Knowledge-Based Systems xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

One global optimization method in network flow model for multiple object tracking

Zhenyu He [a], Yuxin Cui [a,b], Hongpeng Wang [a,c], Xinge You [d,*], C.L. Philip Chen [d,e]

[a] Department of Computer Science, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen 518055, China
[b] Department of Computer Science and Engineering, University of South Carolina, Columbia, SC 29201, USA
[c] Shenzhen IOT Key Technology and Application Systems Integration Engineering Laboratory, Shenzhen 518055, China
[d] Department of Electronics and Information Engineering, Huazhong University of Science and Technology, China
[e] The Faculty of Science and Technology, University of Macau, Macau

A R T I C L E I N F O

Article history:
Received 9 November 2014
Received in revised form 14 March 2015
Accepted 19 April 2015
Available online xxxx

Keywords:
Multiple object tracking
Network flow model
Observation model

A B S T R A C T

In this paper, we address the task of automatically tracking a variable number of objects in the scene of a monocular and uncalibrated camera. We propose a global optimization method in network flow model for multiple object tracking. This approach extends recent work which formulates the tracking-by-detection into a maximum-a posteriori (MAP) data association problem. We redefine the observation likelihood and the affinity between observations to handle long term occlusions. Moreover, an improved greedy algorithm is designed to solve min-cost flow, reducing the amount of ID switches apparently. Furthermore, a linear hypothesis method is proposed to fill up the gaps in the trajectories. The experiment results demonstrate that our method is effective and efficient, and outperforms the state-of-the-art approaches on several benchmark datasets.

1. Introduction

Multiple object tracking is an important aspect of compute vision, especially the pedestrian tracking. It has been used in many tasks, such as the video surveillance and automatic drive. Unlike the single object tracking, which only keeps an eye on one object along the frame sequence, multiple object tracking must track all the targets in camera sight and address many complex situations, e.g., object enters and exits, one is occluded by another or some barriers, and so on. In single object tracking, we may just concentrate on how to represent the object which needs to be tracked. However, in multiple object tracking, the focus has been transformed to how to address the data association problem, i.e., how to find the corresponding observation in previous frames or next frames. Moreover, the multiple object tracker has to automatically track targets of a certain category, so that when a target emerges in the scene, the tracker should start a new tracking. If the camera is uncalibrated and monocular, lots of information will be lost when the 3D world is mapped to 2D image, which brings about some intricate occlusions. Different methods have been proposed to address the problem of tracking multiple objects using monocular and uncalibrated camera in recent years. Nevertheless, the factors such as the complicated background, the crowded people and the low-quality video can render the multiple tracking extremely challenging.

With the improvement of the detector [1–4] in recent years, tracking-by-detection [5–7] is becoming feasible and popular. In the framework of tracking-by-detection, it is not indispensable to observe the targets during the tracking procedure any more and we can concentrate on addressing the data association problem because high reliable object detections are given by the detector as the input observations. That is, we only need to connect the target hypotheses generated by the detector as their similarity across frames. Several models based on graph theory have been introduced to simulate the tracking problem, such as the Maximum Weight Independent Set (MWIS) model [6], Generalized Minimum Clique Graphs (GMCP) model [8] and network flow model [9–11]. Actually, the detection result may have many false positives and missing detections. The missing detections are mainly caused by occlusions. Some state-of-the-art detectors [12,13] have a good performance on the partial occlusion. But for the full occlusion, it is helpless. Hence, it is necessary for tracking-by-detection tracker to be able to eliminate the false positive and fill up the missing detections (see Fig. 1).

Tracking in network flow model is a kind of global optimal procedure. In this framework, the object responses, i.e., the nodes,

* Corresponding author.
E-mail addresses: zyhe@hitsz.edu.cn (Z. He), wanghp@hit.edu.cn (H. Wang), youxg@hust.edu.cn (X. You), Philip.Chen@ieee.org (C.L. Philip Chen).

http://dx.doi.org/10.1016/j.knosys.2015.04.018
0950-7051/

Fig. 1. Examples of detection input (left) and tracking output yielded by our approach (right). The input observations have a lot of false positives and false negatives. Besides, the box size and positive often do not fit the object. Our approach can remove the false positives and repair the missing detections.

compose a large and intricate network graph. Two nodes are connected by an arc means that they may represent the same target and the cost of arcs or edges in the graph indicates the agreement between two detection observation. Pirsiavash et al. [5] formulated data association to an MAP problem and solved it as a min-cost flow issue. Rather than applying the proven optimal algorithms directly, such as the push relabel algorithm [14] and successive shortest path [15], they proposed a greedy method on the basis of finding the shortest path with dynamic programming algorithm. Their approach can return all the trajectories in extremely short time, not giving the global optimal solution but a high-quality approximate solution. Our work is mainly inspired by the min-cost flow network described in [5]. We find out that, in our cost flow network model, this greedy algorithm is not high-quality any more. We aim to reformulate the similarity evaluation to adapt the network flow mode to long term occlusions and improve Pirsiavash's approximate greedy method to be not only efficient but also precise. The main contributions of this paper include:

- A new integrated observation model for evaluating the affinity between two detection observations. This integrated observation model not only has better robustness than previous models, but also can deal with the occlusions situation.
- An improved approximate greedy algorithm. For the problem that Pirsiavash's approximate greedy may cause a lot of ID switches, this improved approximate greedy algorithm can remarkably eliminate the ID switches and its running time is much less than optimal algorithms.
- A simple and effective linear hypothesis method for reducing the number of false negatives. After the optimization, a lot of gaps caused by the occlusion exist in the trajectories though the detections have been connected to trajectories. This linear hypothesis mechanism can fill up most of the gaps accurately.

In the rest of this paper, we briefly discuss the related work in Section 2 and describe the construction of network flow model in Section 3. A novel optimization algorithm is proposed in Section 4 and a detailed experimental evaluation of the presented method is given in Section 5. Finally, a conclusion is drawn in Section 6.

2. Related work

Considerable developments have occurred for multiple targets tracking since radar tracking [16] twenty years before. Early approaches follow objects in local strategy [17,18]. That is, they solve the data association frame-by-frame and object-by-object. Once one object is found in one frame, the tracker will keep on looking for it in the next frame based on its state estimate in one or more previous frames. In this framework, kalman filter [19] and particle filter [20] which is known as a sequential Monte Carlo method, are usually applied to connect the object hypotheses.

Many approaches in global strategy are proposed in recent years. These methods usually address the global data association by solving the global optimization problem over a long period window whose range may be from tens of frames to more than one thousand frames. Contrasted to the approaches in local strategy, these approaches have a higher accuracy whereas the temporal delay is inevitable. Besides, since most of globally optimal methods employ the reliable detection output of a high-quality detector as their input, the performance of a globally optimal method can be heavily influenced by the input detection responses. According to their specific models, tracking task can be modeled to dynamic programming problem [21], cost-flow network [22,23], Maximum Weight Independent Set [6], Generalized Minimum Clique Graphs [8], and so on.

In the network flow model, the basic and first procedure is transforming the global data association to a maximum-a

Please cite this article in press as: Z. He et al., One global optimization method in network flow model for multiple object tracking, Knowl. Based Syst. (2015), http://dx.doi.org/10.1016/j.knosys.2015.04.018