Fast and Reliable Tracking Algorithm for On-Road Vehicle Detection Systems

Jang Woon Baek
Smart Vision Research Section
ETRI
Daegu, Korea
jwbaek98@etri.re.kr

Byung-Gil Han
Smart Vision Research Section
ETRI
Daegu, Korea
kilyhan@etri.re.kr

Hyunwoo Kang
Smart Vision Research Section
ETRI
Daegu, Korea
hwkang@etri.re.kr

Yoonsu Chung
Smart Vision Research Section
ETRI
Daegu, Korea
yoonsu@etri.re.kr

Su-In Lee
Daegu-Gyeongbuk Research Center
ETRI
Daegu, Korea
silee@etri.re.kr

***Abstract*—In this paper, we proposes a novel tracking algorithm combining Kalman Filter with mean-shift. Kalman Filter predicts the vehicle position in the next frame. Mean-shift finds the best candidate which has maximum similarity with the tracked vehicle in the predicted area. Kalman Filter updates its state value of vehicle position with the position of the best candidate from the mean-shift tracker. As a result, the proposed algorithm tracks the vehicle without local maximum problem of mean-shift tracker. The proposed algorithm is very fast because it does not perform the redetection process, and it has no detection misses because it finds the best candidate which has maximum similarity with the tracked vehicle in the predicted area. Also, the proposed algorithm has deleting and adding policies for the tracking list management. If a vehicle was consecutively detected in the previous frames, the proposed algorithm assumes that the vehicle exists although the vehicle is not detected in the predicted region at the current frame. If a vehicle was not detected in the previous frames consecutively, the proposed algorithm assumes that the vehicle does not exist although the vehicle is detected in the current frame. We evaluated the performance of the proposed algorithm in terms of processing time and detection ratio. At target board, the proposed algorithm has 40 frames per second, which meets the real time requirements of the ADAS systems. The detection ratio and processing time of the proposed algorithm outperformed our former work.**

Keywords—vehicle detection; vehicle tracking; Kalman Filter; mean-shift

I. INTRODUCTION

Most of the traffic accidents occurred due to driver's inattention according to a survey about the causes of car accidents [1]. In order to reduce the accidents due to driver's inattention, the needs of Advanced Driver Assistance Systems (ADAS) increased. ADAS systems provide the drivers with warning for the traffic accidents, or control the internal vehicle system to avoid accidents or reduce collision impact. The vehicle detection technology is essential in the ADAS systems such as autonomous emergency braking systems (AEBS), forward collision warning systems (FCWS), advanced cruise control systems (ACCS), blind spot detection systems, lane change assist systems (LCAS), and so on. Especially, EURO NCAP has employed the AEB functions to the car safety assessment program since 2014.

In this paper, we focused on the camera based vehicle detection because it provides more information than other sensors, and is very cost effective. Camera based vehicle detection consists of the candidate generation phase, the verification phase, and the tracking phase[2]. In the candidate generation phase, vehicle candidates are generated by using characteristics of a vehicle appearance such as symmetry, horizontal and vertical edge, shadow of a vehicle [3-6]. In verification phase, feature vectors are extracted from every candidate region. The feature vectors are used to classify between vehicles and non-vehicles. The feature vector could be Haar-like features [7], local binary patterns (LBP) [8], modified census transform (MCT) [9,10] or histogram of oriented gradients (HOG) features [11, 12]. In tracking phase, the verified vehicles are tracked in subsequent images using a tracking algorithm. Kalman Filter [12], particle filter [13, 14], mean-shift tracker [15] are used to track objects such as persons, vehicles, and so on. In our former work, we used Kalman Filter (KF) which predicts the position of the verified vehicle in the next frame and updates its state values with the measured value, which means the re-detected position in the predicted area. Using tracking decreases the processing time for vehicle detection because the verification is performed within the predicted area without scanning a full image. However, it sometimes does not detect the vehicle at the predicted area. If the predicted area does not include the entire part of the tracked vehicle, the tracked vehicle cannot be not detected.

II. PROPOSED ALGORITHM

The proposed algorithm combines Kalman Filter with Mean-Shift in order to removes the problems of Kalman Filter based tracking, in which the detection error sometimes occurs at the area predicted by KF. Also, the proposed algorithm manages the tracking list, by comparing the tracked vehicles and the verified vehicles which is periodically detected, and adding the newly detected vehicle to the tracking list, instead of initializing the tracking list at every detection time. Fig.1 shows the flow diagram of the vehicle detection procedure. First, the video query process retrieves an image frame from the camera installed on the car. The image frame is scaled down in order to reduce the processing time. The next step is to detect vehicles in the scaled down image. The vehicle detection is performed periodically because it consumes much processing time. For the intermediate frames, vehicle tracking is used. If we configure the detection period is five, the vehicle detection process is performed per a fifth frame, and the detected vehicles are tracked in the four intermediate frames. We use the cascade classifier to detect vehicles in the image. The cascade classifier is very fast because it is enable to early drop almost negatives at first and second stages. The processing time is very important because of the real time requirement of the ADAS systems. We trained the cascade classifier with MCT features which has much less false alarm than LBP features, and is faster than HOG and Haar feature. The MCT feature has 511 patterns while LBP has 256 patterns. This means that the MCT feature is more discriminant more than LBP feature. The MCT classifier computes the weak classifier by one array reference to the lookup table, while HOG and Haar feature based classifier computes the weak classifier by several array references. The number of positive dataset is 10,000, and the number of negative dataset is 30,000. We obtained the dataset using real video records from the cameras installed in the vehicle. We uses two cascade classifiers. One scans the full image and detects the vehicles which includes many false positives. Another verifies the vehicles detected by the former classifier. Using two phase classification dramatically reduces false alarms. The widow size of the first classifier is 16 pixels * 16 pixels, and the second classifier has 32 pixels * 32 pixels window size. The confidence threshold of the second classifier is higher than one of the first classifier. The verified vehicles from the vehicle detection process are tracked using Kalman Filter. KF initializes its state value with the position of the verified vehicles. And it predicts the position of the vehicles at the next image frame. The proposed algorithm detects the vehicle around the predicted position. After, KF updates the state value with the position of the redetected vehicles. There are sometimes tracking misses in the image region predicted by Kalman Filter. If the predicted area does not include the entire part of the vehicle, the vehicle might not be redetected by a vehicle classifier. In order to this problem, we proposed a tracking algorithm which combines the KF filter and the mean-shift algorithm. The mean shift algorithm is an efficient approach to tracking objects whose appearance is defined by histogram. If the means-shift tracker set the tracked vehicle in the current frame, it searches in the tracked vehicle's neighborhood in the next frame, and find the best candidate maximizing a similarity function. The same processes are repeated in the next pair of frames. In the mean shift algorithm, it is important to set the search window location and the size. If the search window size is total image area, it is easy to enter the local maximum. However, the mean-shift tracker combined with KF can avoid the local maximum problem . KF predicts the vehicle position in the next frame. Mean-shift tracker finds the best candidate which has maximum similarity with the tracked vehicle in the predicted area. KF updates its state value of vehicle position with the position of the best candidate from the mean-shift tracker. As a result, the proposed algorithm tracks the vehicle without local maximum problem of mean-shift tracker. The proposed algorithm is very fast because it does not perform the redetection process, and it has no detection misses because it finds the best candidate which has maximum similarity with the tracked vehicle in the predicted area.

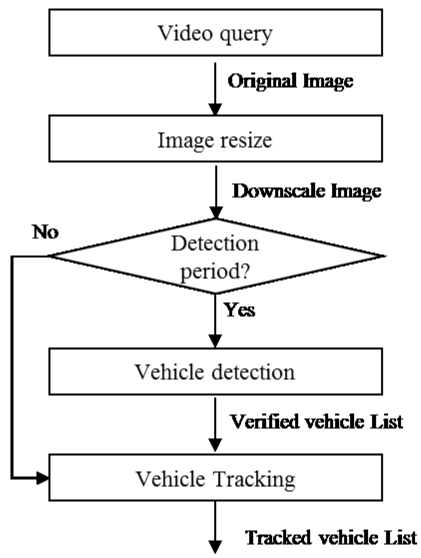

Fig. 1. Vehicle detection algorithm

In our previous research, the vehicle detection algorithm periodically detects vehicles, and initializes the tracking list with the new detected vehicles. However, there are false positive or true negative in detection result. In this paper, we designed the tracking algorithm to maintain the tracking list by comparing the newly detected vehicles and the tracked vehicles. Instead of initializing the tracking list with the newly detected vehicles, the proposed algorithm determines whether a vehicle of the tracking list belongs among the newly detected vehicles. We use intersection ratio between the tracked vehicles and the new detected vehicles in order to determine the redetection of the tracked vehicles. If a vehicle in the tracking list is redetected, the proposed algorithm update the vehicle position with the new position . If a vehicle in the tracking list is not detected, the vehicle removed from the tracking list. Other than the tracked vehicles, the newly detected vehicles are added to the tracking list. The proposed algorithm has deleting and adding policies for the tracking list management. The proposed algorithms does not removes the vehicle from the tracking list which was consecutively in the previous frames although the tracked vehicle is not detected in the current frame. In the deleting and adding policies, we use a reliability point (RP) in order to reduce false detections and to increase detection ratio. If a vehicle was consecutively detected in the previous frames, the proposed algorithm assume that the vehicle exists although the vehicle is not detected in the predicted region at the current frame. If a vehicle was not detected in the previous frames consecutively, the proposed algorithm assume that the vehicle does not exist although the vehicle is detected in the current frame. The proposed algorithm increases the RP value if the tracked vehicle is redetected, but it decreased the RP value if the tracked vehicle is not detected. The proposed algorithm set the maximum RP value, the minimum RP value, and the threshold value. The RP value cannot be higher than the maximum RP value, and lower than the minimum RP value. The tracked vehicle enter invalid state if its RP value is lower than the threshold values, and it is deleted if it RP value research the minimum RP value. Although the tracked vehicle is not detected in current frame, it is considered as a detected vehicle if its RP value after decreasing one point is higher than the threshold value.

We tested our vehicle detection algorithm at target board which has 1GHz i.MX6Quad and 1GB RAM. We evaluated the performance of the proposed algorithm in terms of processing time and detection ratio. At target board, processing time per an image frame is 66ms without the proposed tracking algorithm. In case of using the proposed tracking algorithm, average processing time per an image frame (detection period = 5) is 21ms, which means that 48 frame per second can be processed. This processing time is suitable for the ADAS systems. We extracted ground truths (Gt) from the video file which was recoded from the cameras at the subject vehicle. We uses PASCALVOC measure as detection result (Dt). The size of image may affect the result of detection performance. We tested the performance of the detection ratio according to the image size. Table 1 shows the detection ratio with different image sizes and different detection periods. The detection ratio of the proposed algorithm is over 96 % at both original image size and downscale image size.

TABLE I. DETECTION RATIO WITH DIFFERENT TRACKING

	KF	**KF + Mean Shift**
Detection period = 5frame	# of Gt = 280 # of Dt = 255 Detection ratio = 91%	# of Gt = 280 # of Dt = 255 Detection ratio = 96%

We will improve the detection ratio at vehicle tracking time by optimizing Kalman filter parameters and adjusting detection period. Also, we will apply road surface detection to the proposed algorithm in order to removing false alarms outside the road.

REFERENCES

[1] National Transportation Safety Board, Special Investigation Report - Highway Vehicle- and Infrastructure-based Technology For the Prevention of Rear-end Collisions. NTSB Number SIR-01/01, May 2001.

[2] Zehang, S., Bebis, G., Miller, R., "On-road vehicle detection: a review," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28(5) pp. 694-711, 2006.

[3] C. Txomakas and W. Seelen, "Vehicle Detection in Traffic Scenes Using Shadows," IR-INI 98 06, Aug. 1998.

[4] B. Dai, Y. Fu, T. Wu, "A Vehicle Detection Method via Symmetry in Multi-Scale Windows," IEEE Conference on Industrial Electronics and Applications, pp. 1827-1831, 2007.

[5] Z. Sun, R. Miller, G. Bebis, and D. DiMeo, "A Real-time Precrash Vehicle Detection System," IEEE Workshop on Applications of Computer Vision, pp. 171-176, 2002

[6] J. Baek, K. Kwon, and S. Lee, "Mono-camera based Vehicle Detection Using Effective Candidate Generation," Proceedings of ISCE2014, 2014.

[7] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," Conference on Computer Vision and Pattern Recognition, 2001.

[8] S. Liao, X. Zhu, Z. Lei, L. Zhang, and S. Z. Li, "Learning Multi-scale Block Local Binary Patterns for Face Recognition," Proceedings of International Conference on Biometrics, 2007.

[9] B. Han, J. Lee, K. Lim, and Y. Chung, "Real-Time License Plate Detection in High-Resolution Videos Using Fastest Available Cascade Classifier and Core Patterns," ETRI Journal, Vol. 37, No. 2, pp. 251-261, Apr. 2015.

[10] K. Ban, J. Kim, and H. Yoon, "Gender Classification of Low-Resolution Facial Image Based on Pixel Classifier Boosting," ETRI Journal, Vol. 38, No. 2, Apr. 2016.

[11] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," Conference on Computer Vision and Pattern Recognition, 2005.

[12] J. Baek, E. Lee, M. Park, and D. Seo, " Mono-Camera Based Side Vehicle Detection for Blind Spot Detection Systems," Proceedings of ICUFN2015, 2015.

[13] J. Choi, D. Moon, and J. Yoo, "Robust Multi-person Tracking for Real-Time Intelligent Video Surveillance," ETRI Journal, Vol. 37, No. 3, Jun. 2015.

[14] M. Yildirim, I. Ince, Y. Salman, J. Song, J. Park, and B. Yoon, "Direction-Based Modified Partical Filter for Vehicle Tracking," ETRI Journal, Vol. 38, No. 2, Apr. 2016.

[15] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time Tracking of Non-Rigid Objects using Mean Shift," IEEE Conference on Computer Vision and Pattern Recognition, 2000.