Transportation Research Part C 56 (2015) 80–93

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Vehicle detection grammars with partial occlusion handling for traffic surveillance

Bin Tian [a,*], Ming Tang [b], Fei-Yue Wang [a]

[a] State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 31 March 2014
Received in revised form 13 December 2014
Accepted 22 February 2015

Keywords:
Computer vision
Grammar model
Occlusion handling
Part-based object detection
Vehicle detection

ABSTRACT

Traffic surveillance is an important topic in intelligent transportation systems (ITS). Robust vehicle detection is one challenging problem for complex traffic surveillance. In this paper, we propose an efficient vehicle detection method by designing vehicle detection grammars and handling partial occlusion. The grammar model is implemented by novel detection grammars, including structure, deformation and pairwise SVM grammars. First, the vehicle is divided into its constitute parts, called *semantic parts*, which can represent the vehicle effectively. To increase the robustness of part detection, the semantic parts are represented by their detection score maps. The semantic parts are further divided into *sub-parts* automatically. The two-layer division of the vehicle is modeled into a grammar model. Then, the grammar model is trained by a designed training procedure to get ideal grammar parameters, including appearance models and grammar productions. After that, vehicle detection is executed by a designed detection procedure with respect to the grammar model. Finally, the issue of vehicle occlusion is handled by designing and training specific grammars. The strategy adopted by our method is first to divide the vehicle into the semantic parts and sub-parts, then to train the grammar productions for semantic parts and sub-parts by introducing novel *pairwise SVM grammars* and finally to detect the vehicle by applying the trained grammars. Experiments in practical urban scenarios are carried out for complex traffic surveillance. It can be shown that our method adapts to partial occlusion and various challenging cases.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

With rapid development of urbanization, traffic congestion, incident and violation pose great challenges on traffic management systems. Video-based surveillance has attracted lots of attention in intelligent transportation systems (ITS). Reliable and robust vehicle detection is a fundamental component for traffic surveillance. Different vehicle appearances and poses make it difficult to train a unified detection model. Under complex urban environments, bad weathers, illumination change and poor/strong lighting conditions will degrade the detection performance dramatically. Especially for traffic congestion, vehicles are occluded by each other so that separate vehicles will easily merge into a single vehicle.

Moving object detection is widely used for the vehicle detection in ITS. These methods can be classified into background modeling, frame differencing and optical flow (Stauffer and Grimson, 1999). They can handle illumination change and apply

* Corresponding author.
E-mail addresses: bin.tian@ia.ac.cn (B. Tian), tangm@nlpr.ia.ac.cn (M. Tang), feiyue@ieee.org (F.-Y. Wang).

http://dx.doi.org/10.1016/j.trc.2015.02.020
0968-090X/© 2015 Elsevier Ltd. All rights reserved.

to multi-mode background. There are some drawbacks in that they are unable to detect a stationary vehicle and the detected moving object is not necessarily a vehicle. Therefore, lots of research utilizes the visual features of the vehicle to detect it in a still image (Sun et al., 2002; Cheng et al., 2012). Features such as gabor, color, edge and corner are usually used to represent the vehicle. Then, they are fed into deterministic classifier and generative model to recognize vehicles. Besides, Cheon et al. (2012) employed a two-step method including hypothesis generation and hypothesis verification to locate the vehicle. This method works well during sunny daytime, but may fail during poor lighting conditions such as nighttime.

Many recent studies on part-based model have been developed to recognize vehicles. According to a human cognitive study, a vehicle is considered to be composed of a window, roof, wheel and other parts (Lam et al., 2007). These parts are usually learned and detected by using their appearance, edge and shape features (Lin et al., 2009; Lin et al., 2012). After part detection, the spatial relationship, motion cue and multiple models are usually used to detect vehicles. Winn and Shotton (2006) decomposed an object into several local regions to detect them. The relationships between them are used to improve detection performance by introducing a layout conditional random field (LCRF) model. Hoiem et al. (2007) further expanded this method by using 3-D models to label the learning samples. In addition to specifying the constitute parts artificially, Felzenszwalb et al. (2010) proposed a discriminatively trained deformable part model which can select the parts and learn their model automatically. This model was designed for general object detection instead of vehicle detection. Tehrani Niknejad et al. (2012) employed this model for on-road vehicle detection, which is composed of five components, including the front, back, side, front truncated and back truncated. Each component contains a root filter and six part filters which are learned by using a latent support vector machine (LSVM) and histogram of oriented gradient (HOG) features. This method constructed a model for each vehicle pose according to its aspect ratio, which can only consider limited vehicle poses. It was applied to in-vehicle cameras instead of a fix camera for traffic surveillance. Different from the aforementioned methods, our method considers the characteristics of the traffic scenario and vehicle to cope with both the vehicle deformation and partial occlusion. It also introduces score features for vehicle parts to produce more robust detections by decreasing false alarms.

The issue of object occlusion has attracted much research interest in recent years. Due to high density of vehicles and low camera angle for monitoring urban traffic, severe occlusion will lead to false detections of occluded vehicles. Common occlusion handling methods utilize the visible information of occluded objects while ignoring the features of occluded parts (Pang et al., 2004; Huang and Liao, 2004; Kanhere et al., 2005; Pang et al., 2007; Song and Nevatia, 2007; Zhang et al., 2008; Kanhere and Birchfield, 2008; Wang et al., 2009; Shu et al., 2012; Tang et al., 2013). Huang and Liao (2004) utilized vehicle motion vectors to detect occluded vehicles. Different parts of one vehicle have the same motion vectors, while the parts from different vehicles do not share the same motion vectors. Based on this observation, motion vectors of local areas are computed first. If neighboring areas have similar motion vectors, they belong to the same vehicle. Otherwise, they belong to different vehicles. In Zhang et al. (2008), considered that occlusion is caused by information loss in the projection of a 3-D scene onto a 2-D image plane, and the key point of handling occlusion is estimation of the lost information. Their framework consists of three levels for occlusion handling: the intraframe, interframe, and tracking levels, which are sequentially implemented. These methods can deal with moving vehicles, but it might fail for stopping ones, e.g. queuing vehicles in congested traffic. Kanhere et al. (2005) and Kanhere and Birchfield (2008) presented a method for segmenting and tracking vehicles on highways using a camera that was relatively low to ground. In their framework, feature points are extracted first. Then, distribution of these points is estimated to recognize occluded vehicles. In Shu et al. (2012), used a part-based model to detect person. The subset of parts are selected to maximize detection probability. This significantly improved detection performance in the case of occlusion. In Tang et al. (2013), to address occlusion problem, Tang et al. proposed a double-person detector that allows to predict bounding boxes of two people even when they occlude each other by 50% or more. This method learns double-person detectors by directly training occluded samples. The methods in Shu et al. (2012) and Tang et al. (2013) provide innovative ideas of handling object occlusion. In this paper, we will handle vehicle occlusion in practical traffic scenes, by first detecting visible parts and then inferring the occluded parts.

In this paper, we present a vehicle detection method by designing vehicle detection grammars and handling partial occlusion. The grammar model is implemented by novel detection grammars, including structure, deformation and pairwise SVM grammars. First, the vehicle is divided into its constitute parts, called *semantic parts*, by considering the characteristics of traffic scenario and vehicle. To increase the robustness of the part detection, the semantic parts are represented by their detection score maps. The semantic parts are further divided into *sub-parts* automatically. The part division procedure results in a two-layer architecture, as shown in Fig. 1. The two-layer division of the vehicle is modeled into a grammar model. Then, the grammar model is trained by a designed training procedure to get ideal grammar parameters, including appearance models and grammar productions. After that, vehicle detection is executed by a designed detection procedure with respect to the trained grammar model. Finally, the issue of vehicle occlusion is handled by designing and training specific grammars. The strategy adopted by our method is first to divide the vehicle into semantic parts and sub-parts, then to train the grammar productions for the semantic parts and sub-parts by introducing novel pairwise SVM grammars and finally to detect the vehicle by applying the trained grammars. The main contributions of this work lie in: (1) the characteristics of traffic scenario and vehicle are analyzed to derive efficient part division for better detection and occlusion handling; (2) novel vehicle detection grammars are realized by employing pairwise SVMs; (3) partial occlusion is handled with the proposed grammar model. The details of the method will be discussed in later sections.

The remainder of this paper is organized as follows. The base-line algorithms, DPM (Felzenszwalb et al., 2010) and object detection grammars (Girshick et al., 2011), are reviewed in Section 2. In Section 3, we present the vehicle detection

Download English Version:

https://daneshyari.com/en/article/6936710

Download Persian Version:

https://daneshyari.com/article/6936710

Daneshyari.com