2016 2nd International Conference on Science in Information Technology (ICSITech)

Vehicle Detection and Tracking Based on Corner and Lines Adjacent Detection Features

M.D. Enjat Munajat[1], Dwi H. Widyantoro[2], Rinaldi Munir[3]
School of Electrical Engineering and Informatics Institute of Technology Bandung,
[1]mdenjatm@unpad.ac.id, [2]dwi@stei.itb.ac.id, [3]rinaldi@informatika.org

***Abstract*—This paper discusses a new method in detecting moving objects, which is differ from most of the methods used such as Gaussian Mixture Model, and Haar-Like approach. The focus is on utilizing corner detection and line adjacent detection features through thresholding process creating black and white images to detect the corner of each object in each frame. The process divides a frame length into 4 parts, whereas the first part acted as initiation process of moving object recognition while the rest of the frame functioning as vehicle tracking, speed measurement, and number of vehicles calculation. The initiation process started by identifying corner spots of the moving objects that must be recognized as a single object. The lines surpassing through two points are later identified to determine whether those spots have dark color (0) or light color (1). The moving objects is represented by light color (1) and the walking objects is represented by dark color (0). A group of corner spots, identified and connected by two-point-line equation to be recognized as one unified object by using corner and line adjacent method. The identified vehicle objects can be more easily tracked and identified by the average speed in order to obtain the number of passing vehicles. The research result shows that in the initiation phase, the corner and line adjacent features able to detect moving object and distinguish it with different objects. Furthermore, in tracking phase, system is able to track the vehicle position, measuring the speed and number of vehicles. The system is proven to be able to recognize the moving objects quickly and accurately resulting in the more feasible process of speed measurement and tracking.**

Keywords*— *vehicle detection; motion; corner detection; corner and lines adjacent; vehicle tracking; vehicle counting

I. Introduction

The detection of vehicles is an issue that often arises in the field of object detection research. The vehicle detection is a necessity, mainly to detect the density of vehicles in one area, enabling the traffic control to minimize congestion. Nowadays, a number of methods of moving vehicles detection have been developed either manually or automatically. The most widely used method for automatic system is the background subtraction which spots a descriptive pattern of moving objects in order to be separated from the background to find the desired object. The detection was made by comparing the information pixel-to-pixel in every movement in the frame. Here are some methods used in applying traffic control based on background images subtraction approach: Color median [1], Frame differencing [2] [3] [4], Mixture of Gaussian (MOG) [5] [6], wavelet differencing [7], the kernel-density calculation [8] [9], and sigma-delta filter [10]; Vehicle linearity [14]. [11] application of frame differencing method with thresholding by comparing the pixel intensity of vehicle on the highway. [6] application of MOG model with shadow elimination method based on the color reflection and gradient features. Recently, a sigma-delta filter model has been proposed [10] to continuely recognize the idling vehicle as a foreground object and to be able to reduce the computational load. However, the substraction background model can not resolve the common problems such as vehicle images being blocked , the presence of shadows, variations of light, camera vibration, and climate change [11]. Some researchers propose some ways to eliminate the shadows in the video by applying an additional system running on the background substraction process [12]. However, background subtraction process also has its disadvantages as it does not have the ability to accurately detect objects on its center, which is very crucial at the start of detection process.

Based on the previous research mentioned above, this research tries to develop a method enable to adapt shadow change, obstruction, and lighting, with high accuracy by using a static monocular camera. Therefore, this research proposes a method to detect, track, and measure the speed of a moving object based on the corner and line adjacent features. This method requires a lower cost, but really effective and accurate in recognizing moving objects in a fast and real-time condition.

II. Related Works

A. Vehicle Detecting

Research on vehicle detection is already established, some use a multi-camera method as [13] which is expected to be able to capture objects unable to be captured by the first camera would be able to be captured by the next cameras. The process of [13] even uses up to 15 cameras in the reading process of passing vehicles. The method applied in a scenario of skipping the vehicle 1, 2, 3, 5, 15, proven to be able to reach 95.2% reading result of vehicle validation but dropped to 60% if the proposed method not applied. Similarly [2], defines a new feature called “vehicle linearity” to classify the type of vehicle. This new feature is very useful for distinguishing “van truck” from the “truck” even without using 3-D information, using Gaussian Mixture Model (GMM) which is widely used. The process carried out on GMM is to identify the background then examine the movement of the foreground in each frame. One of which is [14], which has similar work in the field of vehicle detection in 1997. Here Beymer, et al [14] utilizes a tracker on the network of Texas Instruments C40 DSPs (Digital Signal Processing). Indeed, in this case, it seems that [14] uses build-in Chips in addition to algorithms for the vehicle detection process. In his experiment, Beymer executed the experiment using 44 hours traffic video recording. At a glance, it looks like the method used by [14] refers to the Gaussian Mixture Model (GMM) method and bears a resemblance to the method used by [2]. It seems that the research of Beymer, et al [14] is significant to trigger some studies similar to [2] and [15].

B. Vehicle Tracking

In the field of vehicle tracking, some researchers use multiple sensor while others use the development of set approaches such as Kalman Filter, Ada Boost, and so forth. However, the tendency of tracking methods used in the tracking process is also applied in the detection process. In principal, those methods are correlated with each other, but the difference in the development by each researcher generate a different kind of approach. Research developed by [16] is more likely intended to see the monitoring process of traffic jam and accident detection at intersections. Although [16] also carries out the process of detection in the research but it is focusing more on the tracking process. [16] adopted the Spatio Temporal - Markov Random Fields (ST-MRF) based on the Hidden Markov Model (HMM), in which [16] applied pixel scheme of 8 x 8. The experiment was performed at congested intersection with the duration of 25 minutes. Basically, the research of [16] is focusing more on object detection. By utilizing the object detection on each frame by using Spatio Temporal - Markov Random Fields (ST-MRF), it allows the system to be able to do more tracking analysis to various behavioral patterns analysis of each vehicle. For example, it would be able to detect vehicle collision, vehicle passing through, congestion, illegal U-turns, and reckless driving. Although the method proposed by [16] is quite simple but at that time (in 2000), it reaches up to 93% -96% success result.

III. PROPOSED METHODS

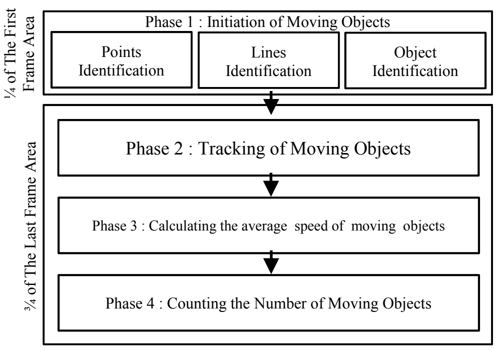


Fig. 1. Process Phase

Based on the methods described above, this reserarch tries to approach the measuring of congestion level by using surveillance camera. The camera is located above the road with the lens position is focusing down the road. From Fig. 1, it can be seen that the surveillance camera mounted on the pole with a particular height and visibility enabled it to capture the actual condition of the road. At this point its purpose is to record the actual road condition in real time. Ideally, the camera always records the condition of the street. Result of the recording is then split into multiple frames to be processed. The process separates it into four phases; Phase 1. Initiation of moving objects; Phase 2. Tracking moving objects; Phase 3. Calculating the average speed of moving objects; and phase 4. Counting the number of moving objects.

IV. EXPERIMENTAL RESULTS

A.Phase 1: Initiation Moving Objects

At this phase, the flow as more detailed described in Fig. 2, started by immediately converting the result of the video reading into grey-picture format. By utilizing thresholding feature, the image boundary is more vivid and firm in Black and White (B/W) format. We assign the thresholds value about 30 for linking threshold and 50 for upper threshold. Image of thresholding result (B/W), then processed with corner detection reading by utilizing the Kanade-Lucas-Tomasi (KLT) method, hence its corner spots will be detected as shown in Fig. 6.b. The detection result of the corner spots is later masked with the original image as shown in Fig. 6.c, so it will be distinctive among the angle of detection depicted in the original image. Afterwards, the system records the *x* and *y* position of each corner spots. On the recording result of the *xy* position, then it creates straight lines connecting each point. In order to eliminate too many unnecessary lines, a filtering process of the line length is needed here, assuming that the length of the line that exceeds the length of the vehicle can be eliminated. Therefore, will leave only lines that correspond to the length and width of the vehicle shown. To connect the points of (x_1, y_1) with (x_2, y_2),..., $(x_n y_n)$ the equation used on two points is shown in the eq. (1) - (3):

$$m = \frac{y_2 - y_1}{x_2 - x_1} \tag{1}$$

$$y - y_1 = m(x - x_1) \tag{2}$$

$$y = mx + c \tag{3}$$

With *m* as the line gradient, (x_1, y_1) is the starting point and (x_2, y_2) is the endpoint. By using Eq. (3), we obtain line equation through two points, with limit value of *x* is $(x_2 - x_1) \ll x \ll x_2$ and y value is $(y_2 - y_1) \ll y \ll y_2$. With this *x* and *y* limit values, it can be acknowledged that each point is white '1' or black '0'. With this light and dark color information, it can be concluded that if the bright color '1' dominates along the line (over 70%), it is certain that this line represents an object, otherwise, if it is less, then it can be eliminated. Such pattern can be illustrated in Fig. 3. The information of dark/light value from these points is then recorded and compiled by a connecting it with the value of its environment. Each connection of the points, formed in notation becomes images as shown in Fig. 4. In that notation, it can be seen that if points connection has linked with other points, with assumption that the line has a larger number of white points (about 70%), while the rests do not have any connection. The correlation of the conectivity pattern between points is designed in the form of a matrix, to form a model as shown in Fig. 5. Based on this pattern, it can be discovered that there are two big black blocks representing the pattern of connecting points as in Fig. 4. The black blocks represent the connection of each point, which directly represent each object. It is shown that there are two large objects on the matrix that do not directly represent both large objects mentioned in Fig. 3.

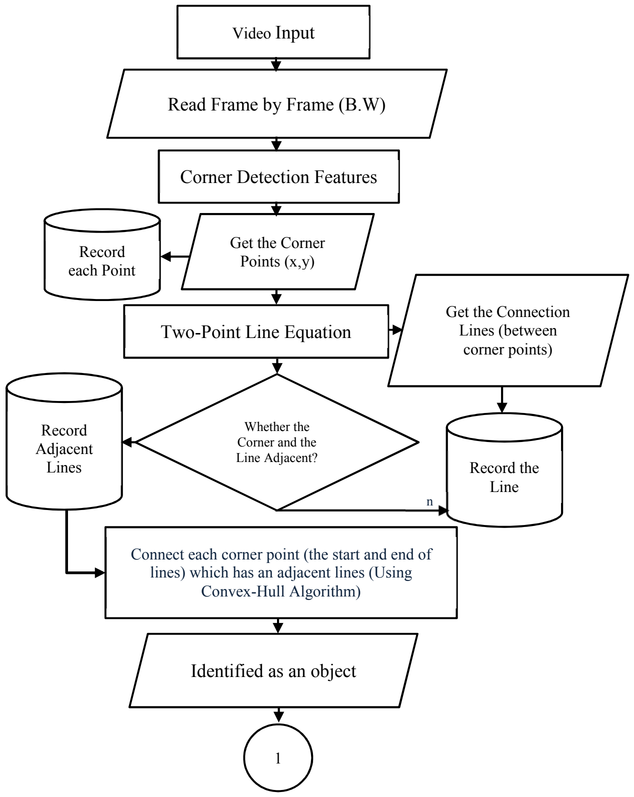


Fig. 2. The Initiation Phase

In this experiment, we use a highway video recording from one of the main roads in Bandung City, Indonesia. The video was recorded by 16 MP camera with 25 fps (frame per second) capacity. The Computer using Windows 7 platform and also C# and EmguCV ver.2.90. The video recording is analyzed frame by frame. The display is changed into greyscale and underwent thresholding process, so that only the dark and bright parts that would be able to be displayed (as shown in Fig. 6.b). Dark picture presents static condition or highway, while the bright condition (white) presents a moving object (such as a motorcycle or a car). After the thresholding process, the system detects positions of a corner in every frame, in order to obtain plenty of corners (presented in the form of '+' at each corner detected). The connection of each corner is presented by a line whose equation must be obtained between its two points. Having obtained the equation for each point, Then the system will process the line adjacent to be formed by inspecting points passed by the line, whether the content of point is 70% of light (representing object) or dark (static condition representing the highway).

In accordance to the method proposed in Fig. 2, one-third of the frame is used for the initiation phase so that in one-third part of the frame the recognition process of object can be completed. The condition indicates that the system is able to determine the connection between different objects in each frame and is able to store the same connection among the same objects so that the situation of each object can be differentiated until these points become a single unified object. Having obtained a single unified object, then the system will track the moving object in order to obtain the average motion of each object. Fig. 6 indicates that the proposed method is able to separate two adjacent objects, because the adjacent algorithm is clear enough to recognize the unity of the same object. The similar pattern applied in order to examine the consistency of the point and line movement, produced during the notation process of corner point. Therefore, when two coincided vehicles occured, the distance of line delta which tends to be unconsistent will show that there are two different objects. Thus, it makes the system easier to group consistent corner and line. Each grouping signs each different object, so by using convex-hull algorithm the border between two lines is clearer (Fig. 6). The method is strong enough to differentiate two nearby or coincide objects, which usually become issues on object detection research.

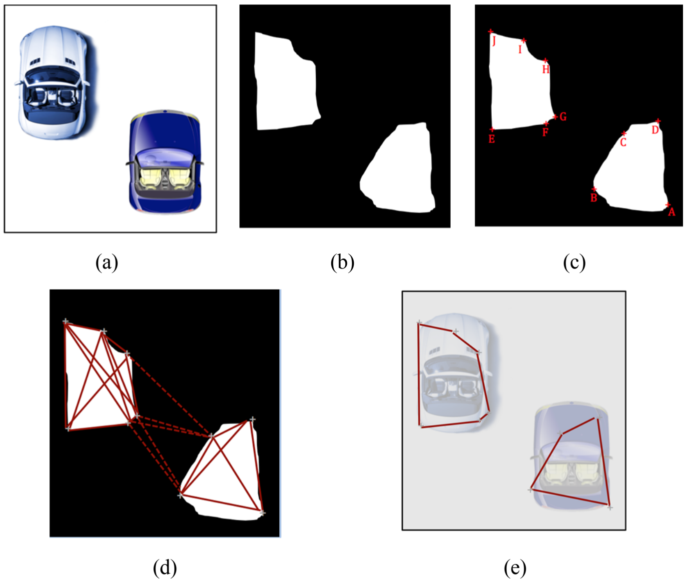


Fig. 3. Model of vertex detection and the connection among the vertices: (a) position of cars on the highway; (b) the bright patterns representing vehicle object; (c). bright patterns whose vertices are detected; (d). The line connection connecting each vertex; (e) car limit by convex-hull algotrithm

$$A \rightarrow B, C, D$$
$$B \rightarrow A, C, D;\ \ B \nrightarrow \mathrm{F,G}$$
$$C \rightarrow A, B, D;\ \ C \nrightarrow F, G$$
$$D \rightarrow A, B, C;\ \ D \nrightarrow F, G$$
$$E \rightarrow F, G, H, I, J$$
$$F \rightarrow E, G, H, I, J;\ \ F \nrightarrow B, C$$
$$G \rightarrow E, F, H, I, J;\ \ G \nrightarrow B, C, D,$$
$$H \rightarrow E, F, G, I, J; H \nrightarrow C, D$$
$$I \rightarrow E, F, G, H, J$$
$$J \rightarrow E, F, G, H, I$$

Fig. 4. Notation of Vertex Connection

Fig. 7, indicates comparison between the proposed method and the background-subtraction method which is commonly used. It can be seen that the detection of the vehicle with the proposed method is able to separate two coincided vehicles. In this condition, there are four objects (three cars and one motorcycle) which will continue to be recognized as different objects until the process is over, so that a coincided condition, shadow and etc., will not interfere the reading process.

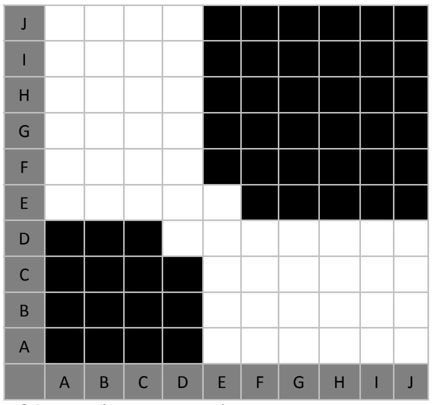


Fig. 5. Model of Connection among Points

B.Phase 2 : Tracking of Moving Objects

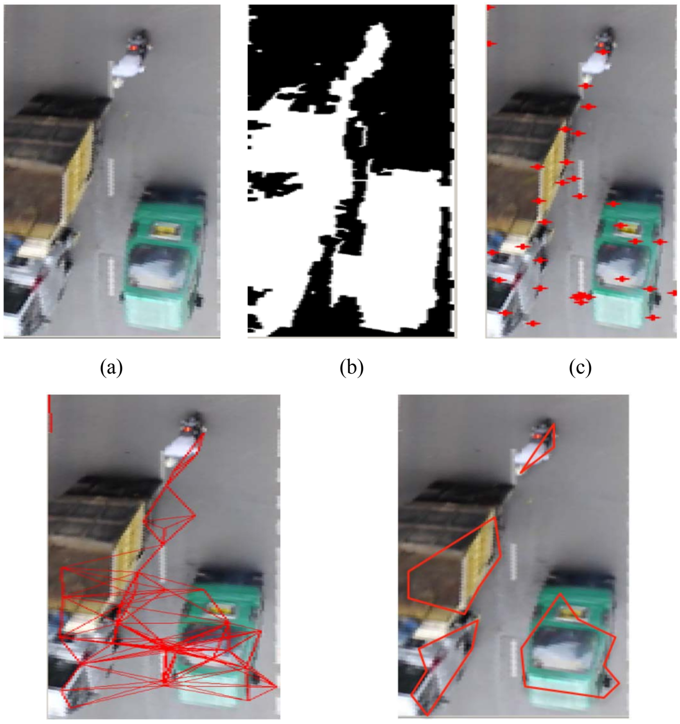
Once the object is able to be recognized as one unified object/vehicle, then the system should be able to do the tracking process towards the moving vehicle object. The proposed system for performing this tracking process is shown in Fig. 7. As an object recognized as one unified object /vehicle, then its movement in each frame will be tracked. The starting point of tracking is at the starting position in three quarter of the rest of the frame, because we know that a quarter of the initial frame is an initiation phase for object recognition. Afterwards, the movement of object in each frame will be easily tracked up to the last position in the frame. This is necessary as a source in determining of vehicle speed and the number of vehicle process. The next process is to track the movement of polygon (vehicles). We used Lucas-Kanade tracker with paramid for this purpose [17], which will track the new position for each polygon.

C.Phase 3: Calculate the Average Speed of Moving Objects

The measurement of vehicle speed is performed by (1) identifying the start and end of tracking lines, (2) detecting vehicles when it enters the starting point, (3) tracking the vehicle until it arrives at the end point, and (4) calculating the average vehicle speeds passing the end point. The start and end points are determined as part of the setup process along with the road boundary identification. A manual process during the setting of tracking lines is measuring the distance between the two lines in real condition. One of the main issues occurred is tracking vehicles passing the tracking path. The main task here is to count the number frames needed for a vehicle from starting point to reach the end of tracking line. Once the vehicle arrives at the end of tracking line, we can calculate the vehicle speed by using information on the distance between start and endlines (as identified during the setup process) as well as the frame rate. In particular, the vehicle speed is indicated as follows :

$$Speed = \frac{Frame_{rate}\ x\ Distance}{\#Frame} \tag{4}$$

Where *Distance* is the distance between start and end of tracking lines while *#Frame* is the number of frames needed for a vehicle to go from starting line towards the end of the line. Any existing algorithms on vehicle tracking can be used for this purpose.



(a) (b) (c)

(d) (e)

Fig. 6. Display of the System Using Proposed Method: (a) Original Image Condition; (B) Having Thresholding Process; (C) Corner Detection; (D) Connection of Corner Point in Line Form; (E) Limit of Vehicle with Convex-Hull Algorithm

(a) (b) (c)

Fig. 7. Comparison : (a) Tresholding Object Result; (b) Background-Subtraction Method; (c) Proposed Methods

To prove the accuracy of the calculation speed, we analyze the video frame by frame using the proposed method and compared with other methods [oka], as shown in Fig.9. Then, to validate the actual speed, we used the motorcycle as a target of measurement (see the circle marks on Fig.9), which has a constant speed at 30 Km/ H. Using the same video, then we try to analyze. Based on the results of the calculations in Table 1, shows that the speed of the proposed method is more precise than the methods proposed by other methods [18].

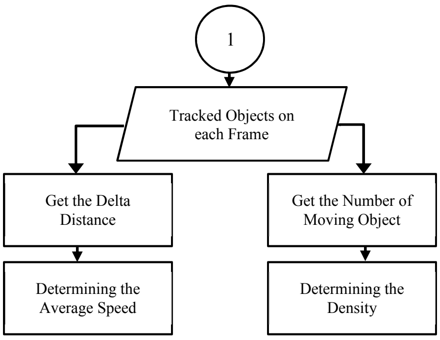


Fig. 8. Tracking Scheme for Obtaining the Vehicle Number and Average Speed

(a) (b)

Fig. 9. Original image with targeted vehicles (red circle marks): (a) Video time 00.15, with a low level of illumination and shadows of trees covered most of road ; (b) Video time 02.05, with high illumination levels and a little shade of trees covered most of road

TABLE I. SPEED VALIDATION

Second	**Method by [18] (Km/H)**	**Proposed Method (Km/H)**	**Real Speed (Km/H)**
00.15	33-35	30-31	30
02.05	34-35	30-31	30

D.Phase 4 : Count the Number of Moving Objects

The last phase is the calculation of the number of vehicles passing by, which will be easily acquired if the vehicle recognition process at the initiation phase is well obtained. In this case, the first three-quarter of the frame length is used as the start and end point of the calculation. During this process, each vehicle passing at the start point (first three-quarter of the frame length) was calculated and validated on the endpoint. If in the process of initiation (first quarter of the frame length), the system is already able to separate the identity of each moving object, the system will continue to monitor the number of vehicles that were detected at the initiation point for tracking and calculation later at the endpoint of the frame. After all objects are successfully identified, the next process is to start vehicle tracking process by identifying the velocity. As we already know, from the vehicle object reading result, the information about the distance between starting point and ending point of every tracked object, which is called distance range, can be obtained. From the collection of distance range of every object, the information of average distance can be acquired. Based on distance range, the velocity of each vehicle can be easily recorded, so the average of moving vehicle can be measured. Fig. 8 shows the result of tracking process to the measurement of distance range conducted in the 01.11-01.12 minute. In this condition, the road tends to be congested and is being dominated by car, so the average speed of the moving car is +/- 1 Km/H. Moreover, to measure the number of vehicle will be easier for the system as the identity of each vehicle will occured. Hence, the number of vehicle crossing the line has its own identity connected to the object detected at the initiation phase. This whole process relies on the initiation phase, if it fails, it can be assured that the tracking process and vehicle velocity measurement will also be unsuccesful.

V. DISCUSSION

In Indonesia, highway can also be used by two-wheeled vehicles (motorcycles) other than four-wheel (car). Therefore, when the road is crowded, then the speed of the cars moving will automatically getting slower simultaneously, but this condition can be interfered by the motorcycle running at rapid speed through the gap between cars. These incident able to obstruct the appropriate average speed. Our method is detect any polygon shapes of cars and motorcycles. Using a frame size of 640x480 pixels, the polygon shape of motor is represented by area that is less than 600 pixels, otherwise the polygon of car area above 600 pixels, so the system is able to separate the speed of motorcyles (Fig. 10) and cars (Fig. 11). As shown in Fig. 10 that information of the average distance and deviation standard will be able to minimize abnormal speed which in this case is caused by motorcycles running at rapid speed on particular conditions when congestion occured or objects moving slowly. Certainly this is undesirable because it would cause error in reading cars as dominating objects. To overcome this matter, the calculation of deviation standard can minimize the abnormal distance by detecting whether the distance of a moving object exceeds the deviation standard distance. If so, then certainly it is abnormal when a vehicle (motorcycle) capable of crossing through the congestion, which should be minimized as it will cause reading errors. The minimizing process by using the $Av.Speed_{Filtered_i}$ equation, in order to get the speed that would be able to minimize the unexpected effect of the speed (in this case the motorcycle speed when the car stops due to the congestion).

$$Std_{Dev_i} = Std.Dev(Speed_1, Speed_2, Speed_3, \dots, Speed_N) \quad (5)$$

$$Av.Speed_{Filtered_i} = \frac{(Av.Speed_i + Std_{Dev_i})}{2} \quad (6)$$

By minimizing abnormal distance (as a noise) the result will provide a relatively normal speed, so that reading process of the average distance can be more precise. Fig. 10 provide data argument, which is capable of giving a closer result to the actual condition, where the speed tends to be the same without the intervention of motorcycles. Data in the figure shows the speed at minute 02:11 at a street in the area of Dago.

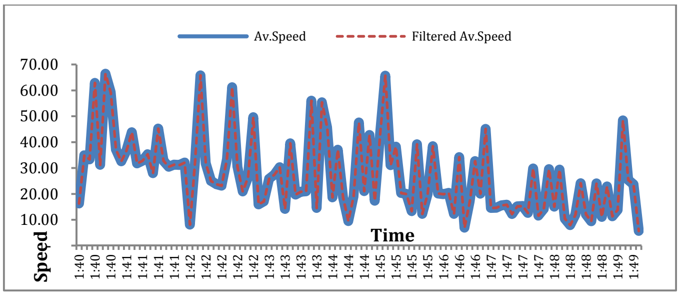


Fig. 10. Car Speed (not affected by motorcycle)

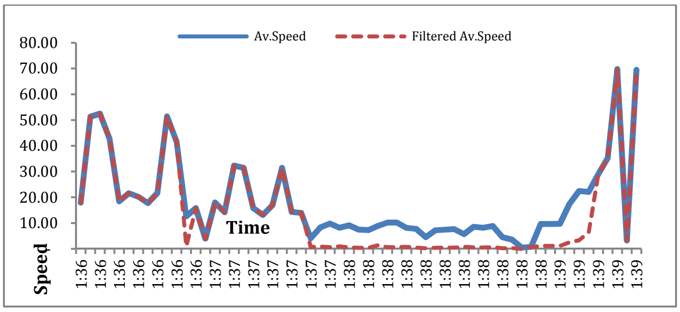


Fig. 11. Speed of Car (affected by motorcycle)

Actual value of the motor speed in Fig. 10 extremely varied, so for the convenience reading we retrieves the value of average speed only. There can be seen speed of dots read by systems with different levels of speed. The speed should be able to interpret the average of both rows. With the equation (5) and (6), the average speed able to narrow the above average motorcycle speed, hence, the reading process can be more accurate. The speed of motorcycle should be a concern because it could interfere the normal speed during a congested road condition.

VI. Conclusion

The developed system has the capability to analyse low or high traffic, and even able to cope with the coinciding vehicles and shadows. However, the developed system, requires more attention at the initiation phase as it is the most crucial phase. This phase is vulnareble to reading errors when two or more objects are adjacent with constant distance delta, but this condition is very rarely occured. If the initial phase is being well processed, it can be assured that the vehicle speed and the number of vehicles calculation will be trouble-free, which leads to vehicle speed calculation by using the average speed of vehicles passing in each frame. This could only means that it will involve more than one vehicle at a particular time. In addition, another small problem is the camera position, as the position and location of the camera must be free from disruption of wind or trees. Moreover, the direction of the lens mostly should leads towards the highway in order to avoid disruption from pedestrian or other moving object outside the highway, which could be misread and mislead and affected the reading. For future work, the system uses data records from point corner to determine the road boundaries automatically.

VII. Acknowledgement

This paper is the result of the research funded by the Decentralization of Directorate of Higher Education (DIKTI) with PUPT Scheme, 2016.

References

[1] M. Sankari and C. Meena, "Estimation of Dynamic Background and Object Detection in Noisy Visual Surveillance." International Journal, 2011

[2] J.W.Hsieh, S.H.Yu, Y.S.Chen and W.F. Hu, " Automatic Traffic Surveillance System for Vehicle Tracking and Classification," IEEE Transactions On Intelligent Transportation Systems, Vol. 7, No. 2, June 2006.

[3] N. Chaiyawatana, B. Uyyanonvara, T. Kondo, P. Dubey and Y. Hatori, "Robust object detection on video surveillance." IEEE, 149-153.

[4] W.-L. Hsu, H.-Y.M. Liao, B.-S. Jeng and K.-C. Fan, "Real-time traffic parameter extraction using entropy," IEE Proc.-Vis. Image Signal Process., Vol. 151, No. 3, June 2004

[5] E. Bas, M. Tekalp, and F.S. Salman, "Automatic vehicle counting from video for traffic flow analysis." Ieee, 392-397, 2007

[6] L.W. Tsai, Y.C. Chean, C.P. Ho, H.Z. Gu and S.Y. Lee, "Multi-Lane Detection and Road Traffic Congestion Classification for Intelligent Transportation System," Energy Procedia, 13, 3174-3182, 2011

[7] V. Crnojevic, B.Antic and D. Culibrk, "Optimal wavelet differencing method for robust motion detection," IEEE, 645-648, 2009

[8] A. Elgammal, R. Duraiswami, D. Harwood and L.S. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," Proceedings of the IEEE, 90(7), 1151-1163, 2002

[9] Y. Sheikh, M.Shah, "Bayesian modeling of dynamic scenes for object detection," Pattern Analysis and Machine Intelligence, IEEE Transactions on, 27(11), 1778-1792, 2005

[10] M.Vargas, J.M. Milla, S.L.Toral and F. Barrero, "An enhanced background estimation algorithm for vehicle detection in urban traffic scenes," Vehicular Technology, IEEE Transactions on, 59(8), 3694-3709, 2010

[11] K. Robert, "Video-based traffic monitoring at day and night vehicle features detection tracking," IEEE, 1-6, 2009

[12] L. Huang, "Real-time multi-vehicle detection and sub-feature based tracking for traffic surveillance systems," IEEE, 324-328, 2010

[13] C. Istin, D. Pescaru, A. Doboli, "Stochastic Model-Based Heuristics for Fast Field of View Loss Recovery in Urban Traffic Management Through Networks of Video Cameras," IEEE Transactions On Intelligent Transportation Systems, Vol. 12, No. 3, september 2011.

[14] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, "A real time computer vision system for measuring traffic parameters," in Proc. IEEE Conf. CVPR, 1997

[15] W. L. Hsu, H. Y. Liao, B. S. Jeng, and K. C. Fan, "Real-time traffic parameter extraction using entropy," Proc. Inst. Elect. Eng.—Vis. Image Signal Process., vol. 151, no. 3, pp. 194–202, Jun. 2004.

[16] S. Kamijo, Y.Matsushita, K.Ikeuchi, "Traffic Monitoring and Accident Detection at Intersections," IEEE Transactions on Intelligent Transportation Systems, Vol. 1, No. 2, June 2000.

[17] J.Y. Bouguet, "Pyramidal implementation of the lucas kanade feature tracker," Intel Corporation, Microprocessor Research Labs, 2000.

[18] O.Perkasa, D.H. Widyantoro, "Video-based System Development for Automatic Traffic Monitoring," IEEE International Conference on Electrical Engineering and Computer Science, 2014