1}{S_Y} \sum f_{t,V}(x, y) \Big|_{f(x,y) \in R_Y} \quad (5b)$$

$$\bar{B}_G = \frac{1}{S_G} \sum f_{t,V}(x, y) \Big|_{f(x,y) \in R_G} \quad (5c)$$

where S_R , S_Y and S_G are the size of the *red*, *yellow* and *green* regions. At frame F , the average brightness of the lights $\bar{B}_{R,F}$,

$\bar{B}_{Y,F}$, and $\bar{B}_{G,F}$ are computed using (6), from which the state of the traffic light is calculated as

$$Red = \begin{cases} ON, & \bar{B}_{R,F} > \frac{1}{2}\bar{B}_R \\ OFF, & \text{otherwise,} \end{cases} \quad (6a)$$

$$Yellow = \begin{cases} ON, & \bar{B}_{Y,F} > \frac{1}{2}\bar{B}_Y \\ OFF, & \text{otherwise,} \end{cases} \quad (6b)$$

$$Green = \begin{cases} ON, & \bar{B}_{G,F} > \frac{1}{2}\bar{B}_G \\ OFF, & \text{otherwise.} \end{cases} \quad (6c)$$

C. Vehicle Motion Estimation

It should be noted that the stop line only needs to be detected once. As the camera parameters are fixed during the operation, the location of the stop line is fixed, which is used to define the locations of the virtual loops. As it is not used for subsequent analysis, even if the stop line is blocked by vehicles, it imposes no restrictions to the detection of RLR. To detect the stop line, many detection and segmentation methods may be employed. In general, a frame where the stop line is not obstructed must be obtained first. This usually would be a background frame obtained from background estimation methods [22]. From the background, edge detection, or region detection as described by (3) may be used. The difficulties in using edge detection is similar to the traffic light detection, although in this case, the stop line is expected to have an additional length feature that would probably help in the detection accuracy. However, the number of edge points could be large without color discrimination. Therefore, it is also appears that detecting white regions would eliminate the other colored features first. The thinning of the regions and Hough transform may well be applied on the edge map too (Fig. 5).

In our case, the white regions $R_{W,k}$ are detected using (3a), for $k = 1, \dots, n_W$, where n_W is number of white regions detected. Then, the detected regions $R_{W,k}$ are thinned and transformed to the ρ - θ space using Hough transform from which the length $l_{W,k}$ and orientation $\theta_{W,k}$ of the thinned regions (lines) are calculated from the intersecting points. The stop line is determined as the line satisfying the conditions: $\max(l_{W,k})$ and $\max(|\theta_{W,k} - \theta_{road}|)$ where θ_{road} is the road direction, which can be estimated from the direction of motions of the vehicles in an image sequence. This detects the longest line that sustains the largest angular deviation from the road direction as part of the stop line, as it is 90° to the road direction in the real world coordinate. Once identified, the stop line is extended across the FOV to eliminate any possible discontinuities along the line (Fig. 6).

Let us define the VLD in frame F as $N \times N$ blocks given by $S_i(x_i, y_i, F)$ for $i = 1, \dots, M$, where M is the number of VLD in the frame and (x_i, y_i) is the spatial coordinates of the VLD center. The initial assignment of these VLD starts from the stop line at the bottom left-hand corner in the prohibited zone. The prohibited zone is bounded by the stop line and two lines in the road direction that intersect the stop line and the traffic lights as depicted in Fig. 7(a). The fourth line of the prohibited zone is not defined here. The advantage of this is that the number of VLD can be extended right into the junction to take into account of turning vehicles. However, here is apparently no upper bound

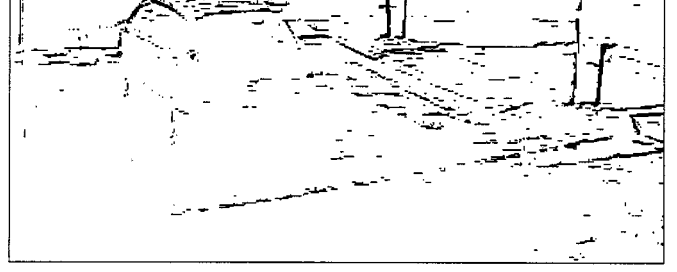


Fig. 5. Edge-detected stop line.

to M , where M is chosen based on N and the number of lanes known *a priori*. If the junction has a large number of lanes, M may become very large. Similarly, M is inversely proportional to N . Typically, M is found to be between 100 and 200. The cases of $M = 100, N = 15$ and $M = 200, N = 13$ are depicted in Fig. 7(b) and (c). The difference between the two cases is that for $M = 100$, the VLD are bigger and are further away from the stop line. For $M = 200$, the VLD are smaller. It is observed that N should not be larger than 15 because large N tends to exaggerate the panning out effect and result in poor match. On the other hand, N should not be too small (< 8) either as M increases with small N causing an increase in computing delay.

In the event that the number of VLD is too large or some of the VLD are assigned outside the road (e.g., pavement), the following reduction approach is used. For an image sequence of N_v frames with vehicles passing the junction and assuming that there is only translation motion between consecutive frames, F and $F+1$, the mean-absolute-difference (MAD) of the i th VLD is given by

$$MAD(x_d, y_d) = \frac{1}{N^2} |S_{i,F}(x_i, y_i) - S_{i,F+1}(x_i + x_d, y_i + y_d)| \quad (7a)$$

$$S_{i,F}(x_i, y_i) = \begin{cases} (x, y) \in I_F | x = x_i, \dots, x_i + N - 1; \\ y = y_i, \dots, y_i + N - 1 \end{cases} \quad (7b)$$

where $S_{i,F}(x_i, y_i)$ denotes the i th VLD in frame F centered at (x_i, y_i) , I_F is the image of frame F and $x_d, y_d : |x_d, y_d| < r$, $|\angle(x_d, y_d) - \theta_{road}| \leq \beta$, where r is the search range and β is a constrained search angle. The values of r and β are dependent on the camera view, frame rate and the maximum allowable vehicle speed of the road. From (7), the motion vector of the i th VLD is calculated as

$$v_i(F, F+1) = (x_d, y_d)^T = \arg \left(\min_{(x_d, y_d)} MAD(x_d, y_d) \right). \quad (8)$$

The mean and standard deviation of the motion vectors' magnitude ($|v_i|$) and orientation ($\angle(v_i)$) of the i th VLD are computed over N_v as follows:

$$\mu_{|v_i|} = \frac{1}{N_v - 1} \sum_{f=0}^{N_v-2} |v_i(f, f+1)| \quad (9a)$$

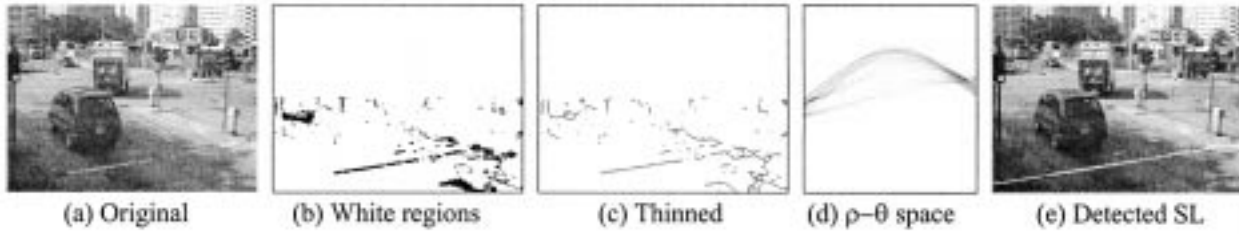


Fig. 6. Region based stop line detection.

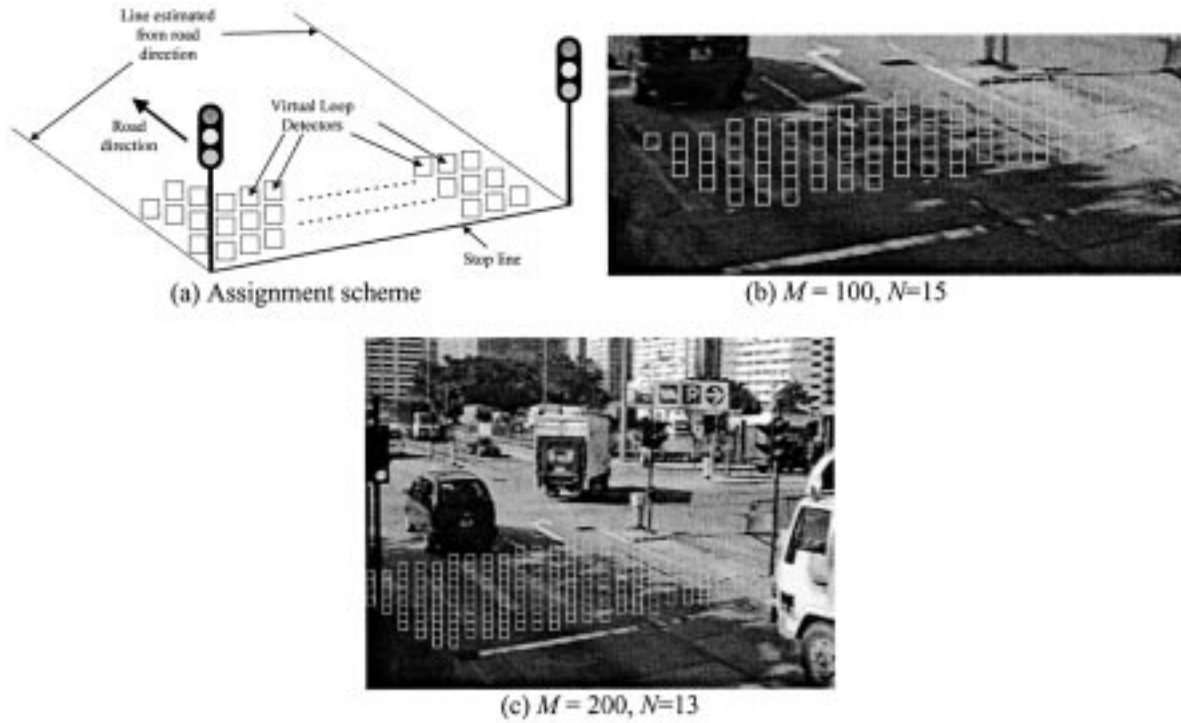


Fig. 7. Initial assignment of VLD.

$$\sigma_{|v_i|} = \sqrt{\frac{1}{N_v - 1} \sum_{f=0}^{N_v-2} [v_i(f, f+1) - \mu_{|v_i|}]^2} \quad (9b)$$

$$\mu_{\angle v_i} = \frac{1}{N_v - 1} \sum_{f=0}^{N_v-2} \angle v_i(f, f+1) \quad (9c)$$

$$\sigma_{\angle v_i} = \sqrt{\frac{1}{N_v - 1} \sum_{f=0}^{N_v-2} [\angle v_i(f, f+1) - \mu_{\angle v_i}]^2} \quad (9d)$$

To select the final set of VLD for vehicle motion estimation, all the initial VLD are examined as follows. First, those VLD having their mean orientation not in the road direction are discarded. Second, the remaining VLD with large orientation standard deviation are also be discarded. Finally, only those VLD with large mean magnitude and small magnitude standard deviation are selected into the final set for subsequent estimation. Fig. 8 depicts these chosen VLD divided into two groups for two lanes.

To estimate vehicle motions in the junction, we first compute the motion vectors $v_i(F, F+1)$ over the VLD set, as depicted

in Fig. 6. For detecting red light runners, we have the following conditions for each VLD:

$$c_i = \begin{cases} 1, & |v_i(F, F+1)| > \frac{1}{2}\mu_{|v_i|} \quad \text{and} \\ & |\angle v_i(F, F+1) - \theta_{\text{road}}| < \sigma_{\angle v_i} \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

If $\sum_{i=0}^{N_B-2} c_i > 1/2N_B$ then a red light runner is detected in that lane, where N_B is the number of final VLD per lane. In essence, if its MV magnitude is larger than half of the mean MV magnitude calculated previously and its direction is aligned with the road direction, then it is possible that part of a vehicle moves over this VLD. If more than half of the VLD satisfy this condition, then a moving vehicle is declared. It should be pointed out that the VLD do not differentiate whether their detected motion is due to vehicles actually passing over them, or due to movement of other kinds such as moving shadow. This explains why the criteria of $|v_i(F, F+1)|$ must be at least half of the mean magnitude (to eliminate motion due to shadow) and its direction must be approximately in the direction of the road (to eliminate pedestrian motion). On the other hand, if the motion is due to

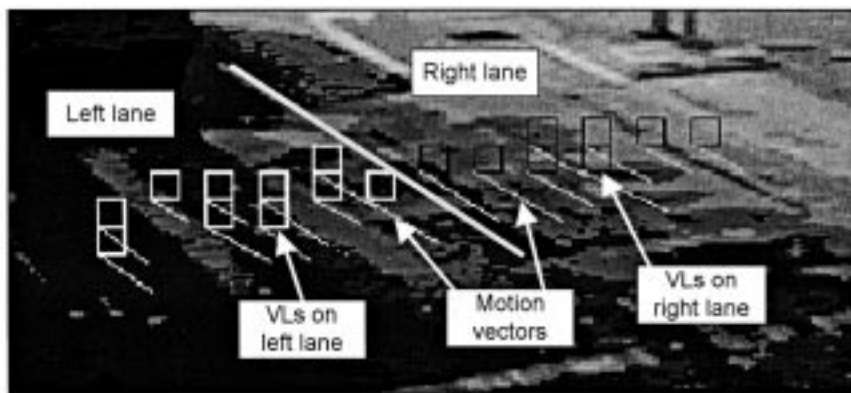


Fig. 8. Final VLD set (ten for left lane, eight for right lane).

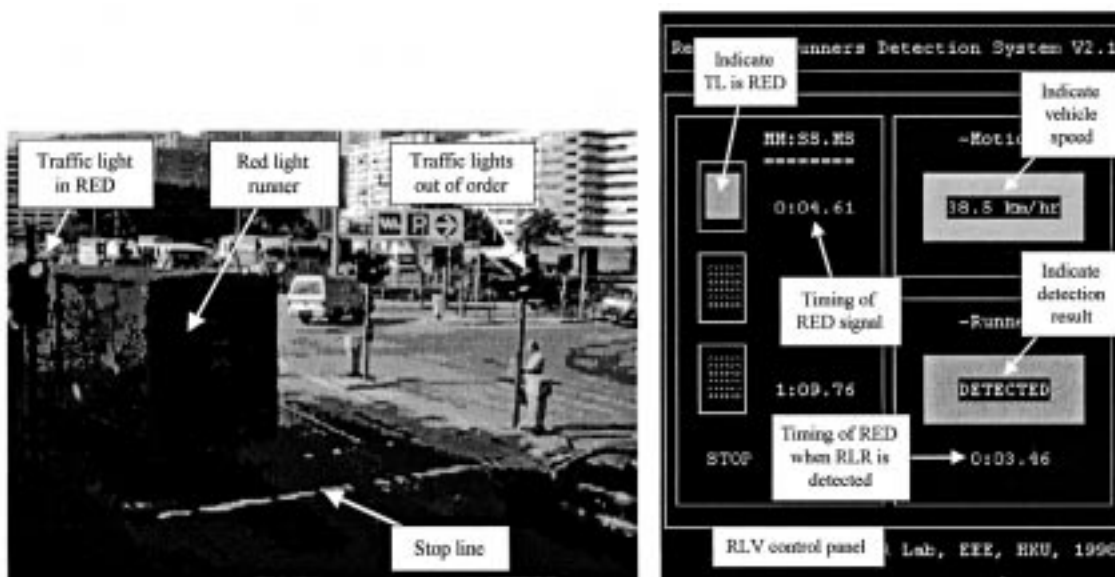


Fig. 9. Red light runner truck.

vehicles not on the road (e.g., moving vehicle on the pavement), this method will fail to detect the difference and the result will become a false alarm. As vehicles normally do not drive on the pavement, the percentage of false alarm due to this is believed to be small.

V. TRIALS AND RESULTS

A prototype was implemented based on the above methodology. The hardware comprises of a commodity 3CCD digital video camera, frame grabber and a Pentium PC. The software was written in C that operates on a 640 × 480 true color video at 11 frames/s. The number of initial VLD, *M*, used was 100 and the VLD block size *N* was 15 and the final set of VLD was depicted in Fig. 8.

The trial site is a typical cross junction in Hong Kong with two lanes in each direction, where a conventional RLC post has been installed. The new prototype was mounted on the camera post, without any connections to the traffic light control system or the ILD. The trial was conducted on a sunny afternoon where shadow from a nearby tree is evident, and the shadows of vehicles falling onto adjacent lanes were also detected. After the

prototype was setup, it was left running for 40 min without interruption. The whole event was recorded on tape for off-line inspection. During the 40 minutes, the traffic lights have changed 20 times and 506 vehicles have passed through the junction. Under normal circumstances, there should be three traffic lights visible under the FOV. However, two of the lights were out of order on the day. The prototype detected the traffic light on the left and used it to build the light sequence. During the trial, the traffic light changed a number of times where fast and slow moving vehicles of all kinds were seen using the junction, e.g., container trucks, fire engines, private cars, scooters, vans, and medium-size goods truck. There were also pedestrians crossing the road, empty handed or pushing a cart. Over this period, two red light runners were detected and these were confirmed by visual inspection. The correction detection rate was 100% and the missing rate is 0% with no false alarm.

Fig. 9 depicts one of the cases of red light running by a truck. It should be noted that there was only one traffic light visible, the contrast of the stop line is poor due to the tree shadow, and the tree shadow caused pseudomotion because of an unexpected breeze. Even under these conditions, the stop line was detected as in Fig. 6 and a set of VLD was reduced as depicted in Fig. 8.

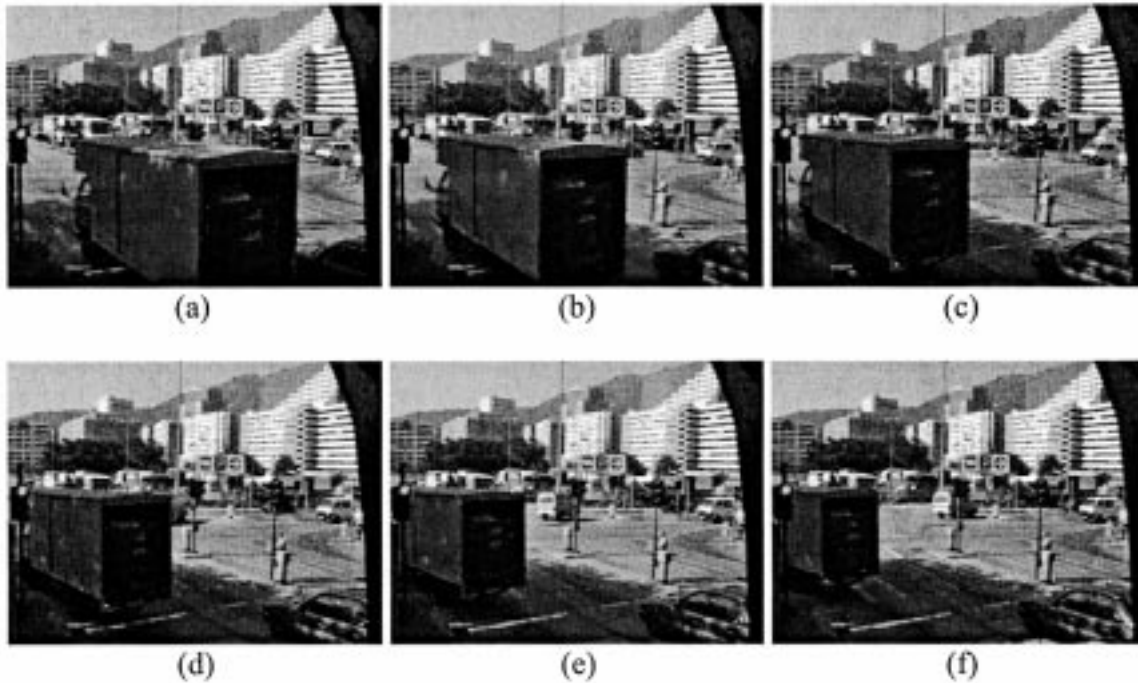


Fig. 10. Red light runner sequence.

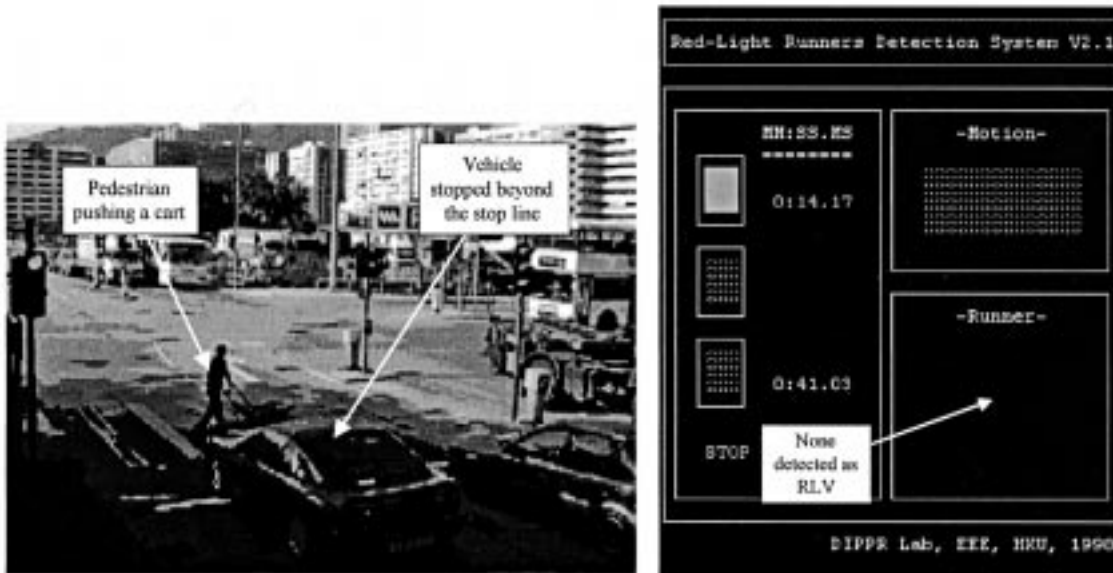


Fig. 11. Pedestrian pushing a cart while crossing the road.

The RLR was correctly identified and the information on its instantaneous speed and the detected time (3.46 s after the light turned red) was also recorded. The image on the left depicts the truck driving clearly beyond the stop line while the light is red. Off-line inspection of the video confirmed that the truck continued to turn left and exited the junction while the light was still red. Fig. 10 depicts the image sequence of this event.

Fig. 11 depicts a case of a pedestrian crossing the road while pushing a cart, and a vehicle crossed over the stop line, but had stopped. The pedestrian with the cart in fact appears as a large object that is moving beyond the stop line. However, this is not considered a RLR as the method is able to differentiate it from

other vehicles, using the motion vector direction. For the vehicle stopped beyond the stop line, as it did not have any motion, the method did not consider it a RLR. Apart from this scenario, there were other pedestrians crossing the road and vehicles turning left or right legitimately. The method performed correctly and robustly.

VI. CONCLUSION

In conclusion, we have presented a novel and effective method for detecting red light runners in this paper. A number of features have been observed. First, without needing any

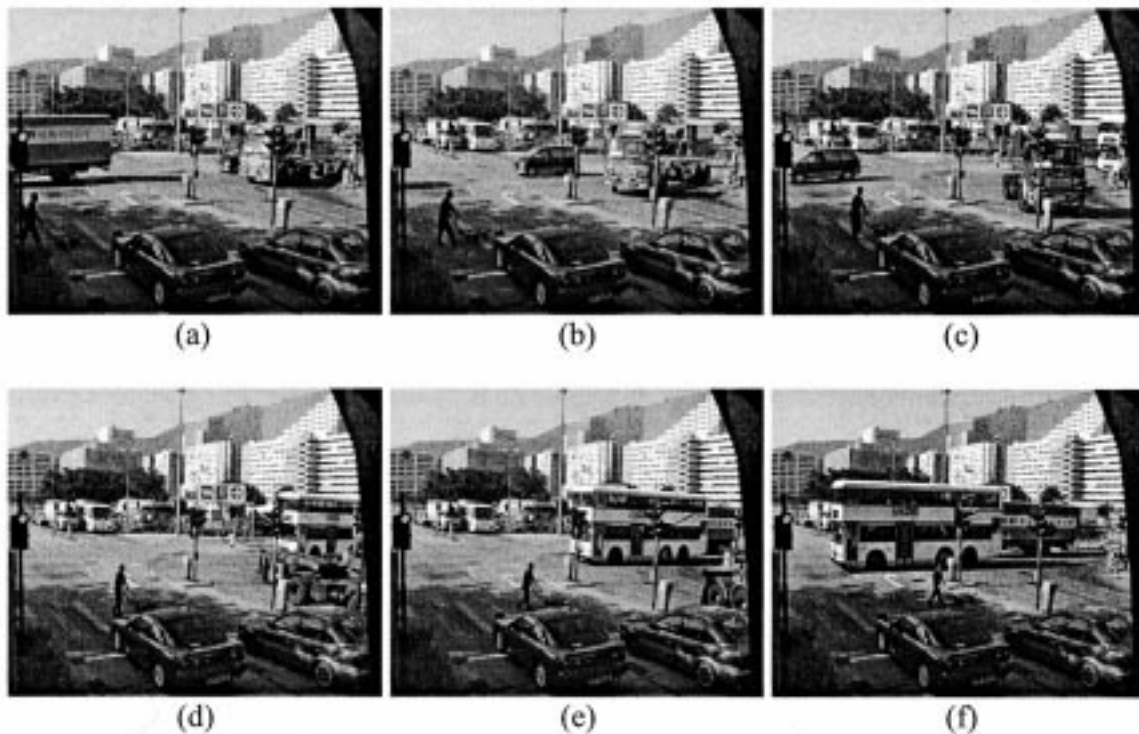


Fig. 12. Sequence of Fig. 11.

connections to the traffic light and buried loop detectors, this method offers a mobile, robust, and cost-effective solution. One of our other trials was to have the prototype mounted on a camera tripod located by the roadside. The set up process was simple and fast, while the correct detection rate was again 100%, 0% missing rate with no false alarm. Second, the new method provides more than ten frames of the RLR incident as compared with only two photos in the RLC. It provides an exact record of the event over time. This is also advantageous if automatic license plate recognition is to be performed. Third, the video analysis concept employed in the new method can be further developed into a multifunctional architecture where other traffic parameters may be calculated or estimated, e.g., vehicle speed and count, which is partially demonstrated in the prototype.

From the trial results, we also observed that the new method is able to cope with a number of hostile but realistic situations. First, it handles the traffic light detection very well under minimum condition. With only one traffic light in view, the method detected it and built the light sequence from it. Second, it deals with the pseudomotions and poor contrast due to the tree shadow appropriately and adequately. Such pseudomotions were the result of tree movement, which affected the accuracy of the motion vectors. However, this is tolerated by using the motion estimation criteria. The poor contrast has adverse effect on the stop line detection and the virtual loop detectors assignment. The trial results showed that as long as the stop line is there, it can be detected even under such adverse condition. Third, the scenario of pedestrian crossing the road is correctly handled by the method. In the incident of a pedestrian pushing a cart while crossing the road, a number of VLD would have been affected because of their locations on the crossing. Constraining the vehicle motion

vectors to be along the road direction simply eliminates the possibility of miscounting pedestrian motion as vehicle motion.

ACKNOWLEDGMENT

The authors would like to thank the Traffic Management Bureau of the Hong Kong Police of the HKSAR and the Road Safety and Standards Division of the Transport Department of the HKSAR.

REFERENCES

- [1] K. M. Lum and Y. D. Wong, "An overview of red-light surveillance cameras in Singapore," *ITE J. Web*, pp. 87–91, 1998.
- [2] C. J. Baguley, "Running the red at signals on high-speed roads," *Traffic Eng. Contr.*, vol. 29, no. 7, pp. 415–420, 1988.
- [3] H. C. Chin, "Effect of automatic red light cameras on red-running," *Traffic Eng. Contr.*, vol. 30, no. (4), pp. 175–179, 1989.
- [4] Hong Kong—Installation of LRT red light cameras completed. Institute of Transportation Engineers. [Online]. Available: <http://www.ite.org/hongkong.htm>
- [5] American traffic systems—Safety results [Online]. Available: <http://www.traffic.com/cameraresults.htm>
- [6] S. J. Thompson *et al.*, "Putting red-light violators in the picture," *Traffic Eng. Contr.*, vol. 30, no. (3), pp. 122–125, 1989.
- [7] Annu. 4th quarter 1997 traffic report. Hong Kong Police Force, HKSAR. [Online]. Available: <http://www.info.gov.hk/police>
- [8] Gatsometer [Online]. Available: <http://ourworld.compuserve.com/homepage/funpages/gatso.htm>
- [9] AutoPatrol RL-200. american traffic systems. [Online]. Available: <http://www.traffic.com>
- [10] Driver Safety Systems Ltd. [Online]. Available: <http://www.dss.co.il>
- [11] "MULTAFOT—Fully automatic red light surveillance system; Technical data sheet," Zellweger Uster AG, Switzerland.
- [12] Redguard. Control Technologies of Central Florida, Inc.. [Online]. Available: <http://www.cttraffic.com>
- [13] Autoscope™—Wide area video vehicle detection system. Autoscope Brochure. [Online]. Available: <http://www.autoscope.com/as-brochure.html>

- [14] VideoTrak-900—Video vehicle tracking and detection system. Peek Traffic Systems. [Online]. Available: <http://www.peek-traffic.com/transyt/products/machine/vt900>
- [15] M. Irani and P. Anandan, "A unified approach to moving object detection in 2-D and 3-D scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, pp. 577–589, 1998.
- [16] N. Hoose, *Computer Image Processing in Traffic Engineering*. London, U.K.: Research Studies, 1991.
- [17] M.-P. Dubuisson and A. K. Jain, "Contour extraction of moving objects in complex outdoor scenes," *Int. J. Comput. Vision*, vol. 14, pp. 83–105, 1995.
- [18] D. Koller *et al.*, "Model-Based object tracking in monocular image sequences of road traffic scenes," *Int. J. Comput. Vision*, vol. 10, no. 3, pp. 257–281, 1993.
- [19] F. Bartolini *et al.*, "Motion estimation and tracking for urban traffic monitoring," in *Proc. Int. Conf. Image Processing'96*, vol. III, Lausanne, Switzerland, Sept. 1996, pp. 787–790.
- [20] H. S. Lai and N. H. C. Yung, "A system architecture for visual traffic surveillance," in *Proc. 5th World Congress Intelligent Transportation Systems*, Seoul, Korea, Oct. 1998.
- [21] W. F. Gardner and D. T. Lawton, "Interactive model-based vehicle tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, pp. 1115–1121, Nov. 1996.
- [22] N. H. C. Yung and A. H. S. Lai, "A fast and accurate scoreboard algorithm for estimating stationary backgrounds in an image sequence," in *Proc. IEEE ISCAS'98*, Monterey, CA, 31 May–3 June 1998, Catalog #98CH36187.
- [23] A. M. Tekalp, *Digital Video Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [24] J. D. Foley *et al.*, *Computer Graphics: Principles and Practices*, 2/e. Reading, MA: Addison-Wesley, 1990.

Nelson H. C. Yung (S'82–M'85–SM'96) received the B.Sc. and Ph.D. degrees from the University of Newcastle Upon Tyne, U.K., in 1982 and 1985, respectively.

He was lecturer at the University of Newcastle Upon Tyne from 1985 to 1990, where he was involved in the research and development of digital image processing and parallel processing. From 1990 to 1993, he became Senior Research Scientist at the Department of Defence, Australia, where he headed a team on the research and development of military-grade signal analysis systems. He joined the University of Hong Kong in late 1993 as an Associate Professor where he leads the Digital Image Processing and Intelligent Transportation Systems Research Group. He is the founding Director of the Laboratory for Intelligent Transportation Systems Research. He has published over 90 research papers. He is a reviewer for the *Institute of Electrical Engineers (IEE) Part G*, *SPIE Optical Engineering*, *HKIE Proceedings*, *Microprocessors and Microsystems Journal*, and *Signal Processing*.

Dr. Yung is a Chartered Electrical Engineer and a member of the HKIE and IEE. He is a Reviewer for the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS. His biography has been published in the *Who's Who in the World*.

Andrew H. S. Lai (M'00) received the B.Eng. degree from the University of Hong Kong, the M.Sc. degree from the University of Surrey, U.K., and the Ph.D. degree from the University of Hong Kong, in 1994, 1995, and 2000, respectively.

He is currently a Research Associate with the University of Hong Kong and a Senior Research Engineer with the Laboratory for Intelligent Transportation Systems Research, University of Hong Kong. His research interests include visual traffic surveillance, digital image processing, and wireless data communication.