On-road Vehicle Detection with Monocular Camera for Embedded Realization: Robust Algorithms and Evaluations

Ravi Kumar Satzoda*, Eshed Ohn-Bar*, Jinhee Lee+, Hohyon Song+, and Mohan M. Trivedi*

*Laboratory for Intelligent and Safe Automobiles,
University of California San Diego, La Jolla, CA-92093
rsatzoda@eng.ucsd.edu, eohnbar@ucsd.edu, mtrivedi@ucsd.edu
+NextChip Co. Ltd.
jini92.lee@nextchip.com, hodal@nextchip.com

Abstract

Vehicle detection is critical operation in automotive active safety systems. Although there are a number of vehicle detection techniques available in literature, computationally efficiency for realization on embedded platforms is not explored and addressed in most existing works. In this paper, we present a computationally efficient vehicle detection algorithm that is particularly designed for architectural translation into efficient embedded hardware. The proposed method uses camera calibration to derive the appropriate window scales that must be used for vehicle detection, resulting in a computational cost reduction of over 10 times. In addition to reduction in sampling windows, the proposed vehicle detection technique uses a novel multi-part based vehicle detection method which detects the vehicles that pose the highest risk to the ego-vehicle. The proposed method is evaluated using different datasets and computational savings are seen in orders of magnitude as compared to conventional sliding window approaches, without compromising on accuracy.

Keywords- vehicle detection, computational efficiency

I. Introduction

Although there is a rich amount of literature on the different tasks involved in drive analytics [1]-[3], and advanced driver assistance systems (ADAS) such as lane detection [4]-[6], and vehicle detection [7]-[9], few works address computational efficiency of constituent algorithms [5][6]. Among the different ADAS operations, detection of on-road vehicles is a critical operation in advanced driver assistance systems (ADAS) such as collision avoidance, lane change assistance etc. [7]. Use of monocular cameras in order to detect vehicles has been explored in studies such as [7]-[9]. While [7] and [8] describe techniques to detect the rear-views of the vehicles using forward looking cameras, [9] describes techniques to detect overtaking vehicles from the rear-view of the ego-vehicle.

In this paper, we present a novel vehicle detection technique that is particularly aimed at improving the computational efficiency of the vehicle detection process. Additionally, considering that vehicle detection is an operation for active safety, it is critical that robustness of the vehicle detection process is not compromised. The proposed method involves part-based detection of vehicle, while leveraging on the camera calibration to reduce the computational complexity.

II. Proposed Method

A. Contextual Information for Detecting On-road Vehicles

The proposed method uses the camera calibration and context for processing in order to detect the vehicles using a forward looking camera from the ego-vehicle. Fig. 1 shows the contextual information for vehicles as seen from the forward looking camera. It can be seen that the vehicle in front of the ego-vehicle is of highest risk to the ego-vehicle followed by the vehicle that is farther away. Also the vehicles appear in a particular manner, i.e. the closest vehicles appear first, followed by the next vehicle, and so on, if the input image frame is scanned from the closest point to the ego-vehicle towards the vanishing point.

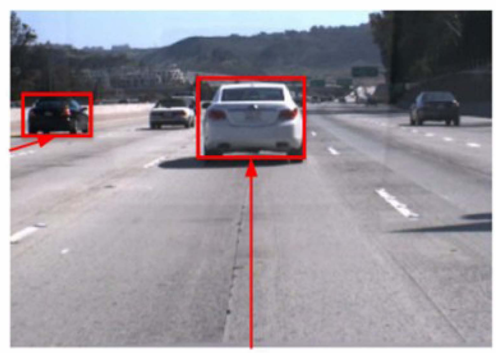

Fig. 1. Contextual information about vehicles seen from forward looking images.

B. Vehicle Detection Method

This contextual information is used to devise the proposed method in the following way. Instead of scanning the entire image using a sliding window of different sizes, which is conventionally done in existing methods, the proposed method first uses the camera calibration to generate an LUT of window sizes. Given the homography matrix H that maps the image space to world plane (inverse perspective mapping - IPM) , a given axle width (w_A) of the vehicle, two points

$P_1^W(x_1^W, y_1^W)$ and $P_2^W(x_2^W, y_1^W)$ in IPM such that $x_2^W - x_1^W = w_A$, we get the window width in the image domain at a given y position (vertical axis along the image domain) using the following equation:

$$[x_1 \quad y_1 \quad 1]^T = \chi H^{-1}[x_1^W \quad y_1^W \quad 1]^T$$
$$[x_2 \quad y_1 \quad 1]^T = \chi H^{-1}[x_2^W \quad y_1^W \quad 1]^T$$
$$w(y_1) = x_2 - x_1; h(y_1) = \alpha w(y_1)$$

where $w(y_1)$ and $h(y_1)$ are the width and height of the sliding window to detect the vehicle at $y=y_1$ (vertical axis) in the image domain). An LUT for all y in the image domain for an $M \times N$ (columns x rows) sized image is generated, which will be used in the next step.

Given the LUT of window sizes, a sliding window I_V with size defined in the LUT is scanned over the image. Each window is divided into two parts P_1 and P_2 such at the height of P_1 is 1/3-rd of the window height. P_1 corresponds to the lower part of the vehicle from the bottom of the vehicle till the bumper of the vehicle. P_2 captures the upper part of the vehicle.

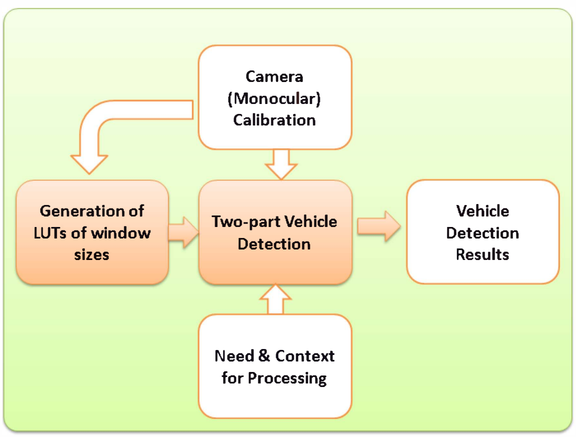

Fig. 2. Proposed vehicle detection method.

HOG features are computed for P_1 first, and a support vector machine (SVM) is used to classify whether it is the lower part of the vehicle. If the classifier detects the lower part, then HOG features are computed for the upper part, and SVM classifier is used to detect the upper part. A vehicle is considered detected if the following condition is met:

$$I_V = vehicle \;\; if \;\; p_1 \times p_2 > T$$

where p_1 and p_2 are the classifier scores for P_1 and P_2 parts of the sliding window I_V, and T is the classification threshold.

III. Performance Evaluation

The proposed method was evaluated for robustness using the LISA datasets in [8], and a comparison was done with the passive learning recogntion in [8]. Table I lists the true positive rates (TPR) and false detection rates (FDR) for the two methods. It can be seen that the proposed method gives similar or better TPR but with nearly 30-40% less false alarms.

In terms of computational complexity, the proposed method gives orders of magnitude lesser computational cost. The computational cost is defined relatively as the total number of windows that are processed for classification. Table I lists the total number of windows that are processed in each method. In

TABLE I
PERFORMANCE EVALUATION

Accuracy Evaluation				
	[8]		This Work	
	LISA-2	LISA-3	LISA-2	LISA-3
TPR	83.5%	98.1%	100%	97.5%
FDR	79.7%	45.8%	53.1%	26.7%
Computational Complexity Analysis				
	Conventional		This Work	
Num. windows	12,943,00		124,692	

conventional sliding window approaches, 10 scales of windows are assumed for an image region given by 1280x301 pixels, with a horizontal skip of 3 pixels. It can be seen that the proposed method involves 1/10-th number of windows, thereby resulting in 10x reduction in computation cost.

IV. Conclusions

In this paper, a novel technique to detect vehicles is proposed that uses the need for detecting vehicles, and the camera calibration to efficiently and robustly detect vehicles. The proposed method is shown to detect vehicles with less than 10 times the computations, while reducing false alarms by at least 30%.

Acknowledgment

We would like to thank our sponsors, particularly Korea Electronics Technology Institute (KETI) and NextChip Co. Ltd.

References

[1] R. K. Satzoda, S. Martin, L. Mihn Van, P. Gunaratne and M. M. Trivedi, "Towards automated drive analysis: A multimodal synergistic approach", *16th Intl. IEEE Conf. on Intell. Trans. Sys. - (ITSC)*, pp.1912-1916, Oct. 2013.

[2] R. K. Satzoda, P. Gunaratne, and M. Trivedi, "Drive analysis using lane semantics for data reduction in naturalistic driving studies", *IEEE Intell. Veh. Symp.*, pp. 293-298, 2014.

[3] R. K. Satzoda and M. M. Trivedi, "Drive Analysis using Vehicle Dynamics and Visual Lane Semantics", *IEEE Transactions on Intelligent Transportation Systems*, 2014 (To Appear).

[4] J. C. McCall, and M. M. Trivedi, "Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation," *IEEE Trans. on Intell. Trans. Sys.*, 7(2):20-37, 2006.

[5] R. K. Satzoda, and M. M. Trivedi, "Selective salient feature based lane analysis," *IEEE Intl. Conf. on Intell. Trans. Sys.*, pp.1906-1911, 2013.

[6] R. K. Satzoda, and M. M. Trivedi, "Vision-Based Lane Analysis: Exploration of Issues and Approaches for Embedded Realization," *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp.604-609, 2013.

[7] R. K. Satzoda, and M. M. Trivedi, "Efficient Lane and Vehicle Detection with Integrated Synergies (ELVIS)", *2014 IEEE CVPRW on Embedded Vision*, pp.694-699, 2014.

[8] S. Sivaraman and M. M. Trivedi, "A General Active-Learning Framework for On-Road Vehicle Recognition and Tracking", *IEEE Trans. on Intell. Transp. Sys.*, 11(2):267–276, 2010.

[9] A. Ramirez, E. Ohn-Bar, and M. Trivedi, "Integrating motion and appearance for overtaking vehicle detection," *IEEE Intelligent Vehicles Symposium*, pp. 96-101, 2014.