

Online Multi-Person Tracking-by-Detection Method Using ACF and Particle Filter

T. Kokul

Postgraduate Institute of Science
University of Peradeniya
Sri Lanka
kokul@mail.vau.jfn.ac.lk

A. Ramanan

Department of Computer Science
University of Jaffna
Sri Lanka
a.ramanan@jfn.ac.lk

U.A.J. Pinidiyaarachchi

Department of Statistics and Computer Science
University of Peradeniya
Sri Lanka
ajp@pdn.ac.lk

Abstract— Automatically detecting and tracking multiple persons in videos is one of the main research interest in computer vision based applications. This paper presents a tracking-by-detection approach for tracking people in dynamic backgrounds with frequent occlusions by combining pre-trained generic person detector, online trained person-specific detector and a motion tracker. The popular aggregate channel features (ACF) are used to train the detectors and target specific particle filter is used as motion tracker. In order to learn right appearance of a target person, person-specific detector learns positive samples from prior frames which are detected by both generic person detector and person-specific detector. Data associations among the coincident detections of the detectors and tracker are used to update the person-specific detector and motion tracker. The person-specific detector searches the target person in a reduced region, which is defined by the associate motion tracker. A careful combination of detections of both detectors and tracker are used to locate the correct target person in the video sequence. Experiments have been carried out on Caltech pedestrian benchmark dataset. The proposed method shows better performance against state-of-the-art tracker while maintaining the tracking speed in real-time.

Keywords—Aggregate channel Features, particle filter, tracking-by-detection, Multi-person tracking

I. INTRODUCTION

Pedestrian tracking is of great interest in many applications such as video surveillance [1], abnormal event detection [2], sport analysis [3] and for a multitude of other purposes. The aim of pedestrian tracking is to find the location of a specific person in a video sequence, while the initial location information of the person is given.

Person tracking systems have been developed for more than two decades and much improvement have been made in recent years [4-6]. It still remains a challenging task, due to many reasons such as changes in pose, viewpoint and background, moving cameras, many persons with similar appearance and the occlusions occurred in scenes frequently. Multi-person tracking is more difficult than single person tracking, since it should handle more data. Online tracking is a complex task than offline tracking since the information processing should take place in real time, which is computationally expensive. Online multi-person tracking plays an important role in many

automatic surveillance systems such as traffic monitoring in airports and railway stations.

Many recent multi-person tracking systems follow the tracking-by-detection models [4] [5], since the person detection research has reached significant improvement and achieved high detection speed. These approaches treat tracking as a repeated detection problem. Such methods apply a detection algorithm such as histogram of oriented gradient (HOG) detector [7] or aggregate channel features (ACF) detector [8] on individual frames and use a data association of detection across frames. These techniques generally perform well in changing backgrounds and moving cameras.

In some multi-person tracking systems [9] initial target is given as the input bounding box to the system and the tracker accordingly learns the model online. Few other systems [4] [5] use a pre-trained person detector to initialise the tracker, which is essential to many realistic surveillance systems, and then to learn the model. The main challenge of using a pre-trained person detector is that the performance of the detector is not reliable always. It sometimes produces false positives and missing detections which leads to false tracking and drifts. Some off-line tracking approaches [10], [11] solve this issue by using information from future frames to locate the person in the current frame with a temporal delay. These techniques are neither realistic nor suitable for online tracking applications.

Most tracking approaches use information from prior frames to learn the appearance model. But insufficient appearance information leads to miss or false tracking of a tracker. Recent study [9] shows that accurate appearance model is more effective than a strong motion model.

In this paper, we propose an online multi-person tracking system, which relies on the careful selection of a generic pre-trained person detector, a person-specific detector and an efficient motion model. Proposed person-specific detector learns online to track a particular person and is able to separate the target person from the background as well as from other persons. We use the information obtained from prior frames to learn the person-specific detector to locate the person in the current frame. In addition to this, we use the association of particle filter based efficient motion model to reduce the false positives and missing detections of the person-specific detector. This motion model also increases the tracker speed by reducing the search space for individual persons.