Rule-Based Multiple Object Tracking for Traffic Surveillance Using Collaborative Background Extraction

Xiaoyuan Su, Taghi M. Khoshgoftaar, Xingquan Zhu, and Andres Folleco

Computer Science and Engineering, Florida Atlantic University, Boca Raton, FL 33431
xsu@fau.edu, {taghi,xqzhu,andres}@cse.fau.edu

Abstract. In order to address the challenges of occlusions and background variations, we propose a novel and effective rule-based multiple object tracking system for traffic surveillance using a collaborative background extraction algorithm. The collaborative background extraction algorithm collaboratively extracts a background from multiple independent extractions to remove spurious background pixels. The rule-based strategies are applied for thresholding, outlier removal, object consolidation, separating neighboring objects, and shadow removal. Empirical results show that our multiple object tracking system is highly accurate for traffic surveillance under occlusion conditions.

1 Introduction

Multiple object tracking (MOT) is important for visual surveillance and event classification tasks [1]. However, due to challenges such as background variation, occlusion, and object appearance variation, MOT is generally difficult. In the case of traffic surveillance, *background variations* in terms of illumination variation, small motions in the environment, weather and shadow changes, *occlusions* in terms of vehicles overshadowed or blocked by neighboring vehicles, trees, or constructions, and *vehicle appearance changes* in terms of different sizes of the same vehicles in different video frames, contribute to inaccurate visual tracking.

As traditional visual tracking methods, feature-based tracking detects features in a video frame and searches for the same features nearby in subsequent frames; Kalman filtering [2] uses a linear function of parameters with respect of time, and assumes white noise with a Gaussian distribution, however, the method with the Kalman filtering to predict states of objects can not be applied to objects in occlusion [3]; particle filtering [4] is appealing in MOT for its ability to have multiple hypotheses, however, its direct application for multiple object tracking is not feasible.

For traffic surveillance videos that generally have stationary background, it is important to segment moving vehicles from the background either when viewing the scene from a fixed camera or after stabilization of the camera motion. With the assumption of a stationary camera, we can simply threshold the difference of intensities between the current image frame with the background image, $I(x,y)-I_{bg}(x,y)$, to segment the moving objects from the background. However, due to background variations, this simple approach may not work well in general. In previous work, normal

G. Bebis et al. (Eds.): ISVC 2007, Part II, LNCS 4842, pp. 469–478, 2007.

(Gaussian) distribution, linear prediction and adaptation [5], and hysteresis thresholding [6] have been investigated to model the background changes.

We proposed a rule-based multiple object tracking system using a collaborative background extraction algorithm for the application of traffic surveillance, which is easy-to-implement and highly effective in handling occlusions in terms of removing outliers and shadows, consolidating objects, and separating occluded vehicles. The collaborative background extraction algorithm collaboratively extracts a background from several independent extractions of the background, which effectively removes spurious background pixels and adaptively reflects the environment changes.

Section 2 is our framework for the collaborative background extraction algorithm and rule-based multiple object tracking system for traffic surveillance. Experimental results and conclusions are in Section 3 and Section 4.

2 Framework

Our multiple object tracking system consists of the following procedure: adaptively extracting backgrounds using collaborative background extraction, generating binary images by differencing frames with background, applying the rule-based tracking strategies, and finally recording features of tracked objects.

2.1 Collaborative Background Extraction

We propose a non-Gaussian, single-thresholding background extraction method called collaborative background extraction, an adaptive background extraction algorithm using a collaborative strategy.

With the assumption that the background will not change significantly in a few seconds, we extract several backgrounds alternatively over a short period of time, e.g., every 60 frames (2 seconds for 30 fps videos), and then integrate these backgrounds into one. By updating the background every few seconds, we adaptively model the background changes.

As in Fig. 1, every single background extraction will produce a background with spurious points in different locations (Fig. 1(a)~(d), black points in the first four background images are intentionally-marked *foreground* points). With the help of collaborative extraction, the final background (Fig. 1(e)) is almost impeccable.

We collaboratively extract a background from four independent extractions of background, in order to produce a reliable background from multiple single extractions. We use the average intensity value of the labeled *background* pixels from the four extractions, and those *foreground* pixels (with values of 0) are then automatically replaced unless none of the four is classified as *background* (Equation 1).

$$bg_k(a,b) = \begin{cases} \frac{1}{|I|}\sum_{i\in I} bg_{k,i}(a,b), & I = \{i \mid bg_{k,i}(a,b) \neq 0; i = 1,2,3,4\} \\ 0, & |I| = 0 \end{cases} . \tag{1}$$

where k is the starting frame number of background extraction, i is the number of the four independent background extractions, I is for a non-zero pixel at (a,b) of the four background files. For example, for a pixel (a,b), when given $bg_{k,1}(a,b)=0$, $bg_{k,2}(a,b)=0$, $bg_{k,3}(a,b)=20$, and $bg_{k,4}(a,b)=30$, we will have $|I|=2$, and $bg_k(a,b)= (20+30)/2=25$.