Image and Vision Computing 48–49 (2016) 42–56

Contents lists available at ScienceDirect

Image and Vision Computing

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

Editor's Choice Article

A variational based model for estimating true tracklets in wide area surveillance☆

Ehsan Pazouki, Mohammad Rahmati*

Amirkabir University of Technology, Computer Engineering Department, Tehran, 15875-4413, Iran

ARTICLE INFO

Article history:
Received 21 April 2015
Received in revised form 13 January 2016
Accepted 4 February 2016
Available online 23 February 2016

Keywords:
Variational model
Wide area surveillance
True tracklet estimation
Multi-object tracking
Camera network tracking

ABSTRACT

In many wide area surveillance applications, tracking objects is usually accomplished by using network of cameras. A common approach to any multi-objects tracking algorithm in a network of cameras comprises of two main steps. First, the movement trajectory of each object, within the field of view of a camera, is extracted and is called object tracklet. Then, the set of tracklets are used to determine the persistent trace of each object. In this paper, we assume that the tracklets are extracted by a conventional tracking algorithm. The occurrence of occlusion between objects, within the viewing scene, leads to various types of errors on the extracted tracklets. If these erroneous tracklets are used in a multi-object tracking algorithm and ignoring the correction phase, then the errors are propagated and affect the results of tracking algorithm. Therefore the true tracklets have to be estimated from the erroneous tracklets. In this paper, we propose a variational model for estimating the true tracklets. The variational principle proposed in this model is established by first introducing a variational energy function. Then the erroneous tracklets are used to estimate the true tracklets through optimizing the energy function. The proposed method is evaluated on two well known datasets and a synthetic dataset which is particularly developed to demonstrate the performance of our algorithm under challenging scenarios. The 10 common metrics, which are used in other multi-objects tracking applications, are used for quantitative evaluations. Our experimental results illustrate that our proposed model estimates the true tracklets which improves the overall association performances.

1. Introduction

The development and use of visual surveillance systems are becoming widespread in the growing cities. An intelligent surveillance system takes advantage of many cameras, which are connected in a network, to cover and monitor a wide area. An important component of these intelligent surveillance systems is the tracking module. The tracking module is used for visual object tracking. The visual object tracking is performed by detecting an object in the frame by frame of a video which is captured from one or many cameras.

There are numerous methods that have been proposed to deal with the object tracking task under different scenarios in the images captured by a single camera [30,2]. However, the multi-objects tracking in a network of cameras is a young and challenging field of research.

The multi-objects tracking task in a network of cameras usually consists of two main steps [1,31,10]. In the first step, objects are tracked in each camera using a single camera tracking algorithm [30,2]. The trajectories of moving objects which are detected within a single camera images are called *tracklets*. Then the extracted *tracklets* are associated to determine the persistent traces of all objects. In the second step, it is usually assumed that the input *tracklets* are accurate and consequently no analysis is required in this stage. A correct *tracklet* refers to the one that all its observed sample points, which are extracted from the consecutive frames, belonging to one object. Therefore, extracting the correct *tracklets* before any association step is an essential requirement. On the other hand, in a real world scenario, many challenging events arise which affect the performance of any tracking algorithm. An important challenge in object tracking is occlusion which is caused by sudden disappearing of objects [30,24] in video frames. In this situation, the tracking algorithms may make a wrong decision and wrong objects are tracked. As a result, the extracted *tracklet* is associated with more than one object. We refer to these *tracklets* as the erroneous *tracklets*.

The erroneous *tracklets* in the extraction stage are resolved by some authors by introducing complex and robust algorithms in the initial step [24,3,25,39,38]. For example, in [24] several models with

☆ Editor's Choice Articles are invited and handled by a select rotating 12 member Editorial Board committee. This paper has been recommended for acceptance by Xiaogang Wang.

* Corresponding author.
E-mail addresses: ehsan.pazouki@aut.ac.ir (E. Pazouki), rahmati@aut.ac.ir (M. Rahmati).

http://dx.doi.org/10.1016/j.imavis.2016.02.001

many parameters are introduced and a success rate of 85% is achieved. In that paper, the authors use different algorithms to handle occlusion in different situations. Algorithms include: (i) a content-adaptive progressive occlusion analysis algorithm, which combines the information provided by spatiotemporal context, reference target, and motion constraints together, (ii) a variant-mask template matching for rectifying the target location, (iii) a drift-inhibitive masked Kalman appearance filter, which accurately evaluates the influence of template drift, and (iv) a local best match authentication algorithm.

In other work [39], the main reliance is based on viewing the multi-camera with an overlapped camera network; however, the cameras in the multi-camera tracking surveillance system are usually non-overlapped, particularly in wide area surveillance systems. Also, in [37], a novel approach is proposed for multi-objects tracking in a single camera and the results of this paper illustrate that the obtained *tracklets* yet embed unavoidable errors.

In all of these methods, the errors are corrected to some extent only under some restrictive assumptions imposed on the behavior of objects and the topology of cameras, to name a few. Under the absence of any assumption, the occurrence of errors on the extracted *tracklets* are inevitable. Thus, the risk of occurring errors in the extracted *tracklets* in the first step is high. Evidently, if the erroneous *tracklets* are not corrected before association, the errors will be propagated to the association results which are the persistent trace of objects. Therefore, an intermediate step is required to correct the errors.

In this paper, we propose a novel variational model as an intermediate step to estimate error free *tracklets* from erroneous *tracklets* before they are used in step-two of the tracking algorithm, as shown in Fig. 1. In this figure, the block diagram illustrates the inclusion of our proposed intermediate step in conjunction with the two conventional steps.

The conventional two steps algorithms for associating *tracklets* which were proposed previously aim to generate the persistent trace of objects in camera networks [1]. In some of these works, the error free ground truth *tracklets* are assumed to be available and therefore they are used in the association process [10,23,13,32,14,26]. In other works, the erroneous *tracklets* are collected and are used as input to the algorithm. Therefore, the intermediate step which is introduced above is necessary to estimate the true *tracklets*.

In [1,31,5], the errors are categorized into two classes and different solutions are proposed for each class. In the first class which is the lost class, a tracking error (**TE**) metric [9] is used to quantify the errors. The **TE** metric is the distance between the model of tracked object in the current observation and the predicted model

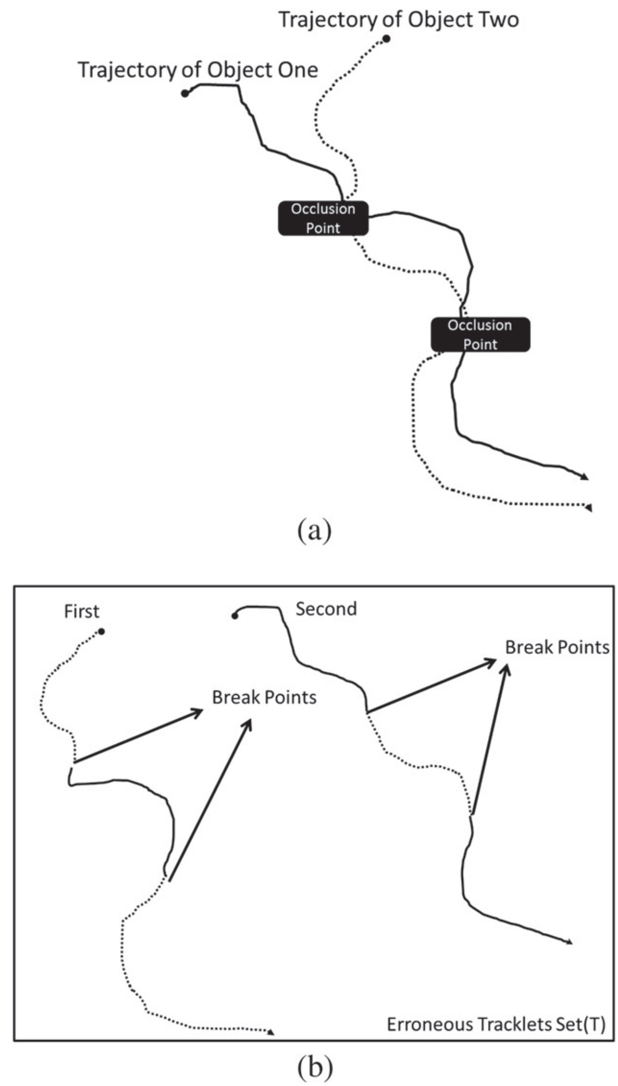

Fig. 2. An example of *tracklet* extraction: (a) an example of occlusion scenario which leads to erroneous *tracklets*; and, (b) erroneous *tracklet* set which is obtained from sample scenario Fig. 2a.

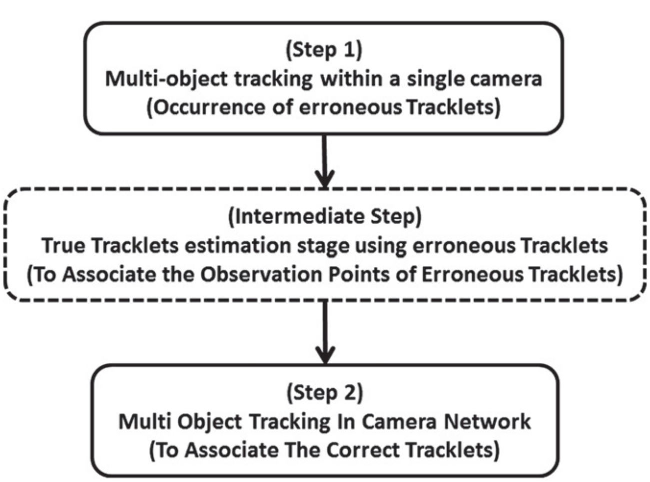

Fig. 1. A block diagram of the conventional two steps multi-object tracking algorithm in camera network which is extended with the intermediate step of our proposed model.

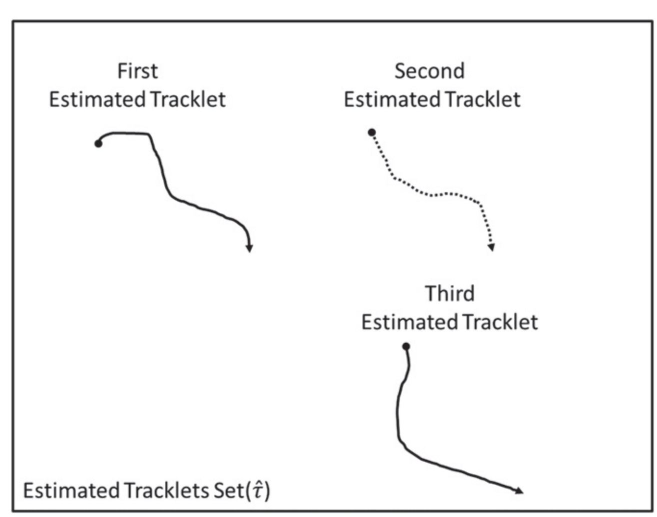

Fig. 3. Estimated *tracklet* set which is obtained from second erroneous tracklet of sample scenario Fig. 2b by using break point algorithm.