



MULTIPLE VEHICLES DETECTION AND TRACKING FOR INTELLIGENT TRANSPORT SYSTEMS USING MACHINE LEARNING APPROACHES

Ngoc Dung Bui¹, Dzung Lai Manh¹, Vu Hieu Tran¹, Binh T. H. Nguyen²

¹University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.

²Ho Chi Minh City University of Technology, HCM City, Vietnam.

ARTICLE INFO

TYPE: Research Article

Received: 29/6/2019

Revised: 31/8/2019

Accepted: 16/9/2019

Published online: 15/11/2019

<https://doi.org/10.25073/tcsj.70.3.29>

* Corresponding author

Email: dzunglm@utc.edu.vn; Tel: 0964978112

Abstract. Video surveillance is emerging research field of intelligent transport systems. This paper presents some techniques which use machine learning and computer vision in vehicles detection and tracking. Firstly the machine learning approaches using Haar-like features and Ada-Boost algorithm for vehicle detection are presented. Secondly approaches to detect vehicles using the background subtraction method based on Gaussian Mixture Model and to track vehicles using optical flow and multiple Kalman filters were given. The method takes advantages of distinguish and tracking multiple vehicles individually. The experimental results demonstrate high accurately of the method.

Keywords: Vehicle detection, tracking, background subtraction, optical flow, Kalman filters.

© 2019 University of Transport and Communications

1. INTRODUCTION

Video surveillance system has become widely deployed in many aspects of life, especially in Intelligent Transportation systems (ITS). Using cameras and the image processing algorithms, the traffic flow can be measured under various environment conditions by detecting vehicles methods [1, 2]. In video surveillance system, there are three fundamental steps of image processing which are image acquiring, pre-processing and analyzing. Result of analyzing step are contents which then can be used for object

recognition. In an ITS with static cameras, motion is used as major factors for object recognition process. A robust moving object detection algorithm must handle the non-idealities of scenes such as changes in illumination, high frequency motion, changes of long-term scene, and also shadows. For the past decade, numerous algorithms were proposed to deal with the above mentioned problems [3, 4, 5]. Computer vision plays a very important role in the development of video surveillance technology. Successful applications of computer vision could be found in many fields such as video surveillance, face recognition, finger and iris recognition, and especially in transportation [6, 7, 8]. Applying computer vision techniques, the features of face, finger and iris can be extract from the image, person can be automatically identified or verified by recognition systems [9]. In video surveillance, series of computer vision algorithms will be applied on the sequence of images from camera to extract the objects or human and analyze their behaviour, characterize and decide their behaviour is normal or abnormal [10]. In transportation, computer vision can be apply to automatically monitor traffic by extract each kind of vehicle and transmit numerical data to the transport management centres [11].

Recently, a lot of camera surveillance systems was deployed [12]. There are two kinds of systems which are semi-automatic and automatic system. With the first one, the camera only capture and store images from the roads, technical staffs will then analyze contents from the image. With the second one, all the surveillances are automatically processed without any interaction from people. This automatic surveillance system can automatically detect moving vehicles, track the vehicles in their lanes and calculate the speed of the vehicles [13]. Many advanced pattern recognition technique are also applied together in the automatic system to detect, track the moving vehicles and measure traffic flow at day and night time by recognize headlight and taillight of vehicles [14].

2. MACHINE LEARNING APPLICATIONS FOR ITS

2.1. Vehicles detection based on machine learning approaches

In any traffic management and planning system, the first and most important step is collecting basic characteristics of traffic flows such as flow rate, speed and density. These characteristics are source for deployment of many intelligent transport systems' applications such as traffic signal controlling, transportation organization and management. During recent years, researches in traffic management and planning system field and to be more specific in vehicles detection and tracking field has become more urgent. Some successful research approaches of will be reviewed in the next part of this paper.

In order to obtain basic flow characteristics from traffic surveillance cameras, a process of analyzing images received from the camera must be created. Normally the process has two main stages: (1) extracting features from the images and (2) detecting and classifying vehicles based on the features. This process is illustrated in the figure 1 with four levels of complexity.

At every levels, the final steps always contains classifiers which will detect and classify vehicles.

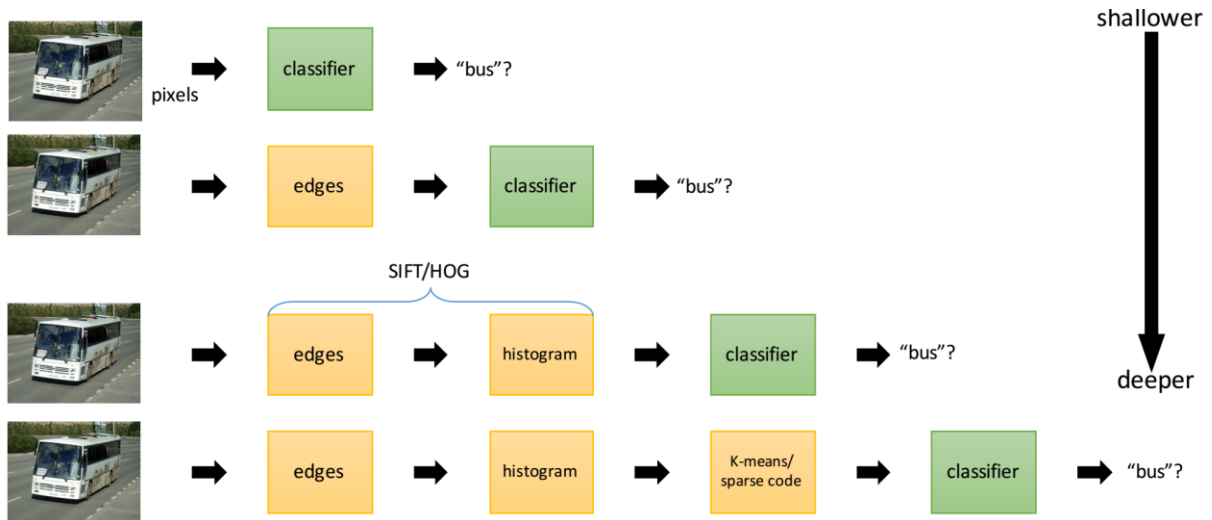


Figure 1. Complexity levels of feature extracting in traditional recognition [6].

(1) Extracting features from the images

Features on the images can be simply a collection of special pixels which have different color or intensity from the neighbor’s pixels. Features are normally pixels on the angle or edge of images’ objects. Some implementation of extracting features have been proposed such as LBP (Local Binary Pattern), HoG (Histogram of Oriented Gradient)... The characteristics of objects can have complex structures, for example an image area where pixels are interlinked follow certain principles. The example of classical principles are distribution of special pixels or the same rule in changing of intensity or light direction. Some machine learning approaches such as SVM (Support Vector Machine), AdaBoost based on Haar-like features have been proposed for these purposes.

(2) Vehicle detection and classification

The next step, the extracted features will be compared with a sample features set, then vehicles will be detected and classified. The set of sample features is built using pattern recognition methods and supervised learning techniques. The most popular supervised learning approaches are neural networks with feed or back propagations using Haar-like features combine with Ada-Boost algorithms.

A fast, popular and effective object-detection method Viola and Jones’s method which use Haar-like features [15]. The proposed Haar-like characteristics are rectangles with dark light areas interleaved as shown in figure 2.

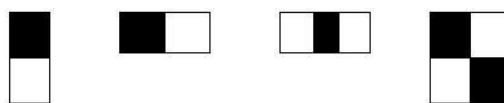


Figure 2. Basic Haar-like features.

Basic Haar-like features can be extended to recognize objects in more effectively ways. There are three groups of Basic Haar-like features: edge, line, center-surround. Extended Haar-like features are showed in figure 3.

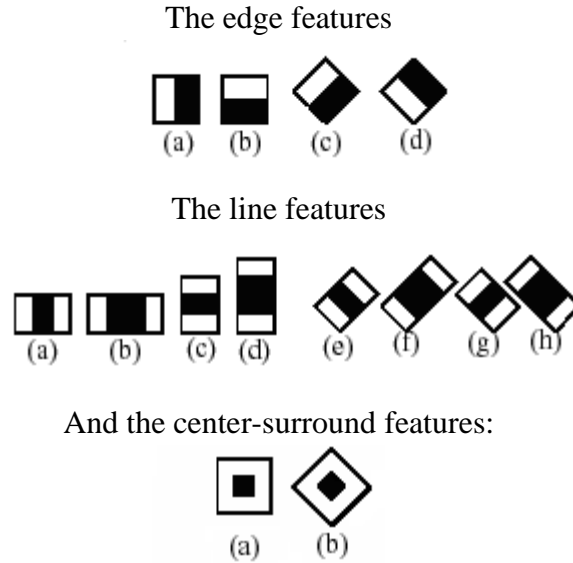


Figure 3. Extended Haar-like features.

Haar-like features' intensity values of pixels are different between pixels in bright areas and dark areas. These values can be quickly calculated based on integral image. Then these values apply AdaBoost algorithm to train strong classifier to identify objects on the image according to the [16]. To obtain a strong classifier, each calculated Haar-like characteristic is used to establish a weak classifier according to the formula number 2.

$$h_i = \begin{cases} +1 & \text{if } V_i \geq T_i \\ -1 & \text{if } V_i < T_i \end{cases} \quad (1)$$

where, V_i is the Haar-like feature value, T_i is the threshold for establishing a weak classifier, the threshold value is the Haar-like feature value of an sample image in the training set. Value $h_i = +1$ if the input image is a vehicle that needs to be detected, in other words, this classifier detected correctly input image. Conversely, $h_i = -1$ means that the input image is not a vehicle.

There is one problem, what is the suitable value of threshold T_i . In other words, which sample in the training data set should be chosen to calculate Haar-like features to set threshold for classifier? In addition, with an input image, size is often much larger than the sample image size, we must consider to utilize a lot sub-windows for the input image. With these sub-windows, only a small number contain vehicles that need to be identified. If you consider all of sub-windows are equally important then it will waste huge amount of computing resources.

Solving these two mentioned problems, the strong classifier is concluded on the basis of many weak classifiers which arranged in a multi-layer structure. Each weak classifier performs classification whether or not a vehicle that needs to be identified in the sub-window under consideration with accuracy is less than 50%. At each layer, the sub-window is removed if the classifier determines there is no vehicle. Conversely, the sub-window will be

moved to the next layer. A sub-window contains a vehicle that needs to be identified if it passes through all layers and is classified by the last layer as containing a vehicle.

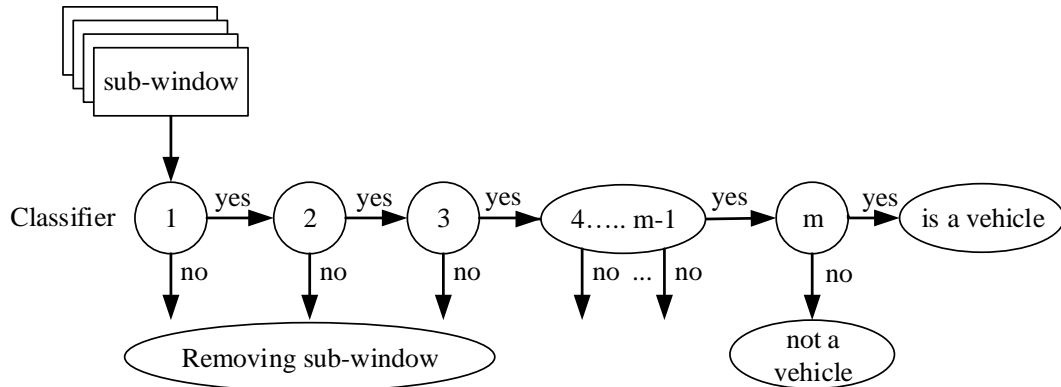


Figure 4. Concluding strong classifier based on multi weak classifiers.

The results of vehicle detection and classification use Haar-like feature according to the cascade model illustrated in the following figure 5.

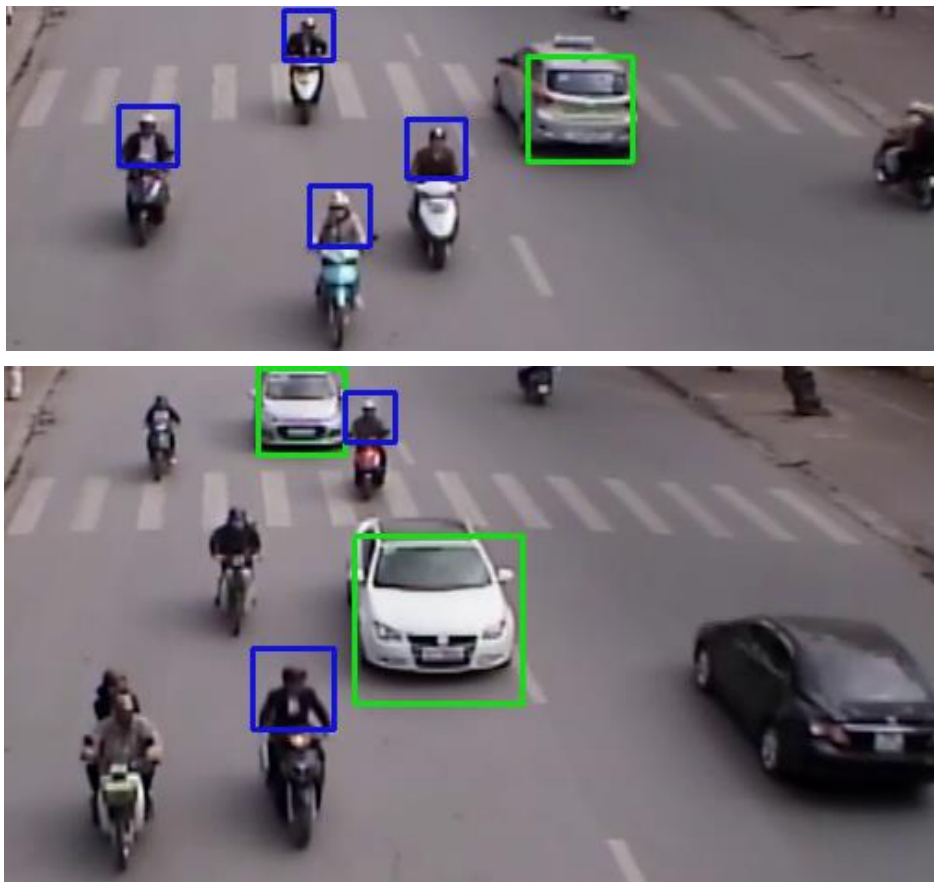


Figure 5. Discover and classify vehicles using Haar characteristics.

In experiment of medium traffic density, the accuracy of traffic vehicle detection and classification method using Haar-like characteristics using AdaBoost algorithms are quite high. However, in high-density traffic conditions, this model has low accuracy because many vehicles are partially hidden and as a result the strong classifier cannot detect these vehicles. It causes limitation of this approach in mixed traffic condition in Vietnam. The mixed traffic

condition is quite common in Vietnamese big cities which have various types of vehicles.

2.2. Vision based approach for vehicle tracking and estimation of traffic flow parameters

Machine learning approaches for detecting and counting vehicles have advantages such as can identify each type of vehicles, allow statistics and classification of vehicles. But there are some disadvantages still exist such as the computational complexity and the low accuracy in high density flow conditions. Moreover, these approaches still lack the abilities to distinguish different type of vehicles, the reason is similarities of vehicles and complex of transportation means. Therefore, this solution is often applied in areas with low traffic density, where vehicles travel clearly in specific lanes such as on highways. In Vietnamese big cities, the traffic flow is mixed and has high density. There are also various types of vehicles that do not strictly follow their lanes. The majority of vehicles are motorcycles. The solutions of motion detection algorithm is based on background subtraction and optical flow [17] have been applied to estimate the average velocity of the traffic flow and the occupancy density of vehicles on the road. Block diagram illustrated estimation process of traffic flow parameters has been shown in figure 6.

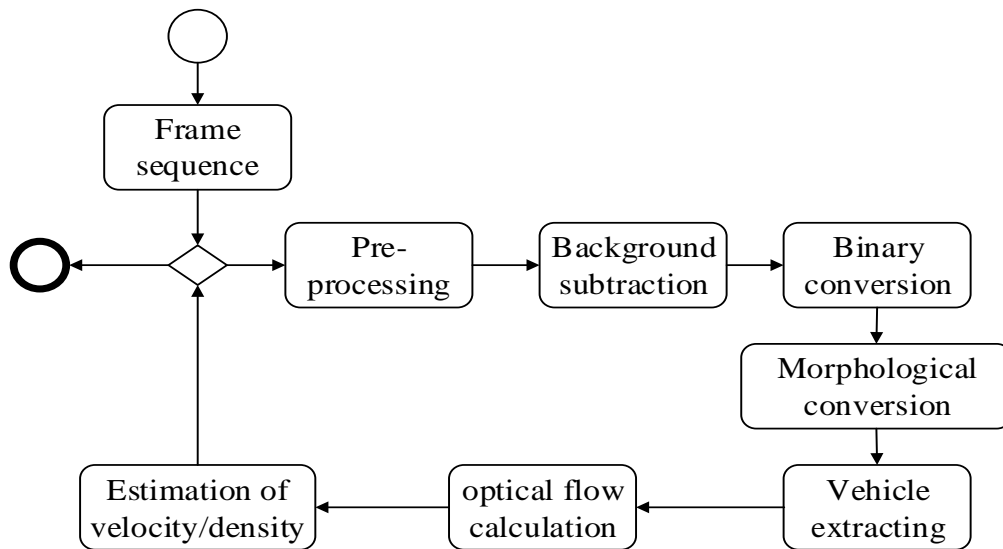


Figure 6. Traffic flow parameter estimation process.

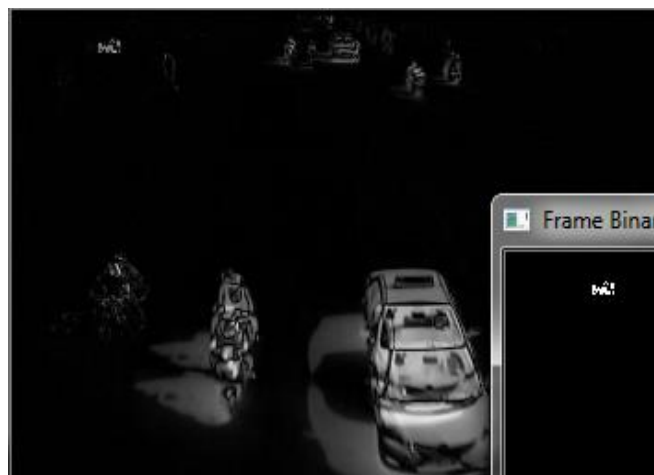


Figure 7. Result of background subtraction.

Image pre-processing: The video frames streamed from the camera, then image pre-processing transformations such as image resizing, color to gray image conversion are performed to reduce the computational complexity.

Background subtraction: The background subtraction algorithm is applied to extract traffic vehicles (foreground objects) from the image background, which are installed according to the mixture of Gaussians model. To detect the moving object from the background, the solution is to calculate the absolute deviation of the intensity of the pixels between two consecutive frames. Through the difference of intensity between two consecutive frames at the same position, it determines whether this pixel belongs to the background or the foreground object.

Binary and morphological transformations: The next step in the process of vehicle tracking from the background is to convert the resulting image to a binary image and apply some morphological transformations to integrate the discrete pixels which belong to a vehicles. These conversions improve the accuracy of vehicle tracking results.

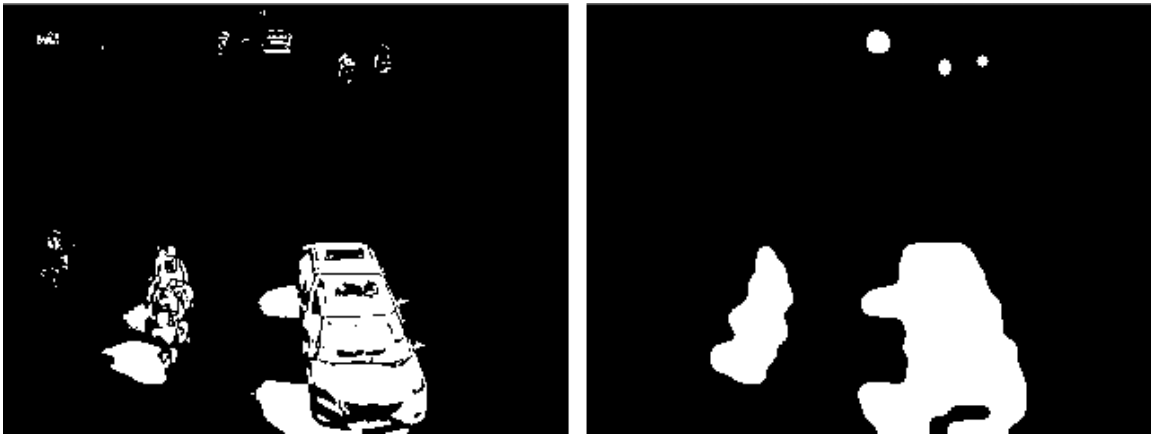


Figure 8. Binary and morphological conversion step.

Vehicle extracting: The edge detection algorithm is applied to localize a moving vehicle, and separating the foreground object from the background. The rectangular boundaries are drawn around the moving object separated from the image background.

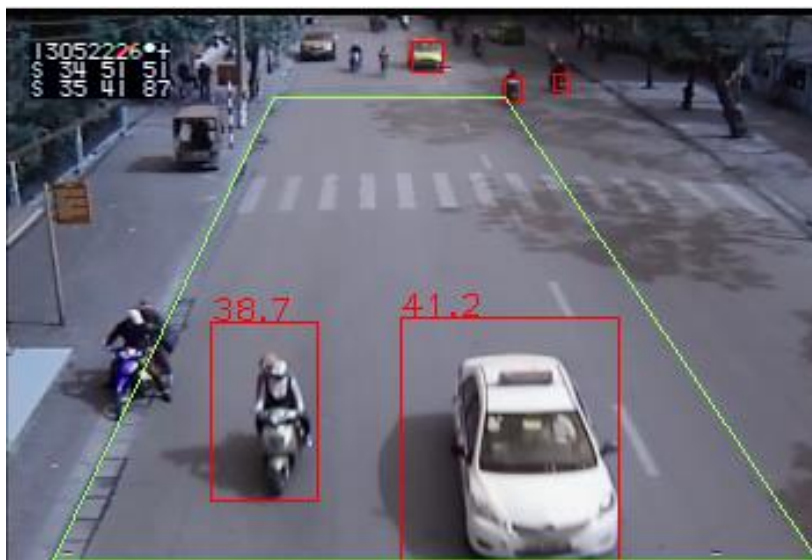


Figure 9. Extracting traffic vehicles.

Optical flow calculation: The Optical flow algorithm [17] is performed to calculate the displacement of the image-based pixels according to the frame flow, as shown in Figure 9, where the detected points are shifted from the previous frame shown. At these pixels, the displacement vector is drawn and shown on the image. These vectors are used to estimate vehicle velocity.



Figure 10. Result of optical flow calculation.

2.3. Vehicles tracking using multiple Kalman filters

Beside Optical flow method, Kalman filter can be used to predict each vehicle in current time. Normally, a Kalman filter is used to estimate the state of a linear system where the state is assumed to be distributed by a Gaussian. It is typically divided into two steps: prediction and correction. The purpose of prediction step is to estimate the state based on the state equation. Similarly the correction step uses the current observations to update the vehicle's state. In this paper, to track multiple vehicle simultaneously, multiple Kalman filters as number of vehicles is used [9]. Each Kalman filter is represented as below:

$$\begin{aligned} x_k &= Ax_{k-1} + w_k \\ z_k &= Hx_k + v_k \end{aligned} \quad (2)$$

where $x = [p_x \ p_y \ v_x \ v_y]^T$, p_x, p_y are the center position of x-axis and y-axis, respectively. v_x, v_y are the velocity of x-axis and y-axis. Matrix A represents the transition matrix, matrix H is the measurement matrix, and T is the time interval between two adjacent frames. w_k and v_k are the Gaussian noises with the error covariance Q_k and R_k . The Kalman filter is process as follow:

Update the state: $x_{k|k-1} = Ax_{k-1|k-1}$

Predict the measurement: $z_{k|k-1} = Hx_{k|k-1}$

Update the state error covariance: $P_{k|k-1} = AP_{k-1|k-1}A^T + Q_k$

To track multiple vehicles in complex transportation, matching between vehicles and measurement should be performed correctly. In this paper, we employ the data association method, which split and merge the vehicles [9]. The overall tracking method is given in figure 11.

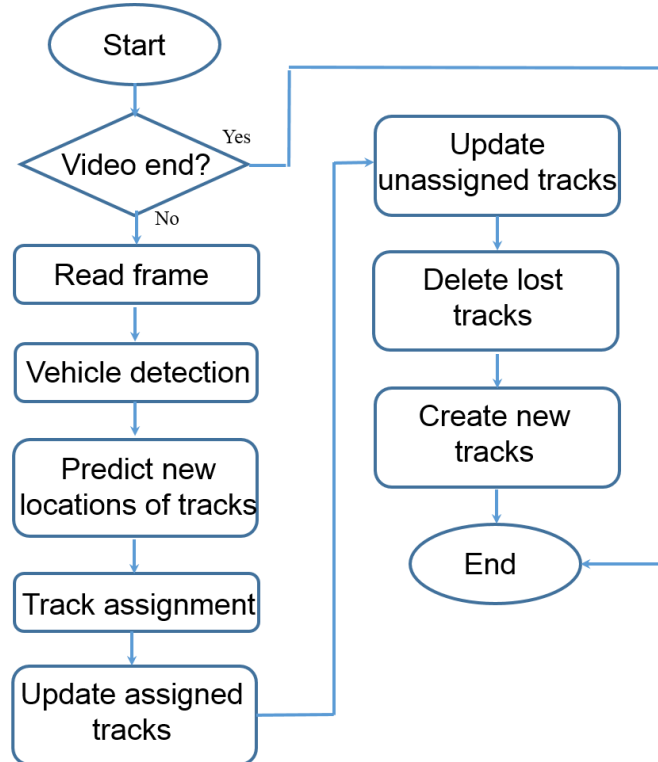


Figure 11. The flow chart of vehicles tracking method.

Figure 12 shows the results for the multiple vehicles tracking. When a car or motorbike comes to the region of the camera, it will be assigned a new tracking object and initialize tracking window for this object. The tracking results of multiple vehicles show the tracking method is able to correctly track the new vehicle in transportation camera surveillance. For the case of several vehicles run near each other, we need data association method to distinguish each vehicles.



Figure 12. Vehicles tracking using multiple Kalman filters.

4. CONCLUSION

In this paper, we presented the detection and tracking method for multiple vehicles based on various methods including background subtraction, optical flow and Kalman filter. All vehicles are detected using background subtraction. For each vehicle, the optical flow and Kalman filter was established and bounding boxes was used as features. The Kalman filter estimates the state based on the state equation and corrects using the current observations to update the vehicle's state. Results of this paper show that this method can be applied in transport management centre for traffic monitoring.

ACKNOWLEDGMENT

This research was supported by a grant from UTC project number T2019-CN-013 TD and T2019-CN-005.

REFERENCES

- [1] M.S., Shirazi, B. Morris, Traffic Flow Classification Using Traffic Cameras. In: Bebis G. et al. (eds) Advances in Visual Computing. ISVC 2018, Lecture Notes in Computer Science, 11241. Springer, Cham, 2018.
- [2] Bas, Erhan, A. Tekalp, F. Salman, Automatic Vehicle Counting from Video for Traffic Flow Analysis, Istanbul, Turkey, 392 – 397, 2007. <https://doi.org/10.1109/IVS.2007.4290146>
- [3] N. T. H. Binh, T. Q. H. Bang, N. D. Bui, Robust and Adaptive Shadow Detection in Surveillance Systems using Gaussian Processes, RIVF, 29-33, 2016
- [4] Yizhong Yang, Qiang Zhang, Pengfei Wang, Xionglou Hu, and Nengju Wu, Moving Object Detection for Dynamic Background Scenes Based on Spatiotemporal Model, Advances in Multimedia, 2017 (2017) 9 pages. <https://doi.org/10.1155/2017/5179013>
- [5] Jin Min Choi, Hyung JinChang, Yung Jun Yoo, Jin Young Choi, Robust moving object detection against fast illumination change, Computer Vision and Image Understanding, 116 (2012) 179-193. <https://doi.org/10.1016/j.cviu.2011.10.007>
- [6] Bruce E. Flinchbaugh; Thomas J. Olson, Emerging Applications of Computer Vision, 1997
- [7] Al-Osaimi; Mohammed Bennamoun; Ajmal Mian, An Expression Deformation Approach to Non-rigid 3D Face Recognition, International Journal of Computer Vision, 81 (2009) 302–316. <https://doi.org/10.1007/s11263-008-0174-0>
- [8] H. Moon, R. Chellapa, A. Rosenfeld, Performance analysis of a simple vehicle detection algorithm, 20 (2003) 1-13. [https://doi.org/10.1016/S0262-8856\(01\)00059-2](https://doi.org/10.1016/S0262-8856(01)00059-2)
- [9] NeeruRathe, A novel approach for lip Reading based on neural network, 2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT), New Delhi, India, 2016.
- [10] Song Yale, Louis-Philippe Morency, Randall Davis, Distribution-Sensitive Learning for

Imbalanced Datasets, 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Shanghai, China, 2013

[11] Chieh-Chih Wang, Cw Thorpe, Arne Suppe, LADAR-based detection and tracking of moving objects from a ground vehicle at high speeds, IEEE IV2003 Intelligent Vehicles Symposium. Proceedings (Cat. No.03TH8683), Columbus, OH, USA, 2003.

[12] Le Hung Lan et al., Application of integrated technologies to monitor and process traffic data to improve operational capacity and road safety in Vietnam, Ministry of Education and Teaching Bilateral Project, 2016.

[13] Andrew H. S. Lai, N. H. C. Yung, Lane detection by orientation and length discrimination, IEEE Trans. Systems, Man, and Cybernetics, Part B, 30 (2000) 539 – 548. <https://doi.org/10.1109/3477.865171>

[14] Yoichiro Iwasaki, Masato Misumi, Toshiyuki Nakamiya, Robust Vehicle Detection under Various Environments to Realize Road Traffic Flow Surveillance Using an Infrared Thermal Camera, The Scientific World Journal, 2015 (2015) 11 pages. <https://doi.org/10.1155/2015/947272>

[15] P. Viola, M. Jones, Rapid Object Detection using a Boosted Cascade of Simple Features. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Hawaii, USA, 511-518, 2001.

[16] Yoav Freund, Raj Iyer, Robert E. Schapire, Yoram Singer, An Efficient Boosting Algorithm for Combining Preferences, 4 (2003) 933-969.

[17] David J. Flee, Yair Weiss, Optical Flow Estimation, In Paragios; et al. Handbook of Mathematical Models in Computer Vision. Springer, 2006.