Signal and Image Processing Letters
Vol. 3., No. 2, 2021, pp. 11-16
ISSN 2714-6677
https://simple.ascee.org/

Vehicle speed estimation using optical flow on traffic video under day and night lighting condition

Ahmad Bramdimas Anggisa [a,1], Adhi Prahara [a,2,*]

[a] Department of Informatics, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[1] ahmad1400018237@webmail.uad.ac.id; [2] adhi.prahara@tif.uad.ac.id
* corresponding author

ARTICLE INFO

Keywords
Vehicle speed
Optical flow
Traffic video
Computer vision
Speed estimation

ABSTRACT

Traffic violation and congestion can happen at day or night. As a preventive measure, CCTV is installed at strategic locations on the road to monitor the traffic violation and congestion. Usually, some speed sensors also installed to measure the speed of vehicles then through a system, it will inform the operator about speedy vehicles or predict a congestion. However, it is not effective because it needs a lot of sensors to be able to monitor the vehicle speed in many locations especially in the highway and before the intersection all the time. This problem leads to the development of intelligent traffic monitoring system using computer vision technology. In this research, an optical flow-based vehicle speed estimation method is proposed. The method takes a CCTV video as an input, defines the road region of interest/ROI, performs orthographic projection transformation to find the ratio of distance, uses optical flow Farneback to track the vehicle movements, and estimates the vehicle's average speed on the road. The method is tested using CCTV video under day and night lighting condition. From the experiment, the proposed method achieves 9.8% of average RMSE.

1. Introduction

Traffic violation and congestion have been increasing and become a concerning problem. Many drivers have violated the traffic speed limit. This behavior may lead to traffic accidents. Based on the publication data from the Directorate General of Land Transportation, in 2018 there were 191,498 cases of traffic accidents. From these data, the highest number of accident involves motorcycles. The accident started with a traffic violation such as violating the vehicle speed limit and human error [1]. Moreover, this problem can happen during the day or night so 24-hours monitoring is important to perform early detection and action. As a preventive measure, the government installed CCTV (Closed Circuit Television) at the strategic locations on the road. Usually, some speed sensors also installed to measure the speed of vehicles. Through a system, it will inform the operator about speedy vehicles or predict a congestion. However, it is not effective because it needs a lot of sensors to be able to monitor the vehicle speed in many locations especially in the highway and before the intersection all the time. This problem leads to the development of intelligent traffic monitoring system using computer vision technology.

Computer vision technology enables us to automatically detect and tracking objects on video. However, many challenges present in this method such as the same vehicle that observed on different CCTV viewpoints produces different vehicle orientations [2], changing environment e.g. city to forest environment and lighting condition especially at day and night [3]. To handle many possibilities of vehicle poses produced by the CCTV viewpoints, Prahara *et al.* proposed vehicle tracking based on road direction [2]. The method groups the vehicle orientation into four directions, estimates the road area and direction, assigns the suitable vehicle detector and tracks the vehicle based on the road direction. From the result, this approach able to detect and track vehicles efficiently based on their orientations. To handle poor lighting condition at night situation, Jazayeri *et al.* use the intensity from the head and tail lights of moving vehicles and from the streets lamps to track the vehicles [3]. The experiment proves that this approach is effective to track vehicle at night.

Vehicle tracking [2]–[10] and vehicle speed estimation [7], [9]–[13] methods have been proposed by researchers and achieves good result. In this research, the object is vehicle on the road and the parameter that will be estimated is vehicle speed. Vehicle speed can be estimated in many approaches such as proposed by Rad *et al.* [11] who compute the vehicle speed after applying background removal to extract the foreground which is the moving vehicles. The speed is computed using the distance measured from the centroid position of the vehicle bounding box in each consecutive frame divided by the frame per second (FPS) of the video then converted into real world speed. Nurhadiyatna *et al.* [12] estimate the speed of vehicle by measuring the Euclidean distance after classifying the vehicle using Principal Component Analysis (PCA) then track the vehicle using Kalman filter. Tourani *et al.* [13] compute the vehicle speed by measuring the displacement of tracked moving vehicles in sequential frames. The moving blobs are tracked using blob tracking algorithm on the detected vehicle produced from Mixture-of-Gaussian background subtraction method.

Modern computer vision uses deep learning model in the method to detect, track, and estimate the speed of vehicle [9], [10]. Hua *et al.* [9] adopted detect-then-track approach. The method uses transfer learning from the existing deep learning model to detect vehicle on the road, tracking the detected vehicle using optical flow then computes the speed of vehicle. The speed is estimated by computing the local movements of corner points from the tracked vehicle in a small tile window to handle the speed normalization in a perspective view of the video. Tang *et al.* [10] also uses deep learning model which is YOLOv2 to detect the vehicle after performing camera calibration to measure the perspective view of video. The detected vehicles are tracked to measure the trajectory that is necessary to estimate the speed of vehicle.

In this research, the average speed of vehicle on the road is estimated using optical flow Farneback method on a pre-determined ROI region. Based on the previous related works, many research suggested detect-then-track approach and it is important to perform calibration of the camera perspective view in order to compute real-world speed. Therefore, the proposed speed estimation method is computed after converting the camera viewpoint into orthographic projection space then track the moving objects on the road. The rest of the paper is organized as follow: section 2 presents the proposed method, section 3 presents the results and discussion and section 4 presents the conclusion of this research.

2. Vehicle Speed Estimation Method

The general step to compute the average speed of vehicles on the road is shown in Fig. 1. The proposed method takes input from traffic video then manually draw a ROI region which is the area on the road that will be the focus to implement the method.

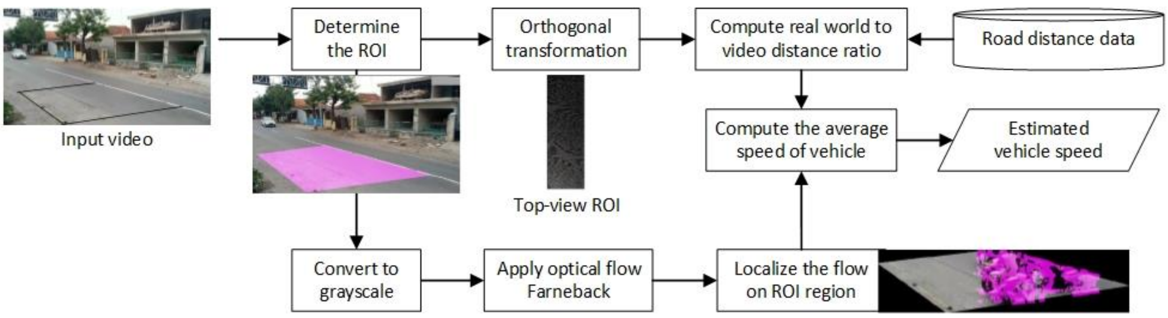


Fig. 1. The general procedure of the proposed method.

The ROI also used to compute the orthographic projection transformation to measure the ratio of real world distance with the distance in the video. The result of this transformation is top-view road area which represent the right proportion of distance. Region inside the ROI is converted into grayscale to apply the optical flow Farneback, frame by frame. The potential regions of moving objects are tracked by the optical flow then filtered to remove the noisy region using thresholding method. The flow is used to estimate the average speed of vehicles in the region. The estimated speed is converted into real world speed using the previously calculated real world distance ratio. The following subsection will explain the procedure in details.

2.1. Dataset and camera calibration

The traffic video data is taken during the day and night. To limit the detection area, a Region of Interest (ROI) is manually drawn on the road area. Region of Interest is the filtered part of the image to limit the area of the frame used for object detection. ROI can also be used as the starting and ending point of the speed calculations when tracked objects enter and exit the detection region. Inside the ROI, a real world distance is measured using a line mark to be used in the real world-video distance ratio computation. Because the position of CCTV camera in the real world is in perspective projection (when parallel lines converged toward vanishing point), an orthographic projection transformation should be performed to map the ROI. In this research, a ROI is drawn on the road area in four control points. A transformation matrix is extracted from the four control points using (1). The transformation matrix is used to transform each points inside the ROI region in the perspective view into corresponding orthographic projection top-view.

$$u = (Ax + By + C)/(Gx + Hy + I)$$

$$v = (Dx + Ey + F)/(Gx + Hy + I)$$

Where u, v is the moving point and x, y is the fixed point, A-H is the transformation coefficients

When I = 1, the equation becomes: (1)

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -ux & -uy \\ 0 & 0 & 0 & x & y & 1 & -vx & -vy \end{bmatrix} [A \quad B \quad C \quad D \quad E \quad F \quad G \quad H]'$$

Or we can write the above equation as $U = X\,T$, hence we get:

$$T = X^{-1}U \text{where } T = [A \quad B \quad C \quad D \quad E \quad F \quad G \quad H]'$$

After the transformation, the motion can be measured accurately and compared with the real-world distance. The actual vehicle speed in the video is set as the ground-truth data. The average speed of vehicle is computed using (2) where v is the estimated average speed of vehicle, s is the average distance in orthographic projection space, r is the ratio of real world-video distance, and t is the duration which computed based on FPS (frame per second) of the video.

$$v = s * r/t \tag{2}$$

2.2. Farneback optical flow

Optical flow method computes the optical flow field in the frame and performs grouping according to the characteristics of the optical flow of the image. The optical flow method generates complete motion information from an object and it is useful for detecting moving objects from the background. The optical flow method tracks point by point of each consecutive frame. There are two types of optical flow namely sparse and dense optical flow. Sparse optical flow computes motion vector for the specific set of features of objects while dense optical flow computes the motion vector for every pixel in the image. There are many optical flow methods such as Lucas-Kanade, Horn-Schunck, and Farneback [14], [15]. In this research, dense optical flow which is Farneback method will be used to compute the motion vector from the video. Farneback method approximates some neighbors of each pixel with a polynomial. By observing the differences in the polynomials that caused by moving objects displacements, the motion of each pixel can be computed.

Farneback method generates an image pyramid, where each level has a lower resolution by down-sampling the previous level resolution. The method tracks the points at various resolution levels starting from the lowest resolution level and continues until convergence. The result is propagated to the next level and refined until it reaches the original resolution. The image pyramid is important to handle large pixel motions that greater than the size of the observed neighborhood. The size of the neighborhood pixels will affect the robustness of the method against image noise and fast motion. The flow vectors of each pixel from the result of optical flow Farneback implementation will be used to estimate the vehicle speed. The orientation and magnitude of the vectors are computed and filtered using a fixed threshold to remove small movements caused by noise.

2.3. Performance evaluation

The performance of the method is evaluated using Root Mean Square Error (RMSE). RMSE compare the estimated speed with the actual speed on each frame and average the result for a specific time duration. Low RMSE means that the predicted data is closer to the actual data. RMSE can be computed using (3) where N is the number of data, y_i is the actual data and $\hat{y}_i$ is the predicted data.

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(y_i - \hat{y}_i)^2}{N}} \tag{3}$$

3. Results and Discussion

The proposed vehicle speed estimation method was developed using Matlab 2017b and runs on a laptop with an Intel Core i5 Processor and 4GB RAM. The method is evaluated using two videos taken at day and two videos taken at night. The duration depends on the time when vehicle enter the ROI region. The test data scenario is shown in Fig. 2 where (a) shows the ROI and marker position on the road and (b) shows the test vehicle that pass through the ROI. The details specification of the test data (from left to right in Fig. 1) is shown in Table 1.

Video1 Video2 Video3 Video4

(a) ROI and marker position on the road

(b) The test vehicle that pass through the ROI

Fig. 2. ROI testing on the road

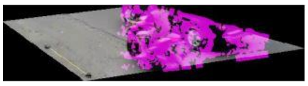
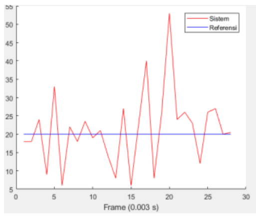
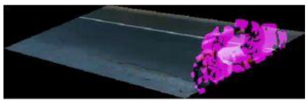
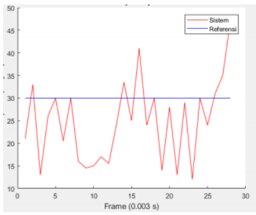
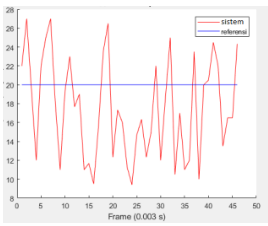
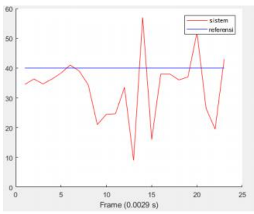
Table 1. Test data specification.

Video	Condition (light)	Camera height from the ground (m)	Vehicle actual speed (km/h)	Real world line marker (m)	Video line marker (pixel)	Duration (frames)
Video1	Day (1800-1900 lux)	2.5	20	2	150	20
Video2	Day (1800 – 1900 lux)	2.5	30	2	150	28
Video3	Night (0-5 lux)	3	20	2	200	47
Video4	Night (0-5 lux)	3	40	2	173	23

Based on Table 1, the difference duration mainly affected by the actual speed of the vehicle. For example, in Video3 with vehicle speed of 20 km/h, the test conducted on 47 frames because the vehicle is traveling at low speed so that many frames are captured on the ROI. However, in Video4 with vehicle speed of 40 km/h, the test conducted in only 23 frames because the vehicle is traveling faster so that only a few frames are captured on the ROI. In the daytime videos (Video1 and Video2) with the speed of 20 km/h and 30 km/h, the test conducted on 20 and 28 frames because the camera location is not high enough so that the ROI area is smaller.

The vehicle speed estimation result is shown in Table 2. Based on Table 2, the estimated speed of Video1 and Video3 is close to the actual speed but on Video2 and Video4 the estimated speed is far from close to the actual speed. However, the average RMSE is 9.8% which means that the method is quite good in estimating vehicle speed. As shown in the plot of the estimated speed frame by frame, the estimated speed is not constant because it is affected by the distance measured from the tracked moving vehicles. Sometimes in a frame there is only a small movement of tracked vehicles hence the speed is also small or even zero. This result affects the RMSE score greatly. However, if the speed estimation is performed only when the vehicles enter and exit the ROI then the result is better. This fact is shown in the plot that the starting point and the end point of the estimated speed (red line) is close to the actual speed (blue line).

Table 2. The result of vehicle speed estimation.

Video	**Moving vehicles on ROI (magenta marker is the tracked moving objects)**	**Plot of the estimated speed in each frame (Red line: estimated speed, Blue line: actual speed, x-axis is the estimated speed, y-axis is the frame)**	**Actual speed (km/h)**	**Average estimated speed (km/h)**	**RMSE (%)**
Video1			20	21.2	10.8
Video2			30	24.7	10.0
Video3			20	17.4	5.8
Video4			40	33.4	12.5
		Average RMSE			9.8

4. Conclusion

The vehicle speed estimation on video has been performed and evaluated. Based on the test results, it can be concluded that from the 4 different videos, the RMSE score from the video at night with an actual speed of 20 km/hour obtains an RMSE of 5.8% and an average estimated speed of 17.4 km/hour, nighttime video with an actual speed of 40 km/hour obtains an RMSE of 12.5% and an average estimated speed of 33.4 km/hour, daytime video with an actual speed of 20 km/hour evaluated with RMSE of 10.8% and the average estimated speed of 21.2 km/hour, and lastly the daytime video with an actual speed of 30 km/hour obtains by an RMSE of 10% and average estimated speed of 24.7 km/hour. The average RMSE from the test data is 9.8% which means the proposed method quite good in estimating vehicle speed on the road. For the future works, the proposed method will be improved specially to reduce the error and produce stable estimated speed each frame.

References

[1] D. G. of L. Transportation, "PDDA 2018," *Directorate General of Land Transportation*, 2018. http://hubdat.dephub.go.id/data-a-informasi/pdda/2843-tahun-2018 (accessed Jan. 01, 2019).

[2] A. Prahara, A. Azhari, and Murinto, "Vehicle pose estimation for vehicle detection and tracking based on road direction," *Int. J. Adv. Intell. Informatics*, vol. 3, no. 1, pp. 35–46, Mar. 2017, doi: 10.26555/ijain.v3i1.88.

[3] A. Jazayeri, H. Cai, J. Y. Zheng, and M. Tuceryan, "Vehicle Detection and Tracking in Car Video Based on Motion Model," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 583–595, Jun. 2011, doi: 10.1109/TITS.2011.2113340.

[4] Bin Tian, Ye Li, Bo Li, and Ding Wen, "Rear-View Vehicle Detection and Tracking by Combining Multiple Parts for Complex Urban Surveillance," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 2, pp. 597–606, Apr. 2014, doi: 10.1109/TITS.2013.2283302.

[5] S. M. Elkerdawi, R. Sayed, and M. ElHelw, "Real-Time Vehicle Detection and Tracking Using Haar-Like Features and Compressive Tracking," Springer International Publishing, 2014, pp. 381–390.

[6] B. Coifman, D. Beymer, P. McLauchlan, and J. Malik, "A real-time computer vision system for vehicle tracking and traffic surveillance," *Transp. Res. Part C Emerg. Technol.*, vol. 6, no. 4, pp. 271–288, 1998.

[7] C. C. Yik, "Vehicle tracking and speed estimation system," *Fac. Comput. Syst. Softw. Eng. Univ. Malaysia Pahang*, 2012.

[8] K. Kovačić, E. Ivanjko, and H. Gold, "Real time vehicle detection and tracking on multiple lanes," 2014.

[9] S. Hua, M. Kapoor, and D. C. Anastasiu, "Vehicle Tracking and Speed Estimation From Traffic Videos," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2018.

[10] Z. Tang, G. Wang, H. Xiao, A. Zheng, and J.-N. Hwang, "Single-Camera and Inter-Camera Vehicle Tracking and 3D Speed Estimation Based on Fusion of Visual and Semantic Features," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2018.

[11] A. G. Rad, A. Dehghani, and M. R. Karim, "Vehicle speed detection in video image sequences using CVS method," *Int. J. Phys. Sci.*, vol. 5, no. 17, pp. 2555–2563, 2010.

[12] A. Nurhadiyatna, B. Hardjono, A. Wibisono, W. Jatmiko, and P. Mursanto, "ITS information source: Vehicle speed measurement using camera as sensor," in *2012 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2012, pp. 179–184.

[13] A. Tourani, A. Shahbahrami, A. Akoushideh, S. Khazaee, and C. Y. Suen, "Motion-based Vehicle Speed Measurement for Intelligent Transportation Systems.," *Int. J. Image, Graph. Signal Process.*, vol. 11, no. 4, 2019.

[14] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *Int. J. Comput. Vis.*, vol. 12, no. 1, pp. 43–77, 1994.

[15] G. Farnebäck, "Two-frame motion estimation based on polynomial expansion," in *Scandinavian conference on Image analysis*, 2003, pp. 363–370.