Multiple Vehicles Tracking in Intelligent Transportation System using Convolutional Neural Network and Kalman Filter

Ngoc Dung Bui
Faculty of Information Technology
University of Transport and Communications
Hanoi, Vietnam
dnbui@utc.edu.vn

Xuan Tung Hoang
Faculty of Information Technology
VNU University of Engineering and Technology
Hanoi, Vietnam
tunghx@vnu.edu.vn

***Abstract*— Vehicles detection and tracking have become an important role to traffic management systems. Recently, many vehicles tracking approaches have already been proposed. However, these approaches were unable to adequately distinguish vehicles from each other when those vehicles look similar and involve in complex transportation conditions. In this paper, a method for tracking vehicles in surveillance cameras is presented. In our method, Convolutional Neural Networks is used to detect vehicles. Also, multiple Kalman filters are used to track those vehicles. The proposed method is designed for distinguishing and tracking multiple vehicles simultaneously. Our experiments show that the proposed mechanism achieves high accuracy even with real time constraints.**

***Keywords—* Convolution Neural Network; Kalman Filter; Vehicles Tracking**

I. Introduction (*Heading 1*)

Camera surveillance provide a flexible way of monitoring the transportation, especially monitor the complex transportation. In intelligent transportation systems (ITS) [1, 2, 3], identifying moving vehicles from camera video stream is a fundamental task. The task of identifying vehicles is normally performed in two steps: vehicle detection and vehicle tracking. In the two mentioned steps, vehicles under monitoring are detected and then tracked by the surveillance system. More specifically, given a video stream recorded by surveillance systems, detection and tracking algorithms will identify target vehicles in consecutive time. The output of these two algorithms will be send to the transport management center for further analysis such as vehicles speed detection, vehicles breaking traffic rule and traffic monitoring [4, 5].

There are number of tracking approaches being used in surveillance systems where vehicle tracking is one of essential case. The most common method for object tracking is using Kalman filters, which are recursive estimators for states of dynamic systems [6, 7]. To increase the accuracy, mean-shift was combined with Kalman filter to predict the search regions [8]. If the system does not fit into linear model, particle filter is an important method to track the object [9]. It combines gray and contour feature particles using fusion algorithm to balance the weights according to the present scene. Motion direction and assignment can be used to track the vehicles in their lanes and calculate the speed of the vehicles [10]. Image segmentation and pattern analysis techniques are also applied in the system to detect and track the moving vehicles at day and night time [11] by recognize headlight and taillight of vehicles. Using cameras and the pattern recognition techniques, the traffic flow can be measured under various environments conditions by detecting vehicles.

The current techniques and algorithms for detection and tracking in transportation surveillance systems are still facing challenges that are not completely solved. Under bad weather and bad transportation conditions, especially where multiple vehicles run concurrently without orders, the tracking algorithms cannot accurately and efficiently track vehicles. This paper propose a method that simultaneously tracks vehicles in a sequence of video frames using multiple Kalman filters. The method detect moving vehicles in each frame and associates these vehicles corresponding to those in successive frames. Particularly, Kalman filters are used to predict vehicle positions and use predicted positions for associations. Experimental results show that the proposed algorithm is able to perform multiple vehicles tracking simultaneously with high level of robustness and efficiency.

The rest of this paper is organized as follow: Section 2 presents a framework of the vehicle detection and tracking. In this section, the proposed method for vehicles tracking using multiple Kalman filters is presented. Section 3 demonstrates the accuracy and robustness of the proposed method. Finally, Section 4 states the conclusions and future works.

II. Method

The goal of this research is to track multiple vehicles in complex transportation situations. In order to achieve this purpose, this paper propose to use multiple Kalman filters to track multiple vehicles concurrently. To do this, firstly, convolution neural network is used to detect vehicles existing in a frame. According to the number of detected vehicles, a corresponding number of Kalman filters are, then, created. Finally, those filters are used to track detected vehicles in successive frames. The general framework of the method is given in figure 1.

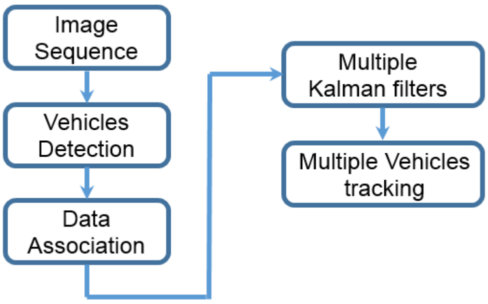


Fig. 1. The framework of the multiple tracking method

A. Convolution Neural Network for Vehicles Detection

Lots of methods vehicle detection methods exist, for example Support Vector Machine (SVM), Gaussian Mixture Model (GMM), or background subtraction. However, those methods are costly in computation and heavily affected by the weather condition, especially when shadows of vehicles appear. In this paper, Convolution Neural Networks (CNN) method is used to detect vehicles running in road. A CNN comprises of convolution and pooling layers [12]. Those layers are then connected to one or more fully-connected layers. Convolution and pooling layers extract the feature maps, which are two dimensional matrices of CNN neurons. With the input image x_i, the output of a convolution layer *j*, denoted by $y_j = b_j + \sum_i (k_{ij} \otimes x_i)$, where $\otimes$ denoted the convolution operator, b_j is a trainable bias parameter, k_{ij} is a convolution layer filter. The feature map y is calculated for any node *y(m,n)*:

$$y(m,n) = k \otimes x = \sum_{u=0}^{U} \sum_{v=0}^{V} k(u,v) x(m+u, n+v)$$

where *k* is the kernel of size *A*B* and *x* is the input image with size *U*V*. The size of the output convolutional is *M*N* where *M=U-A+1* and *N=V-B+1*.

The multi-layer structure of CNN brings advantages to the task of vehicle detection. When frames are processed in convolution layers, those layers incrementally learn features from raw images and outputs of the previous layers, which are high level features such as shapes and edges. Convolution layers, thus, represent an image frame into multiple representations at each convolution layer with different levels of abstraction from low to high. This effectively helps in cancelling out noises and refining detection information. The final step of detection is done at pooling layer at which feature maps are extracted and processed so that vehicles are detected regardless of translation, rotation, scaling and other kinds of geometric transformations. As a result, CNN can provide robust detection regardless of where in road a vehicle is captured and which camera is used to capture the vehicle

B. Vehicles Tracking

We use Kalman filter to predict each vehicle in a specific point in time. Basically, a Kalman filter is used to estimate states of a linear system where states are assumed to be Gaussian random variables. Kalman filter algorithm comprises of two steps: prediction and correction. In prediction step, a state is estimated using a state equation. After that, the correction step takes current observations to adjust and update the estimated state in the prediction step. . In this paper, to track multiple vehicle simultaneously, multiple Kalman filters as number of vehicles is used (Jeong et al., 2014). Each Kalman filter is represented as below:

$$x_k = Ax_{k-1} + w_k$$
$$z_k = Hx_k + v_k$$

where $x = \left[p_x \; p_y \; v_x \; v_y \right]^T$, 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 fames. 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 measurements should be performed correctly. In this paper, we employ the data association method, which split and merge the vehicles [14]. Overall of the tracking method is given in figure 2.

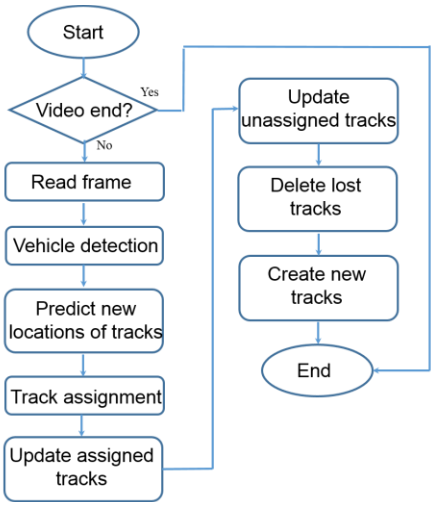


Fig. 2. The flow chart of vehicles tracking method

III. Experimental Results

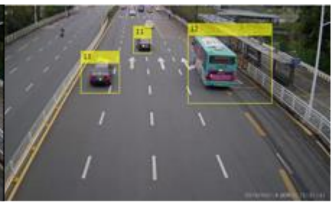
A. Vehicles detection

The first step of object tracking is object detection. The data used in this paper were collected from [15]. Vehicles are detected using Convolution Neural Network. Figure 3 (a) shows the single car in the image captured from camera. Figure 3 (b) shows the car was detected with the bounding box. Figure 3 (c) shows the multiple vehicles including car and bus from camera and the detected vehicles are shown in figure 3 (d).

(a)

(b)

(c)



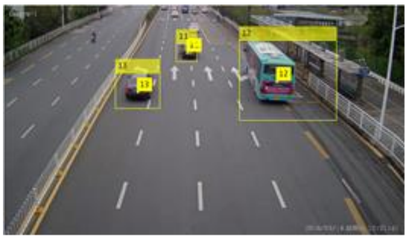
(d)

Fig. 3. (a) input image with single car, (b) the car detected, (c) input image with multiple vehicles, (d) multiple vehicles detected

We initial the track with this object, the Kalman filter is used to estimate the vehicles in the next frame.

B. Vehicles Tracking

(a)



(b)

Fig. 4. Vehicle tracking. (a) single car tracking, (b) multiple vehicles tracking

Figure 4 (a) shows the tracking results for the video of single car was tracked using object detection algorithm presented in the previous section. We use the number at the center of the vehicle for multiple vehicles tracking purpose as shown in Figure 4 (b). The Kalman filter implements two steps: prediction by estimate the state of the object and correction using measurement of object.

IV. Conclusion

In this paper, we presented a tracking method for multiple vehicles based on Kalman filter. For each vehicle, a Kalman filter was established and it uses bounding box as feature. The Kalman filter estimates states based on the state equation and corrects using the current observations to update the vehicle states. Results of this paper show that this method can be applied in transport management system for traffic monitoring.

Acknowledgment

This research was supported by University of Transport and Communications under grant number T2019-CN-013 TĐ.

References

[1] B. E. Flinchbaugh, and T. J. Olson, “Autonomous video surveillance”, Proceedings of the SPIE, vol. 2962, pp. 144–151, 1997.

[2] H. Moon, R. Chellapa, A. Rosenfeld, “Performance analysis of a simple vehicle detection algorithm”, Image and Vision Computing 20. 1–13, 2003.

[3] S. M. Baljit and K. Satish, “A Review of Computer Vision System for The Vehicle Identification and Classification from Online and Offline Videos”, Signal & Image Processing: An International Journal (SIPIJ) Vol.6 (5), 2015.

[4] C. Wang, C. Thorpe and A. Suppe, “Lidar-Based Detection and Tracking of Moving Objects from a Ground Vehicle at High Speeds”, Proceeding of the IEEE Intelligent Vehicles Symposium, 2003.

[5] C. Setchell, and E. Dagless, “Vision-based Road Traffic Monitoring Senso”, IEE Proc. Vision, Image Signal Processing, Vol. 148 (1), pp. 78-84, 2001.

[6] G. S. Manku, P. Jain, A. Aggarwal, and L. Kumar, “Object tracking using affine structure for point correspondence”, IEEE conference of Computer Vision and Pattern Recognition, 704-709, 1997.

[7] L. Xin, W. Kejun, W. Wei, and L. Yang, “A Multiple Object Tracking Method Using Kalman Filter”, IEEE International Conference on Information and Automation, pp. 1862-1866, 2010.

[8] D. Comaniciu, and V. Ramesh, “Mean shift and optimal prediction for efficient object tracking”, IEEE Int. Conf. Image Processing, vol. 3, pp. 70–73, Vancouver, Canada, 2000.

[9] D. P. Bai, , B. B. Lee, “Based-on Particle Filter For Vehicle Detection And Tracking In Digital Video”, Proceedings of the Seventh International Conference on Machine Learning and Cybernetics, Kunming, pp. 12-15, 2008.

[10] H. S. Lai, and H. C. Yung, “Vehicle-Type Identification Through Automated Virtual Loop Assignment and Block-Based Direction-Biased Motion Estimation”, IEEE Transactions on Intelligent Transportation System, Vol. 1(2), 2000.

[11] K. Dhanya, M. Manimekalai, B. Asmin, and G. Vani, “Tracking and Identification of Multiple Vehicles”, 2014.

[12] S. Christian, T. Alexander, E. Dumitru, “Deep Neural Networks for Object Detection”, Advances in Neural Information Processing Systems 26, pp. 1-9, 2013.

[13] J. M. Jeong, , T. S. Yoon, and J. B. Park, “Kalman Filter Based Multiple Objects Detection-Tracking Algorithm Robust to Occlusion”, SICE Annual Conference, Japan, pp. 941-947, 2014.

[14] S. Zheng, “Tracking Multiple Objects In Surveillance Cameras”, Technical Report, 2010.