Information Sciences 278 (2014) 448–465

Contents lists available at ScienceDirect

Information Sciences

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

Multiple object tracking with partial occlusion handling using salient feature points

M.M. Naushad Ali [a], M. Abdullah-Al-Wadud [b], Seok-Lyong Lee [a,*]

[a] Department of Industrial and Management Engineering, Hankuk University of Foreign Studies, 89 Wangsan, Mohyun, Yongin-si, Kyonggi-do, Republic of Korea
[b] Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

ARTICLE INFO

Article history:
Received 3 July 2013
Received in revised form 8 February 2014
Accepted 5 March 2014
Available online 25 March 2014

Keywords:
Multiple object tracking
Salient feature point
Particle filter
Corner detection
Partial occlusion

ABSTRACT

Handling occlusion has been a challenging task in object tracking. In this paper, we propose a multiple object tracking method in the presence of partial occlusion using salient feature points. We first extract the prominent feature points from each target object, and then use a particle filter-based approach to track the feature points in image sequences based on various attributes such as location, velocity and other descriptors. We then detect and revise the feature points that have been tracked incorrectly. The main idea is that, even if some feature points are not successfully tracked due to occlusion or poor imaging condition, the other correctly tracked features can collectively perform the corrections on their behalf. Finally, we track the objects using the correctly tracked feature points through a Hough-like approach, and the object bounding boxes are updated using the relative locations of these feature points. Experimental results demonstrate that our method is proficient in providing accurate human tracking as well as appropriate occlusion handling, compared to the existing methods.

1. Introduction

Visual object tracking has drawn increasing attention in recent years since it has been an important and complicated task in the field of computer vision. It has a wide range of application areas including automatic object detection, object surveillance, activity analysis, and human computer interaction [2,4,19,32]. For example, automated surveillance systems play important roles in monitoring factories, schools, traffic, hospitals, banks, and other facilities, in which the system often includes object detection, tracking, and event analysis according to diverse requirements. Object tracking in a scene is broadly categorized as either single or multiple object tracking. Tracking a single object or some isolated objects is comparatively simpler than tracking multiple objects in the presence of occlusion and/or poor background conditions. On the other hand, various multi-view as well as single view approaches have been proposed to handle the tracking of occluded targets. The multi-view approaches [8,13,17,38] use information from more than one camera to decrease hidden regions in order to recover 3D space information. However, such a setup for capturing videos of the same scene by multiple cameras may not always be possible in practice. The currently available single view approaches can handle isolated objects effectively, while the tracking of multiple objects is severely hampered and may often fail, in the presence of occlusion, especially inter-object occlusion. If a change occurs in the subject's appearance, such as a change of object shape due to slight rotation, most

* Corresponding author. Tel.: +82 313304357; fax: +82 313304093.
E-mail addresses: naushad_iut@yahoo.com (M.M. Naushad Ali), wadud@ieee.org (M. Abdullah-Al-Wadud), sllee@hufs.ac.kr (S.-L. Lee).

http://dx.doi.org/10.1016/j.ins.2014.03.064
0020-0255/

algorithms may fail to track correctly. We address multiple object tracking in the presence of occlusion using the single view approach.

Occlusion has been regarded as one of the main challenging tasks in visual object tracking since it leads to severe degradation of tracking accuracy. In video streams, at times some parts of objects may not be visible due to occlusion. A human can recognize an object even though it is partially occluded. If the object is partially visible, the human brain can reconstruct the entire object using the inference based on the visible portion of the object and the knowledge of the object's general structure. For example, the full body of a human is visible in Fig. 1(a), while only a portion of the body is visible in Fig. 1(b) due to the obstacle. Despite the obstacle, a human can predict the size and shape of the object by assessing the posture of the visible parts. In the presence of occlusion, the sophisticated techniques may be needed to implement the tracking system that analogizes the object recognition mechanism of a human. Many existing tracking methods show reliable tracking for multiple moving objects when the objects are clearly separated from each other and their colors are distinct from those of the background. However, these tracking methods may otherwise fail and an object's tracker may switch to another moving object or to a location somewhere in the background. The aim of this study is find a way to handle such a problem. We focus on tracking multiple moving objects in image sequences when parts of the objects are occluded due to background objects or other target objects. In our approach, a strategy similar to the aforementioned object-reconstruction mechanism of a human is used to handle occlusion.

In this paper, we propose a salient feature point (SFP)-based method to track multiple objects with partial occlusion handling. The proposed tracking algorithm is based on the particle filtering approximation of different feature points of the target objects. Even if some feature points are not successfully tracked due to occlusion, overlap or changes in imaging conditions, the other correctly tracked features can collectively perform the correction for them. The overall structure of the proposed method is shown in Fig. 2. It is composed of three steps: (1) feature descriptor construction for SFPs, (2) object tracking through the prediction of SFPs' locations, and (3) feature descriptor update for SFPs.

To construct a feature descriptor in Step 1, we first represent each target object by a rectangular bounding box in a video. Then, multiple SFPs are extracted for each target object using a corner detection algorithm, and the histogram of oriented gradient (HOG) descriptor for each SFP is generated. To track objects in the consecutive frames in Step 2, a particle filter-based approach is employed to track the SFPs of each object by predicting their subsequent locations. We use the relative location of each SFP to obtain information about the object's shape. The relative locations can also help to determine if an SFP is correctly tracked. We then identify the incorrectly tracked SFPs through outlier analysis and find the bounding box using correctly tracked SFPs. In Step 3, we revise the locations of the incorrectly tracked SFPs based on their previous relative locations, and update the descriptors of the correctly tracked SFPs. Steps 2 and 3 are repeated while subsequent frames remain, and the proposed approach is thus able to successfully track multiple objects in a scene.

The main contributions of the proposed method can be summarized as follows. (1) We present an effective method that addresses one of the main challenging tasks in multiple moving object tracking; that is, tracking in the presence of partial

Fig. 1. Human inference in reconstructing an object in the presence of occlusion: (a) fully visible and (b) partially visible. This example shows that a human can predict a full object (shown using bounding rectangles), even though some portions of the object are not visible.