

An Adapted Point Based Tracking for Vehicle Speed Estimation in Linear Spacing

H. A. Rahim, R. B. Ahmad, A. S. M. Zain

School of Computer and Communication Engineering
Universiti Malaysia Perlis
Kangar, Perlis, Malaysia
haslizarahim@unimap.edu.my

U. U. Sheikh

Computer Vision, Video and Image Processing Research
Lab (CVVIP), Department of Microelectronics and
Computer Engineering
Universiti Teknologi Malaysia
Johor Bharu, Johor, Malaysia

Abstract— Vehicle velocity estimation is an important aspect of intelligent transportation systems. Normally velocity is estimated using dedicated laser speed traps and Doppler radars. Recently, the use of cameras is becoming more common for the purpose of traffic surveillance and smart surveillance system. It is thus the aim of this paper to propose a method for vehicle speed estimation using these existing video cameras. In this paper, we propose a vehicle speed estimation method from video analysis. The method proposed contains several steps; image pre-processing, centroid extraction and tracking. The proposed method transforms the 2D image points into a 3D virtual world to obtain actual vehicle position in 3D space. This is to account for perspective distortion commonly seen in images. Using these 3D points and measuring the time for displacement, the vehicle speed is obtained. Experimental results have shown that the proposed method gives accurate velocity estimation.

Keywords-component; vehicle speed estimation; adapted point based tracking; image processing

I. INTRODUCTION

Smart surveillance system has been defined as the system that utilizes computer vision and pattern recognition technologies for information analysis [1]. The smart surveillance system research development is rapidly growing since the need for more reliable security in public places such as railways, restaurants, schools, shopping malls is accelerating [1-5]. Due to the increasing terrorist attack and crime activities which occurs in our daily life, the recent world events have shifted from “investigation of incidents” to prevention of potentially catastrophic incidents” [2] that require minimum supervision from the human where surveillance activities are mainly carrying out by the computers or embedded systems [1-11].

The smart surveillance system not only monitors the human activities, but it also monitors the environment around it such as moving vehicles and etc. The techniques which are commonly used in smart surveillance are several image processing tasks such as moving object detection, object recognition, object tracking as well as behavior and activities analysis [3]. The events in parking lot area which mostly involve vehicles drive through the parking lot also become

crucial for us to monitor as there are numbers of accidents which happen in the parking lot due to the vehicles' over speed limit. Therefore, vehicle speed estimation algorithm has been designed to be one of the functions of smart surveillance system that able to monitor the vehicles in the parking lot by finding the speed of every vehicle that passes through the parking lot.

Some work on object tracking techniques used to track moving object, has been published. Tai et al. [7] utilized active contour model technique to detect location of vehicle that is integrated with Kalman filtering method to track the vehicle motion. Bramberger [8] proposed a complex scalable multiprocessor architecture targeted for traffic surveillance. This new approach presents a good performance but with a very high computational cost. Bue [6] introduced mean-shift technique to track objects with MPEG-4 compression for real-time surveillance system. Recently, new methods have been introduced in the same field. Litzenberger [5] developed a surveillance system using an asynchronous temporal contrast vision for high speed vision. The pixel array responds in real-time to relative changes in light intensity. A continuous clustering of Address-Event (AE) and tracking of clusters are used for tracking algorithm.

This paper mainly describes about vehicle speed estimation algorithm [11] that utilizes adapted point based tracking technique for object tracking. The performance of the proposed algorithm has been tested on real video sequence and some analysis has been performed.

The paper is organized as follows. Section 2 explains the derivation of vehicle speed/velocity from the kinematic equation. Section 3 outlines the velocity estimation algorithms which include centroid extraction and object tracking. Section 4 discusses the experimental results and Section 5 presents the conclusion of proposed algorithm.

II. VEHICLE DETECTION

The vehicle speed algorithms is acquired from the kinematics equation where a vehicle is assumed to act as an object with mass, and the forces react on it is ignored. In order

to estimate the velocity of the detected vehicle, a set of n 2D coordinates is obtained from video frame: $[\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n] \in \mathbb{Z}^2$. These 2D coordinates represents 3D coordinates in world space transformed to screen space $[\gamma_1, \gamma_2, \gamma_3, \dots, \gamma_n] \in \mathfrak{R}^3$. The transformation of camera viewpoint and projective transformation have caused the transformation operation of a 3D homogenous point in world space to 2D point in an image/screen space [12]. From the kinematic equation, the velocity is derived as the rate of change in the displacement with respect to the change in time, Δt by considering any two points λ_n and λ_{n-1} [11]:

$$v_\tau = \frac{\lambda_n - \lambda_{n-1}}{\tau_n - \tau_{n-1}} \quad (1)$$

$$v_\tau = \lim_{\Delta\tau \rightarrow 0} \frac{\Delta\lambda}{\Delta\tau} = \frac{dL}{dt} \quad (2)$$

The unit of computed velocity in (1) is in pixels/second (pixels s^{-1}) as the λ_n represents the pixel coordinate of the image, extracted from the video frame. Thus $\lambda_n = (x, y)$ is in pixels. Nonetheless, the velocity unit must be changed to the actual world unit, $v = m s^{-1}$, meters/second (or km/hour).

In order to convert the unit of velocity to actual world unit, the pixel coordinates obtained from the image must be converted to world coordinates in meters. The relationship of 1 pixel to the distance in meters in real-world must be defined. Therefore, a camera calibration process is required in determining the relationship of 1 pixel to the distance in meters [11].

A. Camera Calibration

A plane (ground) must be ‘fit’ to the scene on which the moving vehicles are actually on. The ground is assumed to be flat. The definition of ground as flat plane is denoted as, $\Pi = [\pi_1, \pi_2, \pi_3, \pi_4] \in \mathfrak{R}^3$ and the equation of the plane can be formulated as [12]:

$$\pi_1 X + \pi_2 Y + \pi_3 Z + \pi_4 = 0 \quad (3)$$

$$\sum_i \pi_i \neq 0, \quad \{i=1, \dots, 4\} \quad (4)$$

The plane $\Pi_\infty = [0, 0, 0, 1]^T$ is defined as plane at infinity [12].

B. Ratio between Pixels and Length in Meters

After calibrating the camera and considering the ground as flat plane, the ratio between the pixels and length in meters can then be defined by assuming that the ratio of world to pixel space is given as $\frac{\omega}{\varsigma}$, the velocity in actual world unit is derived as:

$$v_\tau = \frac{(\lambda_n - \lambda_{n-1}) \frac{\omega}{\varsigma}}{\tau_n - \tau_{n-1}} ms^{-1} \quad (5)$$

This is only true if there is no perspective distortion where the length in pixels and the distance in world coordinate are linear. The ratio of $\frac{\omega}{\varsigma}$ across the image must be constant so as to hold the linearity property. To maintain the linear spacing of image space to world space, the camera must be calibrated as stated in Section 2.A

III. VELOCITY ESTIMATION ALGORITHMS

After the velocity formula has been determined, the next processing step is to process the incoming video in order to attain values of λ , and τ . Since normal video camera is used, the video sequence has a frame-rate of between 25 frames-per-second up to 30 frames-per-second that constitutes temporal resolution of 33.33 ms to 40 ms. The velocity can be computed after several steps are carried out; moving object detection, centroid extraction and object tracking. For moving object detection, the frame differencing technique has been adopted to detect the moving object.

A. Centroid Extraction

From the difference image obtained in moving object detection technique, several moving blobs, B will be obtained and these blobs can be identified using “connected component” operation on the image.

For every blob $b \in B$ in the image, the bounding box and centroid, $\lambda = (x_{cb}, y_{cb})$ are extracted. Centroid is computed by calculating the blob’s spatial moment and is given in (6) and (7) [12]. This centroid will be used for tracking object.

$$u_{pq} = \sum_{x_b=0}^{m-1} \sum_{y_b=0}^{n-1} x_b^p y_b^q : (x_b, y_b) \in b, \forall b \in B \quad (6)$$

$$x_{cb} = \frac{u_{10}}{u_{00}}, y_{cb} = \frac{u_{01}}{u_{00}} \quad (7)$$

B. Object Tracking

In the case of many objects in the scene, it is important to track the objects to resolve the correspondence for vehicle speed estimation. The objective of tracking is to establish correspondence between blobs $b \in B_\tau$ at time τ and blobs in the consecutive frames $b \in B_{\tau+n}; n \in \mathbb{Z}^+$. With robust tracking, problems such as partial or full occlusion can be handled. Generally, there are two types of tracking; based on correspondence matching and the other is based on explicit tracking that makes use of position prediction and motion estimation.

In this work, object occlusion is not considered as it can rarely happen. The tracking method proposed in this work is an adapted point-based tracking. Centroid points from previous frame and current frame are associates based on several parameters such as minimum distance or direction. For example, a point in the previous frame, p , will be associates with a point in the current frame based on the minimum Euclidean distance (8). However, in order to determine the

minimum distance, all the distance between reference point p and all other points $c_1, c_2 \dots c_n$ must be established.

$$E_d = \sqrt{(p_x - x_{cb})^2 + (p_y - y_{cb})^2}; \forall b \in B \quad (8)$$

Once the correspondence of the two points in the current and previous is resolved, the displacement can be calculated, $\lambda_n - \lambda_{n-1}$ and then the velocity from (5).

IV. EXPERIMENTAL RESULTS

Experimental results of the proposed vehicle speed estimation are presented to demonstrate the real-time velocity estimation of random vehicles. The algorithm is implemented using C++ and executes on a 2.6 GHz Pentium Intel dual core with 4 GB ram memory and 500 GB hard disk. The experiment was set up to monitor the speed of vehicles moving along the selected road lane. The real-world scene was recorded in 7 seconds [11].

Figure 1 illustrates the demonstration of vehicle estimation algorithm in a single lane. Several video sequence frames were taken for two moving object which were detected by the system for every three frames. In figure 1(a) – (d), a car was just approached into the video frame and it has not been detected yet. The car then moved closer to the video camera in figure 1(e). When the car was first detected, the speed was set to reference point, 0 km/hr before the algorithm started to calculate the car's velocity. Figure 1(e) – (g) shows the speed of the identified car, determined by the algorithm that accelerated from 25 km/hr to 33 km/hr. The red rectangle line with the estimated speed was displayed on the identified car to track the movement of the car. Meanwhile, in Figure 1(h), it is showed that two objects have been detected and tracked where the motorcycle was later to be tracked. This exhibits the capability of the proposed algorithm to track more than one moving object per video frame.

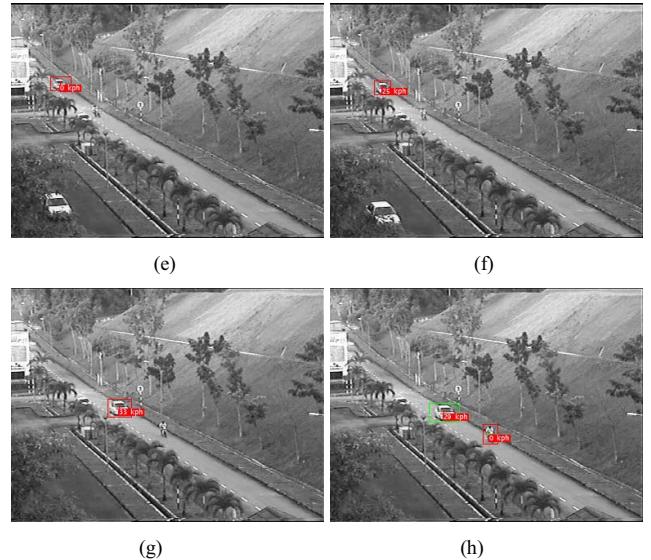
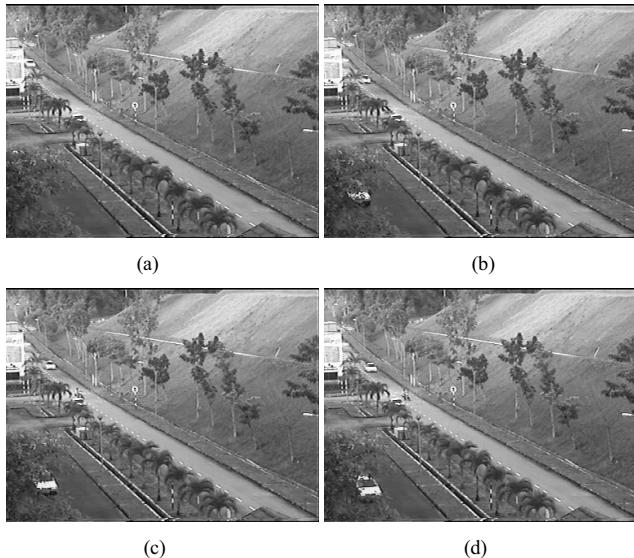
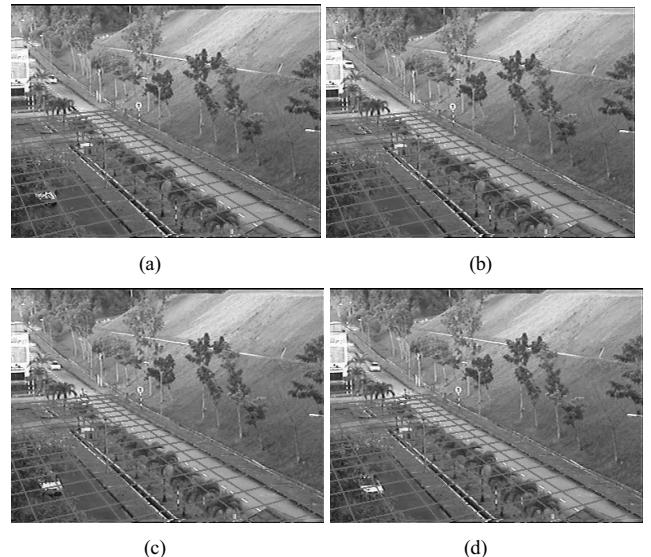


Figure 1. Experimental results of vehicle estimation of identified car and motorcycle in selected road lane

Figure 2 shows images which are presented in 3D with perspectives where the velocity calculation takes account of perspective projection. A plane mesh, indicated by green ray tracing was drawn on the images according to the perspectives of the scene. The 3D position of detected car and motorcycle, shown in figure 2(h) is very important for further work such as for vehicle classification.



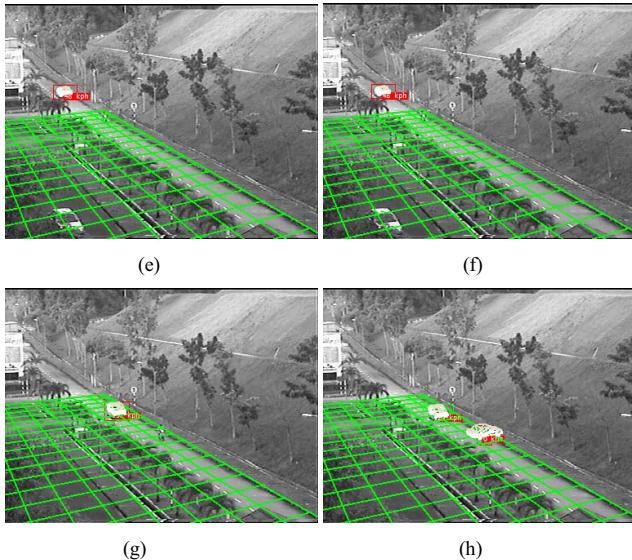


Figure 2. Vehicle speed estimation in 3D with perspective projection

V. CONCLUSION

The algorithm and system setup for vehicle speed estimation has been presented in this paper. The algorithm used for this system is being determined by applying several processing steps which involve moving object detection, centroid extraction and object tracking. The actual experiments have been done in real time by testing the proposed algorithm using desktop PC.

The experimental results demonstrate that the algorithm is capable of computing the multiple objects' speed simultaneously in real-world scene. The 3D perspective projection of velocity is also being included in the captured images. However, this vehicle speed algorithm has not been validated with actual vehicle speed. The accuracy of estimated speed cannot be determined since ground truth is not available yet. The validation of the result will be done in the future stage. The algorithm is limited to several constraints where occlusion handling and curve have not been considered in this case as this system will be used for surveillance system in parking lot.

ACKNOWLEDGEMENT

This research is fully funded by National Science Fund of Malaysia, ScienceFund, under grant 01-01-15-SF0095 and by Ministry of Invention, Science and Technology (MOSTI). The authors would like to express their gratitude to MOSTI.

REFERENCES

- [1] C. F. Shu, A. Hampapur, M. Lu, L. Brown, J. Connell, A. Senior and Y. Tian, "IBM smart surveillance system (S3): An open and extensible framework for event based surveillance", Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance, 2005, pp. 318-323.
- [2] A. Hampapur, L. Brown, J. Connell, S. Pankanti, A. Senior and Y. Tian, "Smart surveillance : Applications, Technology and Implications", Proceedings of 2003 Joint Conference of the Fourth International on Information, Communication and Signal Processing and the Fourth Pacific Rim Conference on Multimedia, vol. 2, 15-18 Dec 2003, pp. 1133-1138.
- [3] M. Valera, and S. A. Velastin, "Intelligent distributed surveillance systems: A review", Proceedings of IEEE Conference on Vision, Image and Signal Processing, vol. 152, 8 April 2005, pp. 192-204.
- [4] A. Hampapur, L. Brown, J. Connell, A. Ekin, N. Haas, M. Lu, H. Merkl and S. Pankanti, "Smart video surveillance: exploring the concept of multiscale spatiotemporal tracking", IEEE Signal Processing Magazine, vol. 22, no. 2, pp. 38-51, March 2005.
- [5] M. Litzenberger, A. N. Belbachir, P. Schon and C. Posch, "Embedded smart camera for high speed vision", Proceedings of First ACM/IEEE International Conference on Distributed Smart Camera, 25-28 Sept 2007, pp. 81-89.
- [6] D. Bauer, A. N. Belbachir and N. Donath, et al., "Embedded Vehicle Speed Estimation System Using an Asynchronous Temporal Contrast Vision Sensor," EURASIP Journal on Embedded Systems, vol. 2007, Article ID 82174, 12 pages, 2007.
- [7] J. C. Tai, S. T. Tseng, C. P. Lin and K. T. Song , "Real-time image tracking for automatic traffic monitoring and enforcement applications," Image Vis. Comput., vol. 22, pp. 485, Jun 2004.
- [8] M. Bramberger, B. Rinner and H. Schwabach, "An embedded smart camera on a scalable heterogeneous multi-DSP system", Proceedings of the European DSP Education and Research Symposium (EDERS2004), Nov 2004.
- [9] D. A. Bue, D. Comaniciu, V. Ramesh and C. Regazzoni, "Smart cameras with real-time video object generation", Proceedings of International Conference on Image Processing, vol.3, 24-28 June 2002, pp. 429-432.
- [10] I. Diaz, M. Heijligers, R. Kleihorst and A. Danilin, "An embedded low power high efficient object tracker for surveillance systems", Proceedings of First ACM/IEEE International Conference on Distributed Smart Camera, 25-29 Sept 2007, pp. 372-378.
- [11] H. A. Rahim, U. U. Sheikh, R. B. Ahmad, A. S. M. Zain and W. N. S. F. W. Ariffin, "Vehicle speed detection using frame differencing for smart surveillance system", in press.
- [12] U. U. Sheikh and S. A. R. Abu-Bakar, "Three-dimensional pose estimation from two-dimensional monocular camera images for vehicle classification", Proceedings of 6th WSEAS International Conference on Circuits, Systems, Electronics, Control and Signal Processing, 29-31 Dec 2007, pp. 356-361.