Short Paper

A Novel Approach for Vehicle Detection Using an AND–OR-Graph-Based Multiscale Model

Ye Li, Meng Joo Er, and Dayong Shen

***Abstract*—In this paper, a novel approach for detecting multiscale vehicles with time-varying vehicle features based on a multiscale AND–OR graph (AOG) model is proposed. Our approach consists of two steps, i.e., construction of a multiscale AOG model and an inference process for vehicle detection. The multiscale model uses global features to describe low-scale vehicles and local features to represent high-scale vehicles. Meanwhile, multiple appearances, such as sketch, flatness, texture, and color, are used to represent the global and local features. By virtue of the use of both global and local features as well as multiple appearances, our model is more suitable for describing multiscale vehicles in complex urban traffic conditions. Based on this multiscale model, an inference process using local features (local process) is integrated with a process using global features (global process) to detect multiscale vehicles. To evaluate the performance of our proposed method, a validation experiment, a quantitative evaluation, and a contrasting experiment are conducted. The experimental results show that our proposed approach can efficiently detect multiscale vehicles. In addition, the results also demonstrate that our approach is able to handle partial vehicle occlusion and various vehicle shapes and has great potential for real-world applications.**

***Index Terms*—AND–OR Graph (AOG), multiscale model, vehicle detection.**

I. Introduction

Vehicle detection plays an important role in intelligent transportation systems, and it revolves around providing vehicle information such as vehicle count, driver assistance, and traffic flow prediction [1]–[3]. One of the common vehicle detection methods is based on computer vision [4]–[6]. These methods are able to detect vehicles by extracting the vehicle features in traffic images [7]–[10]. However, the image size, resolution, and viewing angle (the angle between the horizontal plane and the line from a camera to a vehicle) of a vehicle are different when the vehicle is at different vehicle–camera distances. We call this phenomenon a multiscale problem, which has posed a challenge for vehicle detection due to the time-varying nature of vehicle features.

Manuscript received March 15, 2014; revised June 2, 2014 and August 29, 2014; accepted September 8, 2014. Date of publication October 7, 2014; date of current version July 31, 2015. This work was supported in part by the National Natural Science Foundation of China under Grant 71232006, Grant 61233001, Grant 91024030, and Grant 61304200 and in part by the Ministry of Transport of China under Grant 2012-364-X18-112. The Associate Editor for this paper was L. Li.

Y. Li is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences, Beijing 100864, China (e-mail: liye.cas@gmail.com).

M. J. Er is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: EMJER@ntu.edu.sg).

D. Shen is with the Research Center for Computational Experiments and Parallel Systems, National University of Defense Technology, Changsha 410073, China (e-mail: dayong.shen89@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2014.2359493

To circumvent the multiscale problem, many methods have been proposed [11]–[15]. In [11], a newly chosen tracking target (vehicle motion information) is used to update target templates for detecting multiscale vehicles. However, the vehicle motion information is found to be not suitable for congested traffic conditions and is not used in our proposed method. In [12] and [13], an active basis model (ABM) to represent and detect image objects is proposed. The method uses multiple ABMs learned at different scales to detect multiscale objects from images. This method could also be used to detect multiscale vehicles. However, the ABM method utilizes only single global edge information for object detection, and its detection accuracy is degraded in multiscale vehicle detection under complex urban traffic conditions. In [14], a hybrid image template to model and detect image objects is proposed, and multiple hybrid image templates at different scales to detect multiscale objects are constructed. However, their templates make use of only global object features and are not suitable for shape variance and occlusion of vehicles in complex urban traffic conditions. In our multitemplate method [15], three templates with multiple types of features to represent vehicles with a low, medium, and high scale are constructed to model the vehicle objects by using global vehicle features. Unfortunately, they cannot be used for vehicle detection with different shapes and severe occlusion. Although the aforementioned methods can deal with multiscale vehicles, the detection accuracy of multiscale vehicles is still not good enough for real-world applications.

To circumvent the aforementioned multiscale problem, a novel AND–OR graph (AOG) multiscale-model-based approach for detecting multiscale vehicles is proposed. The AOG proposed in [16] is a method for image object representation and detection. In an AOG, an image object is hierarchically decomposed into multiple object parts, and these parts are combined to detect image objects. The AOG is expressive for image object representation, and it also adapts to object deformation by using hierarchical and structural composition. A normal AOG is a typical local-feature-based method, which enables the AOG to process vehicle occlusion. However, the AOG may fail to detect a low-scale vehicle due to the blurring of vehicle local features. Therefore, global vehicle features are added into the AOG, and a multiscale vehicle model is proposed. Our model contains both global and local vehicle features and is suitable for the description and detection of multiscale vehicles in urban traffic conditions. In summary, the main contributions of this work are twofold. First, a multiscale AOG employing both global and local vehicle features is proposed. Second, a novel approach suitable for detecting multiscale vehicle under complex traffic conditions is developed.

This paper is organized as follows: Section II describes our approach, which includes construction of a multiscale model and a vehicle inference process. Section III presents a validation of the proposed AOG model, a quantitative evaluation of the proposed approach, and one contrasting experiment. In Section IV, conclusions are drawn, and future works are discussed.

1524-9050 © 2014 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

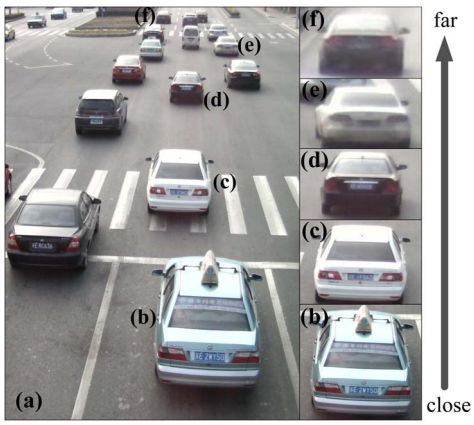

Fig. 1. Examples of multiscale vehicles with various vehicle–camera distances.

II. Proposed Approach

Our proposed approach consists of two steps, i.e., construction of a multiscale model for vehicle representation and an inference process for vehicle detection. In the following subsections, the two steps will be described in detail.

A. Construction of the Multiscale Model

Our multiscale model is based on an AOG, and its construction consists of defining the AOG's graph structure and creating the AOG's probability model. The two steps are described sequentially.

An AOG's graph structure illustrates how the objects are represented in an AOG framework. Before defining the graph structure of our AOG, the characteristics of multiscale vehicles with various vehicle–camera distances are analyzed. Fig. 1 shows examples of multiscale vehicles with various vehicle–camera distances. As shown in Fig. 1(b), when a vehicle is close to the camera, the vehicle is considered high scale and has a large image size, a high resolution, and a large viewing angle from the camera. For this high-scale vehicle, its parts are clearly visible, and the local vehicle features can be used for a description of the high-scale vehicle. Furthermore, our method focuses on urban traffic images where vehicle occlusion is common. It is well known that, under occlusion conditions, the local features perform better than global features. Hence, local features are used to describe high-scale vehicles. By local features, we mean the vehicle window and license plate. As shown in Fig. 1(b), the window and license plate are visible and distinct: even the license plate numbers are very clear. Meanwhile, among different vehicles, vehicle windows have similar shape and a flat region, whereas license plates have similar shape and color.

When a vehicle is far away from the camera, as shown in Fig. 1(f), the vehicle is low scale and has a small image size, a low resolution, and a small angle from the camera. The vehicle is blurred, and most of its details, such as the license plate and lamps, are indistinguishable due to the bad resolution. The low-scale vehicle degenerates into an approximate vehicle contour with some flat image regions. This vehicle cannot be described by local features. Therefore, global features, including vehicle contour and flat vehicle regions, are used to represent the low-scale vehicle. In short, the graph structure of AOG contains global features (vehicle contour and flat regions) for a low-scale vehicle and local features (window and license plate) for a high-scale vehicle.

In an AOG, the graph structure contains three types of nodes, including the AND, OR, and terminal nodes [16]. An AND node A_e^f represents a vehicle or a vehicle part with a decomposable structure,

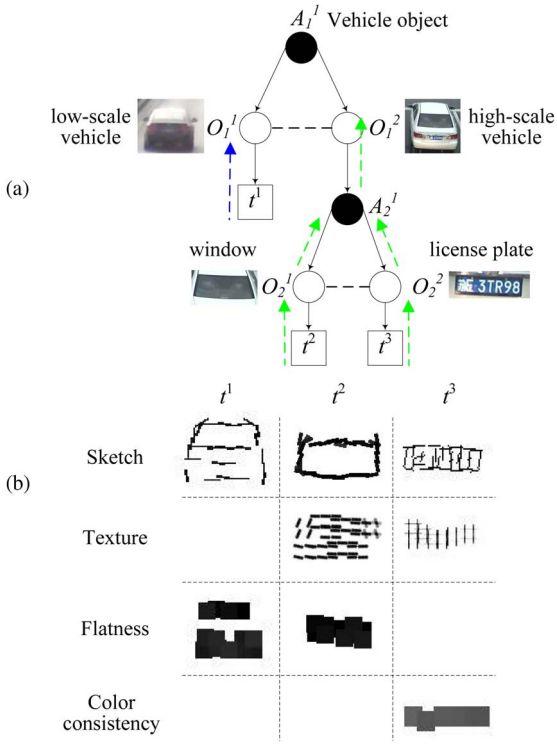

Fig. 2. Proposed multiscale AOG model. (a) AOG's graph structure. (b) Templates with multiple types of appearances for the terminal nodes.

where e denotes its layer index in all AND nodes, and f represents its sequential index in that layer. An OR node O_q^i represents a vehicle or a vehicle part with an alternative structure, where q denotes the layer index in all OR nodes, and i represents the sequential index in that layer. O_q^j means another OR node. A terminal node t^u represents a vehicle or vehicle part directly detected from the image, where u represents the sequential index of t^u. As shown in Fig. 2(a), an AOG containing both a low-scale vehicle and a high-scale vehicle is constructed. In our AOG, the AND node A_1^1, OR node O_1^1, and OR node O_1^2 represent the vehicle object, low-scale vehicle object, and high-scale vehicle object, respectively. The term O_1^1 denotes a link to terminal node t^1, and t^1 achieves direct detection of a low-scale vehicle based on global features from traffic images. The term O_1^2 denotes a link to A_2^1, and A_2^1 is decomposed into O_2^1 and O_2^2, representing the vehicle window and license plate, respectively. The terms O_2^1 and O_2^2 denotes links to terminal node t^2 and t^3, respectively, where t^2 and t^3 are applied to detect the vehicle window and license plate from real traffic images. In summary, the graph structure of our AOG illustrates an ability to represent multiscale vehicles. The AOG's probability model for assisting vehicle detection is described sequentially.

Based on the constructed multiscale AOG, vehicle detection can be accomplished by using the following probability model:

$$p(V|I) \approx \prod_{\langle O_q^i, O_q^j \rangle \subset E_O^V} p\left(O_q^i, O_q^j\right) \prod_{t^u \subset t^V} p(I|t^u) \tag{1}$$

where $V = \{A^V, O^V, t^V, E^V\}$ is a vehicle instance detected from a traffic image I, and A^V, O^V, t^V, and E^V are instance sets of the AND node, OR node, terminal node, and edge in V, respectively. The term E_O^V is an instance set of the edges between the OR nodes in E^V, $p(V|I)$ denotes the posterior probability of vehicle instance V, and $p(O_q^i, O_q^j)$ denotes the probability of integrating the instances of O_q^i and O_q^j during the vehicle detection process. For a detailed explanation

of (1), interested readers can refer to [17]. In this paper, instances of O_q^i and O_q^j are combined by using their relative location and scale. Here, the probability distributions of their relative location and scale are assumed to be Gaussian distributions. Therefore, the term $p(O_q^i, O_q^j)$ contains two Gaussian probability distributions.

In this paper, we utilize hybrid image templates [14] to mathematically model t^u. In (1), $p(I|t^u)$ is the matching probability between the template of t^u and a local image region in I. A general hybrid image template consists of multiple sketch patches, texture patches, flatness patches, and color patches, which describe edge, texture, flatness, and color appearances of an object (or object part), respectively. In our paper, the template of t^1 comprises multiple sketch patches and flatness patches, which describe the contour and flatness region of a low-scale vehicle object, respectively. The template of t^2 consists of multiple sketch patches, flatness patches, and texture patches modeling the shape, flatness, and texture appearance of the high-scale vehicle window, respectively. The template of t^3 contains multiple sketch patches, color patches, and texture patches describing the shape, color, and texture appearance of the high-scale license plate, respectively. In addition, in this paper, the mathematical representation of the sketch patch is a Gabor wavelet element with an orientation $x * \pi/16$ ($x = 0$, or 1, . . ., or 15). The mathematical representation of a texture patch is a gradient histogram of the Gabor filter responses in 16 orientations. The mathematical representation of a flatness patch is a sum of the Gabor filter responses in 16 orientations. The mathematical representation of a color patch is a color histogram in the hue–saturation–value color space.

The graph structure and probability model constitute our multiscale AOG model. In the following subsection, we will introduce an inference process for vehicle detection based on our proposed multiscale AOG model.

B. Training Process for Learning Parameters

The training process includes the following two steps: 1) preparing training images and 2) learning parameters. In step 1, training images are first collected from real traffic conditions, and these images include 60 examples with various vehicle–camera distances. Fig. 3 shows part of the training images. Then, all nodes of our AOG are annotated in training images. Based on these annotated images, AOG's parameters are learned in step 2, and these parameters include $p(O_q^i, O_q^j)$ in (1) and model parameters for terminal nodes $\{t^u\}$. The term $p(O_q^i, O_q^j)$ contains two Gaussian probability distributions of the relative location and scale between O_q^i and O_q^j, and the parameters in these distributions are learned from the training images by using maximum likelihood estimation. Model parameters for $\{t^u\}$ include $p(I|t^u)$ and patches for $\{t^u\}$, and they are learned from the training images by using the information projection principle [14]. Fig. 2(b)shows trained patches for $\{t^u\}$. During the learning process, $p(I|t^u)$ is computed as follows:

$$p(I|t^u) = q(I) \prod_{w=1}^{K_u} \frac{e^{\lambda_{uw} r(I_{uw})}}{z_{uw}} \tag{2}$$

where $q(I)$ is a reference distribution, K_u is the number of patches in t^u, λ_{uw} is a coefficient of the wth patch in t^u, z_{uw} is a normalization constant, and r is a distance function measuring the similarity between the image region I_{uw} and the wth patch. For the detailed definition of $p(I|t^u)$, interested readers can refer to [14]. Therefore, (1) can be written as follows:

$$p(V|I) \approx \prod_{\langle O_q^i, O_q^j \rangle \subset E_O^V} p\left(O_q^i, O_q^j\right) \prod_{t^u \subset t^V} \prod_{w=1}^{K_u} \frac{e^{\lambda_{uw} r(I_{uw})}}{z_{uw}} \tag{3}$$

where $p(V|I)$ is used as the vehicle detection score.

Fig. 3. Examples of the training image.

C. Inference Process for Detecting Vehicles

Based on the constructed AOG, a vehicle inference is utilized to detect vehicles from traffic images. The vehicle inference contains two sequential steps, i.e., s_1 and s_2. In the s_1 step, both a global-feature-based detection process (global process) and a local-feature-based detection process (local process) are carried out, whereas in the s_2 step, the results in the global and local processes are integrated.

The s_1 step is shown in Fig. 2, where the blue and green dashed arrows represent the global and local processes, respectively. The global process directly uses the template of t^1 to filter the traffic image and detects instances of O_1^1. During the global process, the current image is scaled, so that the global process is adapted to vehicles with various scales. Note that a high-scale vehicle is blurred by the scaling so that it matches with the low-scale template. The local process first utilizes templates of t^2 and t^3 to detect the window and license-plate instances and then combines them by using their relative locations and scales to detect instances of O_1^2 from the current image. During the local process, we also scale the image to detect multiscale vehicles. Note that the vehicle window may be still distinguishable when a vehicle is in low scale, which indicates that the local process can also detect some low-scale vehicles.

To test the performances of the global and local processes, a total of 60 vehicle images, including 20 high-scale vehicles, 20 medium-scale vehicles, and 20 low-scale vehicles, were selected. The performances of the global and local processes are separately evaluated using these images. It turns out that a vehicle can be correctly detected only when its bounding box overlaps more than 60% of the region of the ground truth. Fig. 4 shows some results of the two processes, where a tick means that the vehicle is correctly detected and a cross indicates that the result is wrong. In Fig. 4, images at the left of the dashed line show results of the global process, whereas images at the right of the dashed line show results of the local process. The first, second, and third rows show results of low-, medium-, and high-scale vehicles, respectively. The red, blue, and yellow rectangles denote bounding boxes of the detected global feature, window, and license plate, respectively. For clarity, the number of detected vehicles is shown in Fig. 5. Note that, in Fig. 5, the blue dashed line and red solid line show the number of detected vehicles in the local and global processes, respectively. The results show that the global process performs better when a vehicle is farther away from the camera, which indicates that our global features are suitable for the low-scale vehicles. In contrast, the local process has better performance when the vehicle is closer to the camera. Thus, local features are well suited to high-scale vehicles. Moreover, both processes have good performances on the medium-scale vehicles, and their results could complement each other (see the second row in Fig. 4). Above all, the integration of the two processes can take full advantage of both local and global features.

In the s_1 step, instances of O_1^1 and O_1^2 in the current step, as well as their detection scores, are produced. In the s_2 step, instances of O_1^1 and O_1^2 are combined by using their relative location and scale to obtain instances of A_1^1. In the s_2 step, vehicle candidates with both low and high scales are produced, and their detection scores are calculated by using (3). Finally, in the vehicle inference, vehicle candidates with detection scores are extracted, and these scores are compared with a threshold T. The threshold T is estimated from training images by

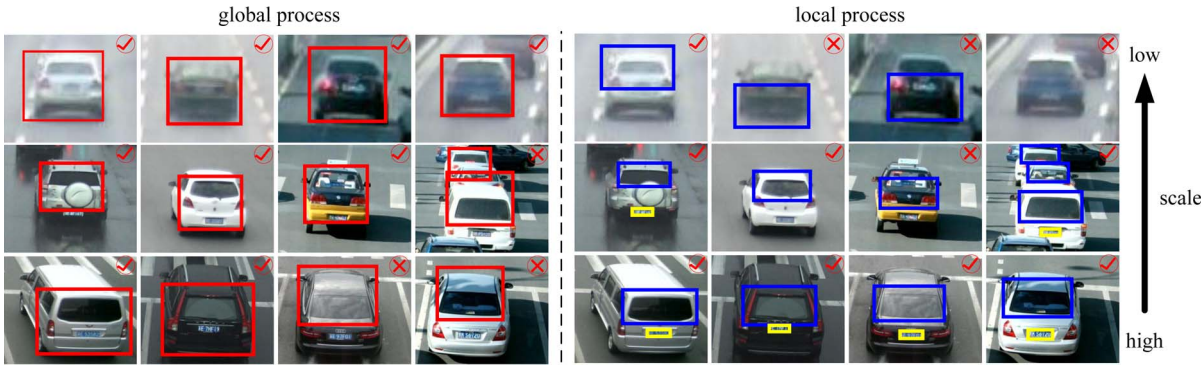

Fig. 4. Performance test for the global and local processes.

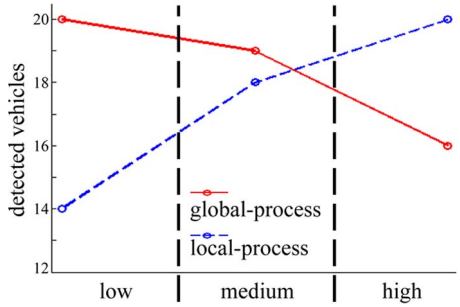

Fig. 5. Quantitative results of global and local processes.

using (3). If a score is greater than T, then that vehicle candidate is seen as a detected vehicle.

III. Experimental Results

In this paper, a validation experiment, a quantitative evaluation, and a contrasting experiment are conducted to evaluate the performance of our proposed approach. The following paragraphs describe the three experiments in detail.

A. Validation of the AOG

In our previous works in [15], a multitemplate method to detect multiscale vehicles is proposed, which provides a common solution for the multiscale problem. That method separately uses three templates to model vehicle objects with low, medium, and high scales. Each template has multiple types of vehicle appearances, including sketch, texture, flatness, and color. However, the AOG is not used in [15]. In contrast, our method in this paper utilizes an AOG to represent multiscale vehicles. As foreshadowed, an AOG is a typical part-based method, which decomposes vehicle objects into parts for detection. Therefore, our method is superior to the multitemplate method because our method is capable of dealing with vehicle occlusion. Another benefit of AOG is that it has been proven to be an effective framework for visual knowledge representation and object detection. The AOG is well suited for detecting vehicles from traffic images. Moreover, the AOG is highly extensible, which makes our method suitable for real-world application.

Fig. 6 shows the comparative experiment with and without the AOG, where (a) is a test image and (b) and (c) represent the detection results of the multitemplate method and the proposed method, respectively. The red rectangles indicate the bounding boxes of the detected vehicles. Note that the red ellipse in Fig. 6(a) shows a severe vehicle occlusion. The multitemplate method models the vehicle by using the features of the entire vehicle (global features), and thus, it fails

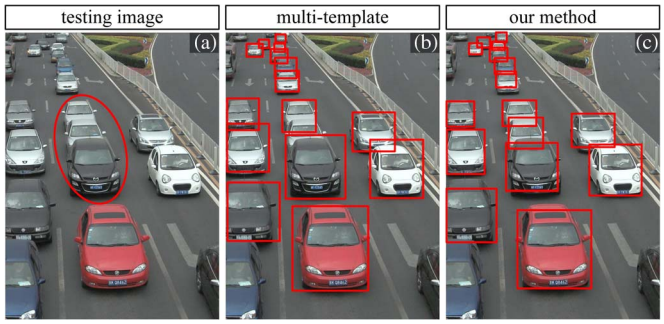

Fig. 6. Results in the validation experiment. (a) Test image. (b) Results of the multitemplate method without AOG. (c) Results of our method.

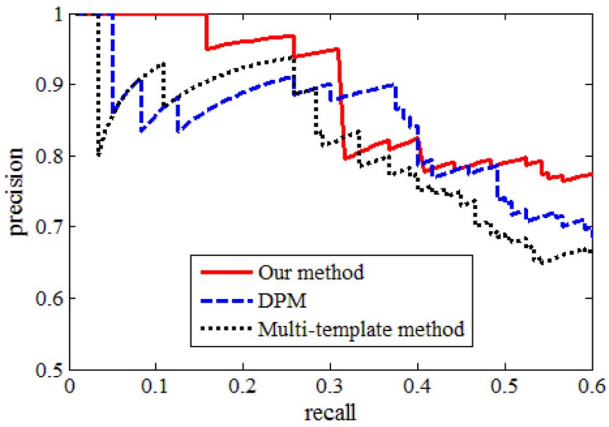

Fig. 7. Precision–recall curves of different methods. The red solid line, blue dashed line, and black dot line show the precision–recall curves of our method, DPM, and the multitemplate method, respectively.

to handle this occlusion. Compared with the multitemplate method, our method is able to deal with vehicle occlusion by virtue of the use of AOG. Moreover, comparative precision–recall curves of the two methods are shown in Fig. 7, which illustrates that our proposed method has higher vehicle detection accuracy than the multitemplate method on the testing set including multiple scales, various vehicle shapes, and vehicle occlusion. In Fig. 7, a precision–recall curve for a deformable part model (DPM) is also shown. The DPM is a method for image object representation and detection, and it performs well when handling multiscale vehicles and occlusion. These curves show that our method is better than the DPM when processing multiscale vehicles. We will introduce a comparative experiment of the DPM and our method in Section III-C in detail.

Fig. 8. Part of negative examples used in this paper.

B. Quantitative Evaluation

In order to quantitatively evaluate our proposed method, we investigate our precision–recall curve on a set of test images. The test image set includes 120 positive examples and 120 negative examples. The positive examples are collected from real urban traffic images containing vehicles with various scales. The positive examples also include vehicles with partial occlusion and various vehicle shapes. The negative examples are images with nonvehicle objects, and these examples are collected from the Pattern Analysis, Statistical Modeling and Computational Learning (PASCAL) Visual Object Classes 2007 data set (http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2007/). Fig. 8 shows a part of negative examples. We perform our method on these test images and record all detection scores and bounding boxes of the detected vehicle instances. A vehicle is correctly detected only when its bounding box overlaps more than 60% of the region of the ground truth. Based on the detection scores and bounding boxes, we draw a precision–recall curve by thresholding these scores at different points. Note that the precision and recall are defined as (true positive)/(true positive + false positive) and (true positive)/(true positive + false negative). As shown in Fig. 7, the red solid line represents the precision–recall curve of our proposed method. Moreover, we also test a multiscale method based on multiple templates [15], and its precision–recall curve is shown with a black dot line in Fig. 7. Comparing the two curves, it is clear that our method is superior to the multitemplate method.

Several detection results in the quantitative experiment are shown in Figs. 9 and 10. Fig. 10 indicates detection results of a test image with occlusion, and the results show that our method can effectively deal with the vehicle occlusion. Fig. 9(a)–(c) shows three test images with multiscale vehicles, and Fig. 9(d)–(f) indicates the bounding boxes of the detected vehicles. Fig. 9(g)–(l) shows the detected patches in the global and local processes, where a darker patch has the higher matching score. The gray rectangles, red lines, and gray lines denote the detected flatness (or color) patches, texture patches, and sketch patches, respectively. These results show that our method can effectively deal with the multiscale vehicles. Moreover, the test images in Fig. 9(a)–(c) contain different weather conditions, including cloudy, sunny, and rainy weather. It turns out that the vehicle shadow in a sunny condition and the vehicle reflection in a rainy condition have slight impact on our method. In addition, the test images in Fig. 9 include vehicles with different vehicle shapes, such as a car, a sport-utility vehicle, and a minibus. The results clearly show that our proposed method can adapt to various vehicle shapes.

C. Contrasting Experiment

In addition to the quantitative evaluation, a contrasting experiment is also conducted. In the contrasting experiment, our proposed method is compared with a method based on mixtures of multiscale DPM [18]. The DPM represents vehicle objects by using a coarse root filter plus a set of higher resolution part filters. Meanwhile, the features similar to the histogram of oriented gradients (HOG) are used in these filters. Thus, the DPM integrates visual appearances at multiple scales, and it detects multiscale vehicles by calculating an image pyramid via repeated smoothing and subsampling. Fig. 11 shows the comparative results. Fig. 11(a) is for a test image, and Fig. 11(b) and (d) shows

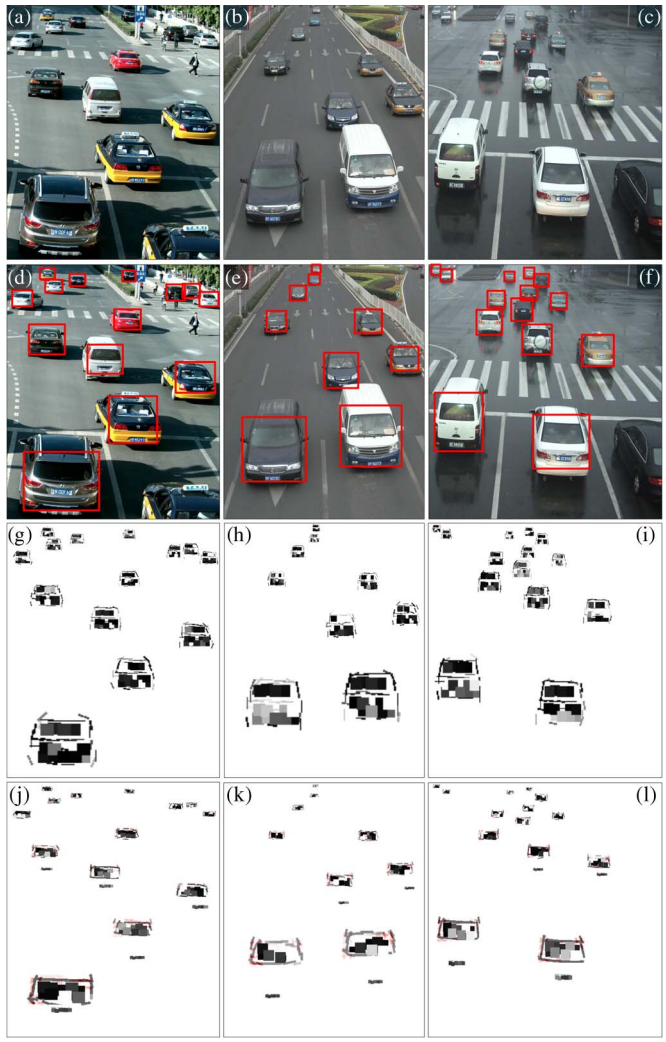

Fig. 9. Results with multiscale vehicles in the quantitative experiment. (a)–(c) Test images. (d)–(f) Our detection results. (g)–(l) Detected patches in the global and local processes.

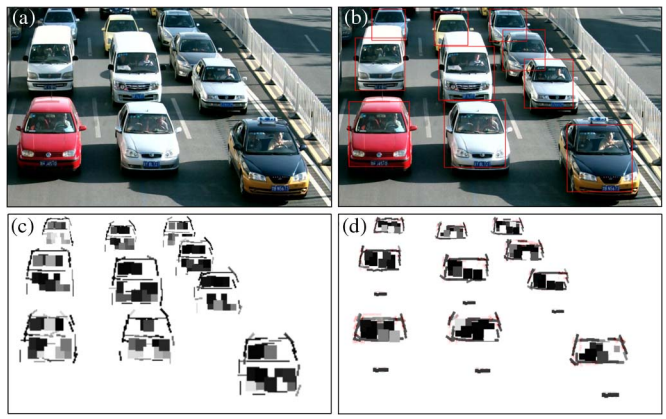

Fig. 10. Results with vehicle occlusion in the quantitative evaluation. (a) Test image. (b) Detection results. The red rectangles show boundary boxes of the detected vehicles. (c) and (d) Detected patches in the global and local processes.

vehicle detection results of the DPM and our proposed method, respectively. Fig. 11(c) is for detection results of the part filters in DPM, and Fig. 11(e)–(f) shows the detected patches in our global and local processes, respectively. As shown in the red ellipse in Fig. 11(a),

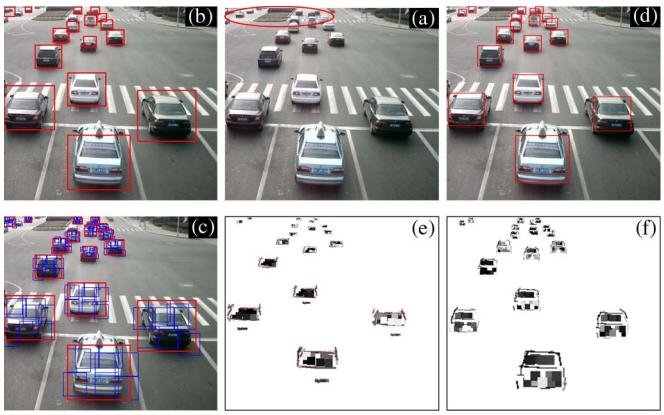

Fig. 11. Results of one comparative experiment. (a) Test image. Vehicle detection results of (b) the DPM and (d) our method. (c) Detection results of particle filters in DPM. (e)–(f) Detected patches in the global and local processes of our proposed method.

there are false and missed detections in the DPM. In contrast, our method is able to cope with these multiscale vehicles, which benefit from the application of our multiscale AOG model. Furthermore, the comparative precision–recall curves of the DPM and our method are shown in Fig. 7, which also illustrates that our method has higher vehicle detection accuracy than the DPM on the testing set including multiple scales, various vehicle shapes, and vehicle occlusion.

IV. Conclusion

In this paper, a multiscale model based on an AOG to circumvent the multiscale problem is proposed. In our multiscale AOG model, the vehicles are divided into two types, i.e., a high-scale vehicle with a close vehicle–camera distance and clear vehicle details and a low-scale vehicle with a long vehicle–camera distance and blurred vehicle appearances. The high-scale vehicle is represented by local vehicle features, including the vehicle window and license plate, and thus, it is modeled by using sketch, texture, flatness, and color. The low-scale vehicle is represented by global vehicle features, including vehicle outline and flat regions, and therefore, it is modeled by using sketch and flatness. The use of global and local features and multiple appearances makes our model more suitable for describing multiscale vehicles in complex urban traffic conditions. Moreover, the global process is integrated with the local process, and a two-step inference process is developed to detect vehicles based on the proposed AOG model. Experimental results show that our method can effectively and efficiently deal with multiscale vehicles, which benefit from the combination of AOG and the global and local processes. Moreover, our proposed method can be adapted to different vehicle shapes and vehicle occlusion.

References

[1] F.-Y. Wang, "Parallel control and management for intelligent transportation systems: Concepts, architectures, applications," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 630–638, Sep. 2010.

[2] L. Unzueta *et al.*, "Adaptive multicue background subtraction for robust vehicle counting and classification," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 527–540, Jun. 2012.

[3] S. Taghvaeeyan and R. Rajamani, "Portable roadside sensors for vehicle counting, classification, speed measurement," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 73–83, Feb. 2014.

[4] N. Buch, S. Velastin, and J. Orwell, "A review of computer vision techniques for the analysis of urban traffic," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 3, pp. 920–939, Sep. 2011.

[5] S. Sivaraman and M. Trivedi, "Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, behavior analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1773–1795, Dec. 2013.

[6] R. Wang, L. Zhang, K. Xiao, R. Sun, and L. Cui, "EasiSee: Real-time vehicle classification and counting via low-cost collaborative sensing," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 414–424, Feb. 2014.

[7] B.-F. Lin *et al.*, "Integrating appearance and edge features for sedan vehicle detection in the blind-spot area," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 737–747, Jun. 2012.

[8] J.-W. Hsieh, L.-C. Chen, and D.-Y. Chen, "Symmetrical SURF and its applications to vehicle detection and vehicle make and model recognition," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 6–20, Feb. 2014.

[9] S. Sivaraman and M. Trivedi, "Integrated lane and vehicle detection, localization, tracking: A synergistic approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 906–917, Jun. 2013.

[10] A. Makris, M. Perrollaz, and C. Laugier, "Probabilistic integration of intensity and depth information for part-based vehicle detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1896–1906, Dec. 2013.

[11] X. Mei and H. Ling, "Robust visual tracking and vehicle classification via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 11, pp. 2259–2272, Nov. 2011.

[12] Y. N. Wu, Z. Si, H. Gong, and S.-C. Zhu, "Learning active basis model for object detection and recognition," *Int. J. Comput. Vis.*, vol. 90, no. 2, pp. 198–235, Nov. 2010.

[13] Z. Si, H. Gong, S.-C. Zhu, and Y. N. Wu, "Learning active basis models by EM-type algorithms," *Stat. Sci.*, vol. 25, no. 4, pp. 458–475, Nov. 2010.

[14] Z. Si and S.-C. Zhu, "Learning hybrid image templates (HIT) by information projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 7, pp. 1354–1367, Jul. 2012.

[15] Y. Li *et al.*, "A multi-scale model integrating multiple features for vehicle detection," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, 2013, pp. 399–403.

[16] Z. Si and S.-C. Zhu, "Learning AND-OR templates for object recognition and detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2189–2205, Sep. 2013.

[17] Y. Li, B. Li, B. Tian, and Q. Yao, "Vehicle detection based on the and-or graph for congested traffic conditions," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 984–993, Jun. 2013.

[18] P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.