J. Vis. Commun. Image R. 22 (2011) 673–685

Contents lists available at SciVerse ScienceDirect

J. Vis. Commun. Image R.

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

Integrated video object tracking with applications in trajectory-based event detection

Hsu-Yung Cheng [a,*], Jenq-Neng Hwang [b]

[a] Department of Computer Science and Information Engineering, National Central University, No. 300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan
[b] Department of Electrical Engineering, University of Washington, 185 Stevens Way, Paul Allen Center, Seattle, WA, USA

ARTICLE INFO

Article history:
Received 24 June 2009
Accepted 30 June 2011
Available online 18 July 2011

Keywords:
Tracking
Kalman Filter
Adaptive particle sampling
Data Association
Trajectory analysis
Event detection
Intelligent Surveillance
Video Analysis

ABSTRACT

This work presents an automated and integrated framework that robustly tracks multiple targets for video-based event detection applications. Integrating the advantages of adaptive particle sampling and mathematical tractability of Kalman filtering, the proposed tracking system achieves both high tracking accuracy and computational simplicity. Occlusion and segmentation error cases are analyzed and resolved by constructing measurement candidates via adaptive particle sampling and an enhanced version of probabilistic data association. Also, we integrate the initial occlusion handling module in the tracking system to backtrack and correct the object trajectories. The reliable tracking results can serve as the foundation for automatic event detection. We also demonstrate event detection by classifying the trajectories of the tracked objects from both traffic monitoring and human surveillance applications. The experimental results have shown that the proposed tracking mechanism can solve the occlusion and segmentation error problems effectively and the events can be detected with high accuracy.

1. Introduction

Computer vision and video analysis techniques play an important role in modern intelligent systems. Video-based systems can capture a larger variety of desired information and are relatively inexpensive because cameras are easy to install, operate, and maintain [1]. With the huge amount of video cameras installed everywhere nowadays, there is an urgent need for automated video understanding techniques that can replace human operators to monitor the areas under surveillance [2]. In video based intelligent systems, a robust tracking system identifies and keeps track of each individual moving object. After tracking results are obtained, a recognition model based on appropriately extracted features can be constructed for the event detection purpose.

For tracking systems, multiple object tracking has been investigated intensively due to its important role in various applications [3]. The tracking problem can be modeled as a dynamic system represented by states in discrete time domain. A Kalman filter [4,5] assumes a linear state transition function and a linear observation model as well as Gaussian noises to obtain optimal state prediction and update formulations. Kalman filtering is based on restricted assumptions and can therefore achieve computational simplicity, which is a desired property for real-time tracking systems. If the linear and Gaussian assumptions are relaxed, the solutions cannot be determined analytically in general and there exist only sub-optimal approximations [6]. The extended Kalman filter (EKF) [7] relaxes the restriction of linearity on the state transition function and observation model, i.e., the system transition function and observation model can be any differentiable functions in EKF. Foresti [8,9] utilized a modified EKF model which does not require *a priori* information to perform tracking. Particle filters further relax the Gaussian assumption to model non-Gaussian and multi-model posterior distributions. For more complex models, particle filters have their merits when dealing with objects that are closer to the camera and having dramatic motion changes, especially in the cases when the moving objects cannot be easily segmented from the background. However, particle filters still need to cooperate with occlusion detection and handling techniques in order to have better performance. Moreover, the computational complexity of particle filters is usually very high. Maggio and Cavallaro [10] designed a combined tracker utilizing particle filters and mean shift techniques to produce a smaller number of samples than traditional particle filters and more reliable results than traditional mean-shift tracking. However, their tracker was not extended to multi-target tracking. Furthermore, without explicit occlusion analysis, particle filters that continue relying on outlier pixels and updating the appearance models can significantly deteriorate the tracking accuracy at occlusion situations [11]. Therefore the ability of handling occlusion cases of their hybrid tracker was quite limited. In fact, although researchers are turning to use

* Corresponding author. Fax: +886 3 422 2681.
E-mail addresses: chengsy@csie.ncu.edu.tw (H.-Y. Cheng), hwang@u.washington.edu (J.-N. Hwang).

1047-3203/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2011.07.001

particle filters because Kalman filtering is only optimal when the linear and Gaussian assumptions are held, many recent works using particle filters [11–15] still use linear functions as the underlying system models. This is reasonable because when the video frame rate is high enough and the camera is set up with a certain distance away from the objects under surveillance, which are often the cases in many surveillance applications, the object motion is close to linear as we look at them frame by frame. Therefore, even with occlusion or segmentation errors, the use of Kalman filter with linear prediction can still be justified. This argument also explains why Kalman filters are still adopted in many recently proposed systems [16–18]. The main issue is how to update the filters with correct measurements when occlusion or segmentation errors occur. We solve this problem effectively by using adaptive particle sampling to construct a measurement candidate list in case of occlusion or segmentation errors. While we use particle sampling techniques to provide reasonable measurement candidates, the mathematical tractability and closed-form solutions provided by Kalman filtering do not need to be sacrificed in the overall tracking model. To associate the measurement candidates with each target object, we propose an enhanced version of probabilistic data association (PDA) approach. The classical PDA [19,20] was originally proposed to associate measurements with targets being tracked in sonar or radar tracking applications, where targets are represented as dots without any definition of shape and measure of area. However, targets in video object tracking applications are moving regions, which cannot be effectively handled by traditional PDA. With appropriate modification, PDA could be very useful for video object tracking systems as well. The enhanced PDA utilizes extra information obtained from video scenes to perform accurate and reliable data association for each target. We will show that even a linear prediction model such as Kalman filtering can work well for various applications with the support of the proposed method.

For event detection, both unsupervised [21] and supervised [22,23] classification methods have been developed. Since unsupervised methods are commonly used to detect anomaly activities instead of some predefined classes of events that we are interested in, supervised event detection remains our prime choice. Common supervised classification models applied to intelligent systems include artificial neural networks (ANNs) and dynamic probabilistic network models such as Hidden Markov Models (HMMs) and Dynamic Bayesian Networks (DBNs) [22,23]. Although an ANN has the advantage of being flexible and being able to model accident sequence consisting of a large number of random processes, HMMs are considered to be more robust against observation sequences with variable length compared to ANN methods. In trajectory-based event detection applications, HMMs can deal with partial trajectories because their superior ability of handling dynamic length data. The ability to manage partial trajectories makes it possible for the system to recognize the ongoing behavior of a moving object while it is still in the scene. In this paper, we demonstrate event detection applications by classifying the trajectories of the tracked objects. An HMM is trained for each event characterized by each class of trajectory. When performing classification, we compute the likelihood that a trajectory belongs to each model and classify the trajectory to the class that has maximum likelihood. The maximum likelihood classification method is able to perform well in multi-class cases and classify the trajectories with high accuracy.

The rest of paper is organized as follows: In Section 2, we propose a tracking algorithm, including the particle sampling for measurement candidate list construction and enhanced probabilistic data association for measurement update. Section 3 elaborates the feature extraction and recognition model for trajectory analysis and event detection. The experimental results are reported and discussed in Section 4. Finally, conclusions are made in Section 5.

2. Multi-object tracking via adaptive particle sampling and enhanced probabilistic data association

Fig. 1 illustrates the framework of the proposed tracking system. A background model is estimated and updated from the video based on the method described in [24] to segment the moving foreground video objects (VO) from the background scene. Both the segmented VOs and the original image frames are inputs to the tracking system. Each module in the tracking system is elaborated in the following sub-sections.

2.1. Tracking list construction

The regions in the video scene that we are interested in monitoring depend on the surveillance environment and the coverage

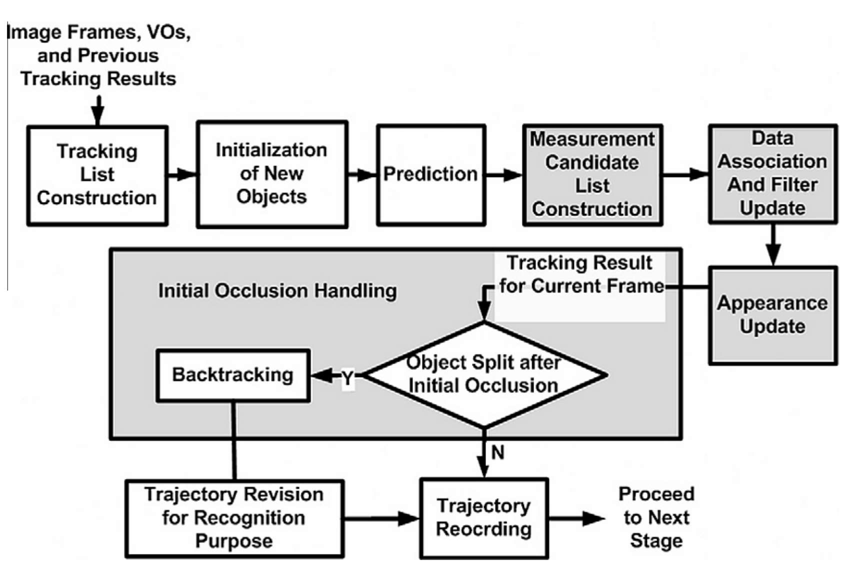

Fig. 1. Tracking system framework.