J. Vis. Commun. Image R. 39 (2016) 82–92

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

A real-time detector for parked vehicles based on hybrid background modeling ☆

Chi-Man Pun *, Cong Lin

Department of Computer and Information Science, University of Macau, Macau

ARTICLE INFO

Article history:
Received 18 December 2015
Revised 15 May 2016
Accepted 17 May 2016
Available online 19 May 2016

Keywords:
Adaptive threshold
Background subtraction
Hybrid background model
Parked vehicle

ABSTRACT

In this paper, a real-time detection system based on hybrid background modeling is proposed for detecting parked vehicles along the side of a road. The hybrid background model consists of three components: (1) a scene background model, (2) a computed restricted area map, and (3) a dynamic threshold curve for vehicles. By exploiting the motion information of normal activity in the scene, we propose a hybrid background model that determines the location of the road, estimates the roadside and generates the adaptive threshold of the vehicle size. The system triggers a notification when a large vehicle-like foreground object has been stationary for more than a pre-set number of video frames (or time). The proposed method is tested on the AVSS 2007 PV dataset. The results are satisfactory compared to other state-of-the-art methods.

1. Introduction

Because demands in many aspects of intelligent transportation systems have increased in recent years, many surveillance cameras have become installed along roads, for instance, on foot bridges, adjacent to traffic lights, in parking areas and on vehicles. Certain intelligent traffic surveillance systems in the literature employ computer vision or data mining techniques [1,2] to process data acquired from these cameras and to solve various automation problems in transportation. On the one hand, some applications [3–6] take advantage of these techniques to obtain knowledge about the surroundings, e.g., pedestrians [3], vehicles [4], lanes [5] and traffic signs [6], using in-vehicle cameras. These methods can potentially assist drivers by presenting scene information, issuing early warnings and even automatically preventing traffic accidents. On the other hand, many cameras have been installed outside vehicles to monitor the transportation status from a broader prospective. The captured video data are semantically analyzed patterns for determining traffic [7,8] or road conditions [9–11] using computer-vision-based methods.

Vehicles parked along the side of a road represent an important problem in traffic surveillance. Parking a vehicle on the side of a main road can slow down traffic streams, block the sight of following vehicles and even lead to traffic accidents. As digital cameras become cheaper and increasingly more surveillance cameras are installed for law enforcement on roads, it has become easier to observe a parked vehicle. However, manually observing these surveillance video remains laborious and costly. A real-time intelligent parked vehicle detection system can help address this problem. Many parked vehicle detectors [12–19] have been proposed in recent decades. Most methods employ fixed cameras and are based on background subtraction methods [16–19]. Generally, background-subtraction-based methods construct a background model off-line using initialization frames, and the background model subtracts the input frames during detection by assuming that the foreground objects lie in the differences. Then, a foreground analysis, e.g., [20], is conducted on the subtracted differences to determine the objects of interest. This framework has been proposed for more than a decade and has been found to be effective in traffic surveillance. However, inevitably, there have been some drawbacks to the conventional background subtraction method [17,18]. Arguably, the most serious problem is that the subtracted foreground (difference) usually contains substantial noise due to the scene background model and noise from cameras. Moreover, the background of real-world settings may change in an outdoor scene as a result of various complicated factors, e.g., changes in illumination and new stationary foreground objects. These factors are likely to introduce noise into the foreground mask. Many parked vehicle detectors that are based of background subtraction attempt to overcome these drawbacks. Lee et al. [12] proposed the detection of parked vehicles in a 1-D data domain. The video frames are first transformed into a 1-D vector. Both

☆ This paper has been recommended for acceptance by M.T. Sun.
* Corresponding author.
E-mail addresses: cmpun@umac.mo (C.-M. Pun), yb17403@umac.mo (C. Lin).

http://dx.doi.org/10.1016/j.jvcir.2016.05.009
1047-3203/

the background subtraction and the foreground analysis are conducted in the 1-D domain. Once the event is detected, the detector transforms the 1-D data back into the original video frame and locates the parked vehicles. Boragno et al. [13] integrated a sophisticated Digital Signal Processor (DSP) to handle the input frames, and a stationary detection filter was designed with constraints applied to stationary objects. Venetianer et al. [14] first grouped the foreground pixels into blobs and filtered the blobs based on calibrated size. Then, the blobs are tracked and subsequently used to generate tempo-spatial objects. Objects are classified by purpose of application. For parked vehicle detection, the vehicles are identified and recorded at a stationary time. Guler et al. [15] employed a multiple-object tracker to determine the location of each vehicle in a scene after obtaining the foreground mask from the background subtraction. Although certain methods [13–15] require users to manually draw the restricted area for the foreground analysis, others [12,21] do not specify the restricted area in the scene for detection. Thus, the detection is equally weighted across the scene in the background subtraction stage, and it is likely that the stationary objects in the unrelated area will trigger false-positive alarms. This might be improved using a more sophisticated foreground analysis method or even pre-trained model. However, real-time performance is crippled by more complicated operations and corresponding higher computational costs.

In this paper, we propose a real-time detection system for parked vehicles along the side of a road. The main idea is that we can determine where the road is and how large the vehicles are by exploiting the motion information of normal activity in the scene; in addition, we further estimate the roadside and the adaptive thresholds for the vehicles. Compared to conventional parked vehicle detectors, the proposed method is capable of automatically generating the road and the restricted area along it. After the weight map of the restricted area is obtained, we apply it to the background subtraction process and generate a substantially clearer foreground mask that is capable of automatically ruling out unrelated foreground objects. Moreover, during initialization, the vehicle size is regressed onto a function curve that outputs adaptive thresholds for a vehicle with respect to the Y-axis. The system triggers a notification when a large vehicle-like foreground object has been stationary for more than a pre-set number of video frames (or time). Because most of the computation is performed during the off-line initialization, the proposed method is straightforward and highly efficient in real time.

2. Proposed detector for parked vehicles using hybrid background modeling

Methods that estimate the dynamic status of objects with fixed cameras can be divided into two classes: object-tracking-based methods [14,15,21] and background-subtraction-based methods. The former methods employ object tracking methods to locate moving objects across frames. Most of these tracking methods construct a generative model for all the vehicles and perform tracking via a detection scheme. After detecting the vehicles, the trajectories are analyzed, and the system determines whether they are moving or stationary. There are certain disadvantages to object-tracking-based methods. First, a sophisticated vehicle detector that learns almost all types of vehicles from the real world, which is a difficult task considering the wide-ranging variance of shapes, colors and textures of vehicles, must be used. Second, false-positive detections of vehicles will result in irreversible false parked vehicle notifications. Finally, when the numbers of tracked objects in the scene increase, the computational cost increases, which leads to unstable performance and decreased efficiency.

Because surveillance cameras are typically fixed, an improved background-subtraction-based method represents a reasonable option. The underlying concept of background subtraction is to detect pixels in the new frame that are different from the prior information of the initialization frames or the background image. However, in contrast to conventional background subtraction methods that merely model the background using initialization video frames, we further extract the motion information from the initialization frames and transform the accumulated motion information into a restricted area map that indicates different weights for detection in the captured scene. Fig. 1 shows the workflow of the proposed method during online detection, and Fig. 2 shows that the trained hybrid background model is extracted in the initialization stage. A hybrid background model is constructed prior to on-line detection, which includes the following three components: (1) a scene background is modeled by a variant of the low-rank method, (2) a restricted area map is extracted from the motion information, and (3) a dynamic threshold curve for vehicles is regressed using the motion information.

2.1. Online detection using hybrid background model

Assuming that the hybrid background model is obtained, as proposed and explained in the next section, the online parked vehicle detection technique employs the hybrid background model to extract foreground objects in the restricted area. Then, we proceed to determine the stationary objects and whether there is any detection. To search for stationary objects, we must know how long the objects have been stationary. To record how long a pixel has been in foreground objects, we propose that the detector maintains a two-dimensional stationary matrix S, whose elements count how long (based on the number of frames) the corresponding pixels are covered by the foreground object. The stationary matrix is initialized as a matrix of zeroes. For clarity, we explain the workflow in Fig. 1 with the following procedures:

Algorithm 1: Real-time detection using hybrid background model

Initialize: stationary matrix $S_0 = 0$, stationary criterion (number of frames).

For frame i = 1 to end frame

Step 1: Perform background subtraction on frame $\boldsymbol{i}$ and compute foreground object labels using restricted area map.

Step 2: Update stationary matrix S_i using (1).

Step 3: Obtain $\overline{S}_i$ by thresholding the stationary criteria.

Step 4: Extract connected components from $\overline{S}_i$ as candidates. In addition, compute their centroids and average widths.

Step 5: Compute adaptive thresholds using the y coordinates of the candidates' centroids. In addition, determine if any detections have occurred.

End

After initialization and the setting of the stationary criteria by the user, we enter a loop of online detection. In **Step 1**, when a new frame at time $\boldsymbol{i}$ is acquired from the camera and input into the foreground extraction stage of the detector, the frame is first subtracted by the modeled background and simultaneously multiplied by the pre-computed map of the restricted area. The differences of each pixel are in the form of a three-elemental vector, from which RGB channels are generated. The differences are pixel-wisely measured by the L_2 norm. The output is a foreground mask $Mask_{foreground}$, which marks the foreground objects in the current frame. The following