# An Intelligent Vision-based Vehicle Detection and Tracking System for Automotive Applications

Yi-Min Tsai, Chih-Chung Tsai, Keng-Yen Huang, and Liang-Gee Chen, *Fellow*, *IEEE*
DSP/IC Design Lab., Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan

***Abstract*—In this paper, we present an intelligent vision-based on-road preceding vehicle detection and tracking system based on computer vision techniques. Pre-processing video stabilization is adopted to improve system reliability and stability. High performance detection is achieved via the machine learning-based method. Our framework is favored for various automotive applications, which yields above 90% detection rate in long range and 99.1% tracking successful rate in middle range.**

## I. Introduction

Owing to the maturity of vision sensors, vision-based systems play an essential role in many vehicular applications. Intelligent Cruise Control (ICC) provides semi-automatic driving mechanism. Collision Warning Systems (CWS) prevent vehicles from sudden crashes. These applications require detection and tracking of on-road vehicles to estimate distance between host and preceding vehicles or to monitor vehicles' behaviors on different lanes.

In these years, machine learning makes progress in object analysis. Several learning-based systems are proposed for on-road vehicle detection and tracking [1]–[5]. Sivaraman et al proposed an active-learning-based method for recognition and a condensation filter for tracking [1]. Sun et al made a summary for vehicle analysis composed of hypothesis generation and verification steps [4]. In this paper, we propose a cooperative detection and tracking framework utilizing both learning techniques and prior-knowledge on vehicle's attributes.

## II. Proposed Framework

The overall system is shown in Fig. 1. Video stabilization is firstly introduced to enhance system accuracy and stability. The following are a hierarchical detection routine and Adaboost-based recognition in dynamic region of interests (ROI). Vehicle prediction and tracking are adopted not only to estimate exact width but to reduce false positives. Moreover, we exploit system specification and capability based on both experimental observations and mathematical derivation.

### A. Video Stabilization

Video stabilization is a technique to remove unwanted oscillation from a shaky video stream. It mainly relies on global motion estimation (GME). GME attempts to estimate the global motion and to separate the motion into intentional motion (IM) and unwanted motion (UM). By subtracting the UM, a stabilized video can be obtained from a shaky condition.

We propose a video stabilization method based on feature analysis. Camera motion model is firstly investigated. Secondly, an image pyramid is constructed by down-sampling. Harris features with SURF-like descriptors are extracted from the selected layer. Features are matched using KD-tree to construct a feature flow field. Lastly, after estimating the GM from the feature flow, a damping filer [6] is utilized to model and predict the unwanted oscillation.

### B. Vehicle Detection and Recognition

To locate vehicles, a window at first progressively scans the ROI of each layer in the image pyramid. The ROIs form strips centering on image horizons. The scanning window scales up and becomes candidates for boost-based recognition. Secondly, Adaboost with Haar-like features is adopted for recognizing vehicles' rears. Each scaled window is verified with boosted cascades composed of trained

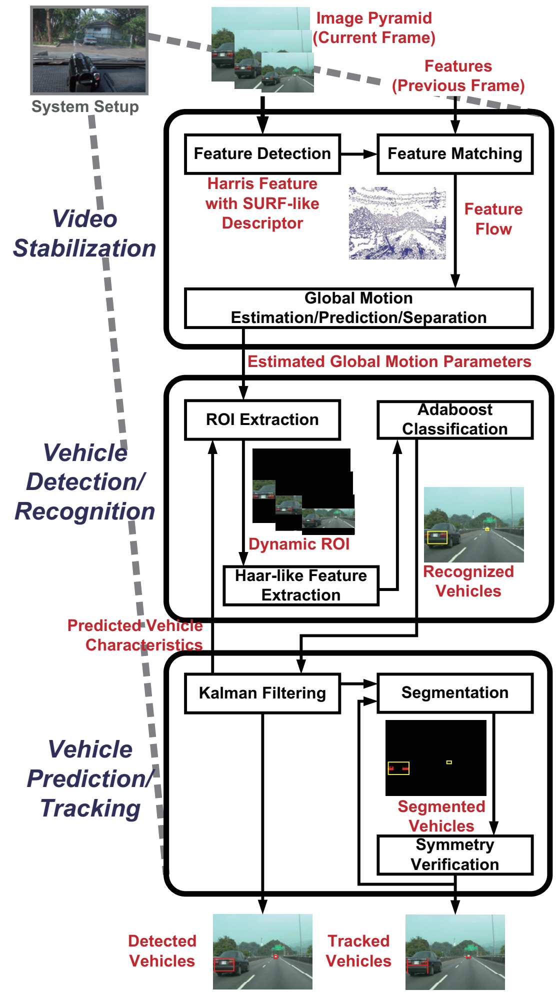



Fig. 1. System flow block diagram

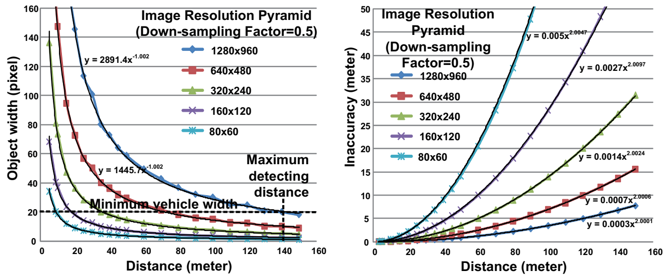
weak classifiers. Non-vehicle windows can be rapidly rejected with early classifiers. More complex classifiers are performed subsequently on vehicle-like windows only if they passed through previous ones.

Vehicles are diverse in object aspect ratio, which suggests variable window size is preferred. Accordingly, multiple cascades are adopted. With multi-resolution strategy, far and near vehicles can be detected in high and low resolution layers respectively. Besides, the system dynamically adjusts ROIs according to global translational motion and position of predicted vehicles in previous frame.

### C. Vehicle Prediction and Tracking

The Kalman estimator is utilized to predict vehicles' characteristics in successive frames. Four state variables, 2D location and size, are used to describe dynamics of detected vehicles and are regarded as feedback. Through Kalman filtering, the system avoids performance degradation caused by detection errors.

We bring up a knowledge-based vehicle tracking considering spatial appearance: Vehicles' rears are in general laterally symmetrical.




(a) (b)

Fig. 2. System property curves and their regression under different image resolutions. A laser range measurer is used for distance measuring. Several calibration images are selected for estimating object width in distinct distance. (a)Relationship between detecting distance and vehicle width. (b)Relationship between detecting distance and induced distance inaccuracy.

Morphological color segmentation in $L^*a^*b^*$ is to extract intact red regions representing rear lights. Symmetry verification explicitly estimates vertical boundaries of vehicles' rears. Because initial region for segmentation is generated from previous tracked results and Kalman prediction, the tracking maintains temporal coherence.

### *D. System Specification and Capability*

Fig. 2(a) describes object width is approximately inversely proportional to the detecting distance for different image resolutions, which is given by,

$$w_l = K{\cdot}W_l{\cdot}D_l^{-1+\sigma} \tag{1}$$

where $D$ is detecting distance. $w$ and $W$ are vehicle and image width in pixel respectively. The subscript $l$ stands for layer index. $K$ is a constant depending on pixel aspect ratio and sensor characteristics. The maximum detecting range is defined as the distance given the minimum vehicle width in pixel set for training. The higher image resolution is, the longer detecting distance system supports.

Fig. 2(b) illustrates the induced distance estimation inaccuracy, $\lambda$, caused by one integer-pixel measuring error on object width under different image resolutions and detecting distances, which is formulated as,

$$\begin{aligned}\lambda_l &= \frac{K{\cdot}W_l}{K{\cdot}W_l{\cdot}D_l^{-1+\sigma}-0.5} - \frac{K{\cdot}W_l}{K{\cdot}W_l{\cdot}D_l^{-1+\sigma}+0.5} \\ &= \frac{K{\cdot}W_l}{(K{\cdot}W_l{\cdot}D_l^{-1+\sigma})^2-0.25} \\ &\simeq K^{-1}W_l^{-1}D_l^{2+\delta}\end{aligned} \tag{2}$$

From Eq. 2, the higher resolution is, the less induced distance inaccuracy to preceding vehicles system conducts. Furthermore, detecting vehicles at far range implies larger distance inaccuracy.

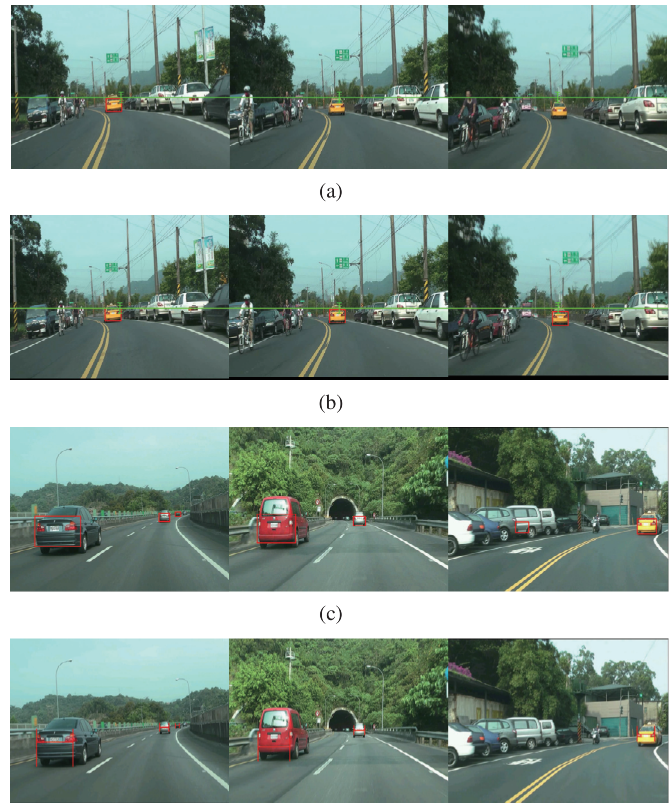
## III. Experimental Result and Demonstration

Real sequences were captured from a CMOS front-mounted camera with 1280×960 resolution. Adaboost cascades are trained in four aspect ratios (20×20, 20×16, 20×12, 20×10). Thus, theoretical maximum detecting distance is about 140 meters. Table I summarizes detection rate (DR) with similar false alarm rate (FAR) in four distance ranges. Proposed system achieves 91.2% DR within 140 meters. Given initial detected objects, successful rate (SR) is defined as average tracked ratio during the entire sequence. Our system yields 99.1% SR within 80 meters. Both DR and SR decrease as detecting distance increases. However, because rear lights are nearly indistinguishable at long distance, SR drops more drastically. This suggests detection mode is preferred for long-distance requirements.

Fig. 3(a) and Fig. 3(b) show the improvement on detection involving stabilization. With video stabilization, vehicles can be accurately located, especially far ones. Fig. 3(c) demonstrates detection results in several driving scenarios, including highway and typical road. The corresponding tracking results are shown in Fig. 3(d). We can observe false alarms are further reduced via tracking mechanism. Precise vehicle width is also obtained for evaluating the relative distance.

TABLE I
Detection rate and tracking successful rate with respect to detecting distance.

| Detecting Distance (meter) | 0∼80 | 0∼100 | 0∼120 | 0∼140 |
|---|---|---|---|---|
| Detection Rate (%) | 97.1 | 96.5 | 94.3 | 91.2 |
| False Alarm Rate (%) | 4.0 | 4.0 | 4.0 | 3.9 |
| Successful Rate (%) | 99.1 | 92.1 | 84.7 | 81.4 |



(a)

(b)

(c)

(d)

Fig. 3. Results for three processing stages. (a)Detection without stabilization. (b)Detection with stabilization. (c)Detection in various situations. (d)Corresponding tracking outcomes to (c).

## IV. Conclusion

We introduce an intelligent vision-based on-road vehicle detection and tracking system. The system achieves high detection rate and tracking successful rate. Convincing results are demonstrated in diverse driving conditions. The processing capability is also explored considering distance criterion. In conclusion, we establish a framework suitable for developing versatile automotive applications.

## References


[1] S. Sivaraman and M.M. Trivedi, "A General Active-Learning Framework for On-Road Vehicle Recognition and Tracking," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 11, pp. 267-276, June 2010.

[2] Z. Sun, G. Bebis, and R. Miller, "On-road vehicle detection using evolutionary Gabor filter optimization," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 6, pp. 125-137, June 2005.

[3] C.-M. Fu, C.-L. Huang, and Y.-S. Chen, "Vision-Based Preceding Vehicle Detection and Tracking," in *Proc. IEEE International Conference on Pattern Recognition*, vol. 2, pp. 1070-1073, 2006.

[4] Z. Sun, G. Bebis, and R. Miller, "On-Road Vehicle Detection: A Review," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, pp. 694-711, May 2006.

[5] W.-C. Chang and C.-W. Cho, "Online Boosting for Vehicle Detection," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 40, pp. 892-902, June 2010.

[6] K.-Y. Huang, Y.-M. Tsai, C.-C. Tsai, and L.-G. Chen, "Feature-based video stabilization for vehicular applications," *Consumer Electronics, IEEE International Symposium on*, 2010.