# Vehicle Detection with a Part-based Model for Complex Traffic Conditions

Ye Li*†, Bin Tian*†, Bo Li*†, Gang Xiong*‡, Fenghua Zhu*‡, Kunfeng Wang*‡
*The State Key Laboratory of Management and Control for Complex Systems,
Institute of Automation, Chinese Academy of Sciences,
Beijing, China
†Beijing Engineering Research Center of Intelligent Systems and Technology,
Institute of Automation, Chinese Academy of Sciences,
Beijing, China
‡Dongguan Research Institute of CASIA,
Cloud Computing Center, Chinese Academy of Sciences,
Dongguan, China

***Abstract*—In complex urban traffic conditions, occlusion between vehicles is a common problem which is challenging to current vehicle detection methods. In this paper, we have proposed a vehicle detection method based on a part-based model which can deal with the occlusion problem. Our method includes two steps: constructing the part-based model and detecting vehicles from traffic images. In the first step, a vehicle is divided into two parts representing an easily-occluded region around license plate and a commonly-visible region around vehicle window. Each part has low intra-class difference and is modeled by hybrid image template (HIT) with multiple types of feature descriptors in this paper. These two parts constitute our part-based model which is beneficial to vehicle detection with occlusion because the occlusion of one part has no impact on the detection of the other part. In the second step, we detect vehicles from the input image. The detection process first identifies the part candidates by using template matching and then combines the part candidates for detecting vehicles. To test our method, we have done several experiments on complex urban traffic conditions with occlusions. The experimental results show that our method can effectively cope with partial occlusion. Moreover, our method can also adapt in slight vehicle deformation and different weather conditions.**

***Index Terms*—Part-based model, vehicle occlusion, vehicle detection**

## I. INTRODUCTION

Vehicle detection task plays an important role in intelligent transportation systems, and provides vehicle information for many applications, such as driver-assistance systems [1] and traffic video surveillance systems [2][3]. Many vehicle detection methods have been proposed [4][5]. Zhang *et al.* [4] applied motion information to detect vehicles from traffic images and their method handled vehicle occlusion on intra-frame, inter-frame, and tracking levels. However, this method is not suitable for the slow-moving urban traffic because vehicles lack motion information. Ref. [5] extracted local features in three significant subregions around vehicle roof and two headlights (or taillights) to detect vehicles. This method can deal with partial occlusion. However, it may fail to detect vehicles in congested traffic conditions because the headlights are easily occluded. In short, vehicle detection methods, which consider vehicle occlusion in urban traffic conditions, are still in great demand.

Sivaraman *et al.* [6] proposed a part-based vehicle model which included two parts using strong classifiers based on Haar-like features. This method detected parts from images and then combined the detected parts for identifying vehicles. The part-based method is proven to be suitable to solve the occlusion problem. Therefore, in this paper we have proposed a vehicle detection method based on a part-based model. Our method includes two steps: constructing the part-based model and detecting vehicles from traffic images. During the model construction, the occlusion is the main concern for part selection and the parts with low intra-class difference and high inter-class difference are highly desired. In addition, a hybrid image template (HIT) [7] is applied to model the parts. This template is deformable and it integrates multiple types of descriptors including sketch, texture, color and flatness. The experimental results will show that our method effectively copes with traffic conditions with vehicle occlusion. Notably, in this paper we mainly focus on the front-view vehicle which is common in traffic video surveillance systems [2].

This paper is organized as follows: Section II presents our method including model construction and vehicle detection. Section III indicates the experiments on complex urban traffic conditions with occlusion. In Section IV, a conclusion and the future work are introduced.

## II. METHOD

Our method includes two steps: the construction of the part-based model and vehicle detection from traffic images. In the following sections, we will introduce the two steps in detail.

### *A. Model Construction*

The model construction includes three steps: part selection, part modeling, and model learning. In the following paragraphs, we will introduce the three steps in detail.

In the part selection step, the occlusion is the main concern and the parts with low intra-class difference and high inter-class difference are highly desired. In consideration of occlusion, we divide a vehicle into two parts representing an easily-occluded region and a commonly-visible region. This division

978-1-4799-0380-1/13/$31.00 ©2013 IEEE

Fig. 1. Complex traffic conditions with vehicle occlusion.

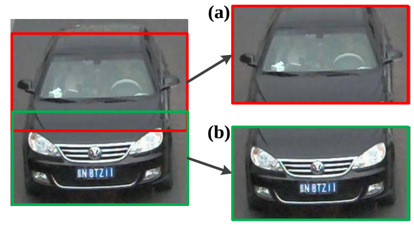


Fig. 2. A vehicle is divided into two parts in this paper.

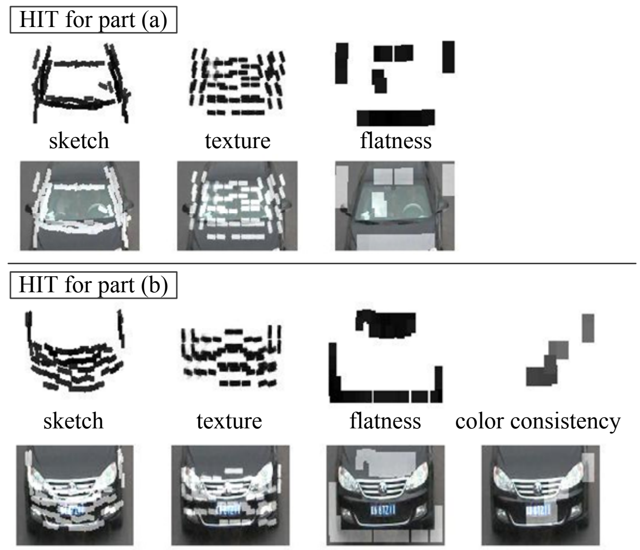


Fig. 3. Two HITs for part (a) and (b) in our model.

is helpful for vehicle detection with occlusion because the lost of one part has no impact on the detection of the other part. In the next paragraph, we will construct the easily-occluded region and the commonly-visible region.

As shown in Fig. 1, the license plates are similar in color and the regions around the license plates have rich edge information. However, the regions around the license plates are easily occluded. Thus, the region around license plate, which is shown in Fig. 2(b), is viewed as the easily-occluded part in our model. Compared with the region around license plate, the region around vehicle window is commonly visible even with severe occlusion in the complex traffic scene. In addition, the front windows among vehicles have similar shape and the roofs and hoods among vehicles have great consistency in flatness. Therefore, the region around front window, which is shown in Fig. 2(a), is used as the other part in our model. Notably, this part does not include the entire roof and hood because their sizes vary with different vehicle types. Up to this point, two parts shown in Fig. 2 are selected to constitute our model, which are called as part (a) and (b) respectively.

After part selection, we model these parts by using HITs in this paper. The HIT is deformable and it integrates four types of features for representing objects. The four types of features include local sketch, texture, color, and flatness. Therefore, the template is suitable for vehicle part modeling. In this paper, the template for each part consists of multiple image patches with various features. These image patches include sketch, texture, flatness, and color patches. The sketch patches are modeled by using Gabor wavelet elements at 16 orientations and they illustrate the edge information. The texture and color patches are modeled by using gradient histogram and color histogram respectively. The flatness patches have near-zero Gabor filter responses at 16 orientations and they illustrate the image regions with void of structures. After part modeling, our model is represented by two HITs corresponding to the two parts and their relative location and scale.

In the model learning step, we learn the model from 20 training images which are collected from real traffic images (shown in Fig. 4). The learning process includes two steps: learning two HITs and computing their relative location and scale. During the HIT learning, the image regions corresponding to part (a) and (b) are manually intercepted from the training images to learn the HITs. The HIT learning process utilizes the information projection principle proposed in [7] and meanwhile the probability distribution of the HIT is also obtained during the learning process. The learned HITs for part (a) and (b) are shown in Fig. 3. The two HITs well describe the features of part (a) and (b), and illustrate the consistency in the color of license plate and the flatness of roof and hood. After HIT learning, in order to compute the relative location and scale between the two parts, we manually label the two

Fig. 4. Training images for model learning.

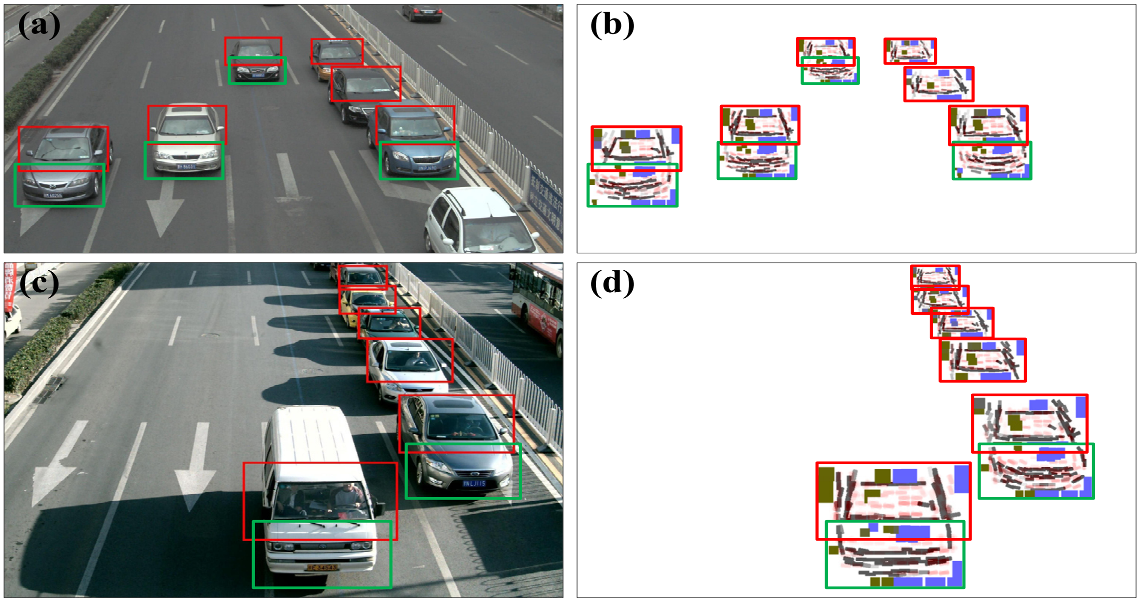


Fig. 5. The experimental results on complex traffic conditions with vehicle occlusion.

parts in training images. In this paper, the probabilities of the relative location and scale between two parts are assumed to obey the Gaussian distributions. The parameters in Gaussian distributions are learned from the labeled training images. After obtaining the above-mentioned probabilities, we define the image data likelihood probability of our model as:

$$p(I|\{p_1, p_2\}) = p(I_{\Lambda_1}, I_{\Lambda_2}) \prod_{i=1}^{2} p(I_{\Lambda_i}|p_i), \tag{1}$$

where $I$ is an input image, $p_1$ and $p_2$ are the HITs of the part (a) and (b) in our model, $I_{\Lambda_1}$ and $I_{\Lambda_2}$ are image regions corresponding to the detected candidates of part (a) and (b) respectively, $p(I_{\Lambda_1}, I_{\Lambda_2})$ denotes the probability of the relative location and scale between a pair of part candidates, $p(I_{\Lambda_i}|p_i)$ is the image data likelihood probability of $p_i$. The $p(I_{\Lambda_i}|p_i)$ is defined as:

$$p(I_{\Lambda_i}|p_i) = q(I) \prod_{j=1}^{N_i} \frac{\exp\{\lambda_{ij} f(I_{\Lambda_{ij}})\}}{Z_{ij}}, \tag{2}$$

where $N_i$ is the number of patches in $p_i$, $q(I)$ is the reference probability, $\lambda_{ij}$ is the coefficient of the $j$-th patch in $p_i$, $f(I_{\Lambda_{ij}})$ measures the distance between the image region $I_{\Lambda_{ij}}$ and the $j$-th patch, and $Z_{ij}$ is the normalizing constant. The detailed definitions of $q(I)$ and $f$ can be found in [7]. The $\lambda_{ij}$ and $Z_{ij}$ are computed during the learning process of $p_i$. According to Eq. 1, the probabilities for all training images are computed and then the vehicle detection threshold is estimated to assist the following vehicle detection.

### *B. Vehicle Detection*

Based on the learned model, we detect vehicles from traffic images. As there may exist multiple vehicles in an input traffic image, an iteration process is applied to detect these vehicles. In each iteration, a vehicle candidate with the maximal probability computed from Eq. 1 is selected and then this probability is compared with the threshold computed above. If this probability is larger than the threshold, the corresponding vehicle candidate is viewed as a detected vehicle object. After finding a vehicle object, the image region corresponding to this object is removed from the input image and the rest image will be used in next iteration for pursuing other vehicles. The iteration process stops until the maximal probability computed in current iteration is less than the threshold. In the following paragraphs, we will introduce how to pursue a vehicle candidate with the maximal probability in each iteration.

In each iteration, we first detect the candidates of the two parts in our model by using template matching procedures. Then we combine the candidates of the two parts by using their relative location and scale. After the combination, multiple vehicle candidates may be generated and their corresponding probabilities are also computed according to Eq. 1. Only the vehicle candidate with the maximal probability is selected in this iteration. Notably, an input image may include vehicles with various sizes, so we scale the image in each iteration for adapting in different vehicle sizes.

The above-mentioned template matching procedure for part candidate detection includes four steps. Take the procedure for part (a) as an example, in the first step the Gabor wavelet elements at 16 orientations are applied to filter the input image for generating edge images at these orientations. In the second step, the detected Gabor wavelet elements are locally perturbed at location and orientation for preferably matching the edge segments in the image [8]. This step outputs the modified edge images at 16 orientations. In the third step, we apply image patches in the HIT to filter the modified edge images and the input image. The color patch is used to filter the input image (using HSV color space in this paper), and other patches are used to filter the modified edge images. During the filter process, the sketch patches are allowed to be locally perturbed at location and orientation for adapting in vehicle deformation. The filter responses for all patches are outputted and used in the next step. In the fourth step, the HIT is utilized to filter the output of the third step for pursuing the part candidates from the input image.

## III. EXPERIMENTS

To verify the effectiveness of our method, we have done several experiments under complex traffic conditions with occlusion. The experimental results are shown in Fig. 5. For the vehicles with occlusion in Fig. 5, their image regions around front windows are visible while the image regions around license plates are occluded. In our method, the part (a) and (b) are separately detected from traffic images and then combined for the detection of vehicle objects. Therefore, the occlusion of part (b) has no impact on the detection of part (a) and it has a small effect on vehicle detection. The experimental results show that our method effectively deals with vehicles with occlusion. In Fig. 5, the parts (a) and (b) of the detected vehicles are shown in green and red rectangles respectively. Fig. 5 (b) and (d) indicate the detection results of sketch, texture, and flatness patches. The gray and black lines represent the sketch patches, the red lines represent the texture patches, and the small rectangles with various colors represent the flatness patches. The color intensity of each patch corresponds to the detection score of this patch. For clarity, the color patches are not drawn in Fig. 5(b) and (d). Moreover, our method can adapt in slight vehicle deformation and different weather conditions. In Fig. 5(d), the vehicles with different types are correctly detected and the vehicle shadow in the sunny day does not impact our method.

## IV. CONCLUSION

In this paper, we propose a vehicle detection method based on a part-based model for complex urban traffic conditions. The proposed method includes two steps: constructing the part-based model and detecting vehicles from traffic images. In the model construction step, we divide a vehicle into two parts which respectively represent an easily-occluded region and a commonly-visible region of the vehicle object. The division makes our method adapt in partial occlusion of the vehicle. Moreover, the two parts are represented by hybrid image templates which integrate multiple types of features including sketch, texture, color and flatness. In the vehicle detection step, the two parts are independently detected and then combined as a vehicle candidate. When one part is occluded, this occlusion has no impact on the detection of the other part and the vehicle can still be detected. The experimental results show that our method effectively deals with the complex urban traffic conditions with vehicle occlusion and adapts in slight vehicle deformation and different weather conditions.

In this paper, we mainly concentrate on the front-view vehicles. In the future work, the rear-view and side-view vehicles will be also considered. Moreover, for preferably accounting for vehicle's structural variability, we will construct hierarchical graph structure based on hybrid image template in the future work.


## ACKNOWLEDGMENT

This work is supported in part by NSFC 71232006, 61233001, 91024030, and 90920305, and CAS 2F12E05 and 2F11N07. We would like to express our sincere appreciation to professor Fei-Yue Wang for his instruction and encouragement.


## REFERENCES


[1] B.-F. Lin, Y.-M. Chan, L.-C. Fu, P.-Y. Hsiao, L.-A. Chuang, S.-S. Huang, and M.-F. Lo, "Integrating appearance and edge features for sedan vehicle detection in the blind-spot area," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 737–747, 2012.

[2] N. Buch, S. Velastin, and J. Orwell, "A Review of Computer Vision Techniques for the Analysis of Urban Traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 920–939, September 2011.

[3] J.-W. Hsieh, S.-H. Yu, Y.-S. Chen, and W.-F. Hu, "Automatic Traffic Surveillance System for Vehicle Tracking and Classification," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 2, pp. 175–187, June 2006.

[4] W. Zhang, Q. Wu, X. Yang, and X. Fang, "Multilevel Framework to Detect and Handle Vehicle Occlusion," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 161–174, March 2008.

[5] C. Wang and J.-J. J. Lien, "Automatic Vehicle Detection Using Local Features - A Statistical Approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 83–96, March 2008.

[6] S. Sivaraman and M. Trivedi, "Real-time vehicle detection using parts at intersections," in *IEEE International Conference on Intelligent Transportation Systems*, 2012, pp. 1519–1524.

[7] Z. Si and S.-C. Zhu, "Learning hybrid image templates (hit) by information projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1354–1367, 2012.

[8] Y. N. Wu, Z. Si, H. Gong, and S.-C. Zhu, "Learning Active Basis Model for Object Detection and Recognition," *Int'l J. Computer Vision*, vol. 90, pp. 198–235, 2010.