J. Vis. Commun. Image R. 37 (2016) 25–31

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

APPOS: An adaptive partial occlusion segmentation method for multiple vehicles tracking

Na Zhao [a], Yingjie Xia [a,b,*], Chao Xu [a], Xingmin Shi [a], Yuncai Liu [a]

[a] *Intelligent Transportation and Information Security Lab, Hangzhou Normal University, Hangzhou, Zhejiang, China*
[b] *College of Computer Science, Zhejiang University, Hangzhou, Zhejiang, China*

ARTICLE INFO

Article history:
Received 14 December 2014
Accepted 20 April 2015
Available online 27 April 2015

Keywords:
Multiple vehicles tracking
Partial occlusion
Segmentation
Optical flow
Mean square error
Location
Foreground extraction
Occlusion detection

ABSTRACT

In traffic surveillance videos, it is common that the vehicles are occluded partially by each other. Such kind of occlusion situation is a challengeable task in multiple vehicles tracking. Various solutions in dealing with the occlusion for vehicles tracking have been proposed in many literatures. However, most of them are specialized on one tracking method and cannot flexibly adapt to the others. In this paper, we propose an adaptive partial occlusion segmentation method (APPOS) for multiple vehicles tracking. In this method, the occlusion detection process is firstly conducted to discover the occlusion. After that, the candidate regions of the respective occluded vehicles are roughly evaluated by the contour's optical flow. Finally, the line scanning which uses color contrast among regions is adopted to accurately locate the vehicles. We evaluate the effectiveness and accuracy of APPOS by the experiments on practical and simulating videos.

1. Introduction

Video-based multiple vehicles tracking is essential for many vision-based intelligent transportation systems (ITS) applications. Although lots of tracking methods have been studied, there are still some challenges. E.g., multiple vehicles tracking will be missed when the occlusion happens by vehicle overlapping or connecting. It is common that the partial occlusion can cause the features of vehicle, such as sizes, textures, and colors, to be changed in the viewpoint of camera. In the vehicle occlusion, there are three kinds of objects to block the vehicle, including other moving objects, background scene objects, and other vehicles. These objects make the tracking trajectories insignificant when the tracked vehicles are not the same one. Therefore, a reliable partial occlusion segmentation method for multiple vehicles tracking is urgently required.

Occlusion segmentation is thought as the research problem in visual algorithms. The moving vehicle firstly exhibits many spatial configurations in a viewpoint. Furthermore, the appearance of the blocked part is replaced by the other object's appearance. Therefore, when one vehicle is occluded other objects, their foregrounds will group together and it will be a challenge to accurately classify the pixels in the foreground. The solution for the occlusion segmentation needs to overcome three difficulties: (i) in the case that the normal vehicle is detected as the occluded vehicle, the result of segmentation can not affect the tracking result; (ii) choose effective features to build the model for the vehicle tracking in lone-term occlusion; and (iii) use a reliable method to locate the vehicle in the pixel level.

Many studies have been conducted to segment occlusion for improving tracking methods by overcoming these difficulties. Harville et al. [1] proposed a solution based on the depth of tracked targets. When a vehicle occludes other vehicles, the vehicle which is closer to camera will have smaller depth. Xia et al. [2] proposed a vehicles overtaking detection method using RGB-D camera to capture the depth data. The depth data is utilized to analyze the posture change of vehicles. Isard and MacCormick [3] proposed a model to continuously predict the location for occlusion segmentation. Comaniciu et al. [4] proposed a real-time tracking method using the mean-shift algorithm on the object appearance to solve the partial occlusion problem. However, all these related studies are specialized on one tracking method and cannot flexibly adapt to the others.

In this paper, we propose an adaptive partial occlusion segmentation method (APPOS) for multiple vehicles tracking. In this method, the occlusion detection process is firstly conducted to discover the occlusion. And the contour's optical flow and line scanning are used to locate the regions of the occluded vehicles. The

* Corresponding author at: Intelligent Transportation and Information Security Lab, Hangzhou Normal University, Hangzhou, Zhejiang, China.
E-mail address: xiayingjie@zju.edu.cn (Y. Xia).

http://dx.doi.org/10.1016/j.jvcir.2015.04.011
1047-3203/

main contributions of APPOS are summarized as: (i) we propose an adaptive partial occlusion segmentation method which can be adapted to most kinds of tracking methods by using the video sequences and vehicle trajectories; (ii) the occlusion segmentation process will not affect the tracking results even if the occlusion is detected incorrectly; (iii) The candidate region evaluating process can not only locate the occluded vehicle, but also get back the tracking-lost vehicle.

The rest of paper is structured as follows. Related work is discussed in Section 2. The Section 3 presents the method of classifying the occlusion situation. Our method is described in Section 4. Experimental results and analysis are given in Section 4. Section 5 draws the conclusions and discusses the future work.

2. Related work

2.1. Vehicle tracking

In last few years, vehicle tracking attracts a lot of researchers and various efficient approaches have been proposed. These methods can be categorized into two kinds. The first leverages appearance features to build the model, which is utilized to evaluate the similarities of region appearance between adjacent two continuous frames. Liu et al. proposed a tracking method based on histograms for local sparse appearance representation [5]. Another kind of approach is generally based on extracting the motion of vehicles. For this kind of method, a point, which refers to the position of a vehicle in the motion, is used as the representation of the vehicle. The association among the moving points, extracted from continuous adjacent frames, are evaluated and the correspondence are resolved for achieving the motion of vehicles as the solution proposed by Serby et al. [6].

The occlusion of vehicles in the image can lead to severe problems in object locating and tracking, especially in following two cases. Firstly, for the appearance-based tracking method, the appearance of a vehicle will exhibit different in two adjacent frames when occlusion exists, and thus make it difficult to associate the same object extracted from the frames. For another case, the motion-based tracking method, some key points of vehicles are hidden to the camera due to the overlapping, and such invisible points will result the inaccuracy of tracking.

2.2. Occlusion detection and segmentation

In order to achieve the better results, lots of modified tracking methods are proposed to mitigate the influence cause by occlusion. In general, the systematical solution to occlusion consists of two stages, which are occlusion detection and segmentation. However, the two stages are not always used together all the time. For instance, the detection-based method [7,8], which is popular in object tracking, does not use all two stages. Xing et al. [9] presented an online detection-based method. Instead of utilizing the particle filter algorithm to generate reliable tracklets, they introduced particle filter to an observation subset by multi-views and the multi-part detector corresponding to all visible object parts. In addition, some approaches leverage the mean shift and Kalman filter for vehicle tracking for tackling the occlusion issues [10,11]. However, it cannot be handled properly if the occluded objects have the similar appearance.

Due to 2D features are ineffective in some scenarios, 3D features are adopted for dealing with the occlusion cases [12]. By learning RGB-D data of traffic scene, Xia et al. [13] combine the depth-based segmentation method and the tracking method based on 2D features and detect and classify traffic-related objects, and the overlapping issues are solved effectively. Due to that occlusion detection should be detected ahead of segmentation, Ali et al. use a rectangle and an eclipse, which are applied to measure the changes of the distance and the size of the object, to detect the occlusion [14]. Zhou et al. [15] utilized block matching method to distinguish occlusions, and the appearance change detection is observed by Bayesian decision theory with different risk possibilities. If an occlusion is declared, the template will not be updated until the object reappears.

As for segmentation, various methods have been presented by researchers worldwide. Appearance matching is a common way in segmentation [16]. Based on the appearance templates of the images [17], the occluded objects can be tracked by associating foreground pixels to different objects. Yan and Essa [18] presented a region-level association process and an object-level localization process to track objects with long-period occlusion. Zhang et al. [19] focused on learning the semantic associations between superpixel sets. Moreover, no assumptions are made about size, shape or motion of objects except that the appearance is assumed to be distinguishable. However, those methods are based on the idea of object permanence and is not suitable for vehicle tracking in surveillance video. Wu et al. [20] and Senior et al. [21] modeled the object using a set of color templates and probability masks to track vehicles. However, if the object is occluded over a long time, the occlusion segmentation will fail.

Dynamic information of object is useful to differentiate distinct objects occluded each other, and then track and predict their motions. Parrilla et al. forwarded a method, which used optical flow algorithm, adaptive filter and (MLP) neural networks to predict the velocities of the object [22]. All in all, in most aforementioned approaches, the occlusion solutions are mainly determined by the tracking method.

3. Proposed method

Our proposed method, APPOS, is mainly consist of foreground extraction, occlusion detection, and occlusion segmentation whose output is used for vehicles tracking. The flowchart is shown in Fig. 1. Each step is specified as follows.

3.1. Categories of occlusion cases

The existing literatures have identified several categories of occlusion cases in the context of vehicle tracking. Prithwijit et al. [23] showed seven occlusion cases based on multiplicity and modes of object overlapping. In our proposed method, there are two aspects to categorize the occlusion cases: occluded objects and position relationship. From the aspect on occluded objects, the categories of occlusion cases are listed as vehicle to vehicle, vehicle to background objects, and vehicle to moving objects. From the other aspect on position relationship of the occluded objects, the categories of occlusion cases include left to right, top to bottom, and diagonal relationship. In the left to right relationship, the occluded objects are divided into two parts by a line perpendicular to X-axis. These two parts are the regions of occluded objects. In the top to bottom relationship, the occluded objects are divided by a line perpendicular to Y-axis. In the diagonal relationship, the foreground of video frames is divided into four parts by a line perpendicular to X-axis and a line perpendicular to Y-axis. Two of them are the regions of occluded objects.

3.2. Foreground extraction

As the preprocessing step of APPOS, the foreground extraction is implemented by the three-frame difference method on video frame sequences. Based on the extracted foreground, the vehicles