Signal Processing 89 (2009) 1844–1849

Contents lists available at ScienceDirect

Signal Processing

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

Fast communication

Adaptive particle sampling and adaptive appearance for multiple video object tracking

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

[a] Department of Computer Science and Information Engineering, National Central University, Chung-Li, Taiwan
[b] Department of Electrical Engineering, University of Washington, Seattle, USA

ARTICLE INFO

Article history:
Received 7 August 2008
Received in revised form
31 January 2009
Accepted 31 March 2009
Available online 8 April 2009

Keywords:
Kalman filter
Adaptive particle sampling
Adaptive appearance
Tracking
Occlusion handling

ABSTRACT

In this work, we propose an innovative method to integrate the Kalman filter and adaptive particle sampling for multiple video object tracking. Taking advantage of both the closed-form equations for optimal prediction and update from Kalman filters and the versatility of particle sampling for measurement selection under occlusion or segmentation error cases, the proposed method achieves both high tracking accuracy and computational simplicity. The adaptive particle sampling, which uses parameters updated by Kalman filters, can thus require only a small number of particles to achieve high positioning and scaling accuracy. Also, the concept of adaptive appearance is applied to enhance the robustness of occlusion handling. The experimental results confirm the effectiveness of the proposed method.

1. Introduction

The tracking problem can be modelled as recursively computing the posterior density of the system state given all the available measurements in a dynamic system represented by states in discrete time domain. In general, the optimal Bayesian solution is analytically intractable. Therefore the Kalman filter (KF) [1] assumes a linear state transition function and a linear observation model as well as Gaussian noises to obtain optimal state prediction and update formulations. Based on these restricted assumptions, KF can achieve computational simplicity, which is a desired property for real-time tracking systems. For video object (VO) tracking, researchers have turned to particle filter-based tracking methods due to the argument that the object movements are not linear. However, 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 KF with linear prediction can still be justified. Actually, many recent works using particle filters [2–5] still use linear functions as the underlying system motion models. This also explains why KFs are still adopted in many recently proposed systems [6–8]. 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 for measurement selection 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 KF do not need to be sacrificed in the overall tracking model. The rest of the paper is organized as follows. In Section 2, we describe the overall proposed tracking framework. In Section 3, the measurement selection procedure is elaborated. Section 4 displays the experimental results and Section 5 gives the conclusion of this work.

* Corresponding author.
E-mail addresses: chengsy@fox1.csie.ncu.edu.tw (H.-Y. Cheng), hwang@u.washington.edu (J.-N. Hwang).

0165-1684/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2009.03.034

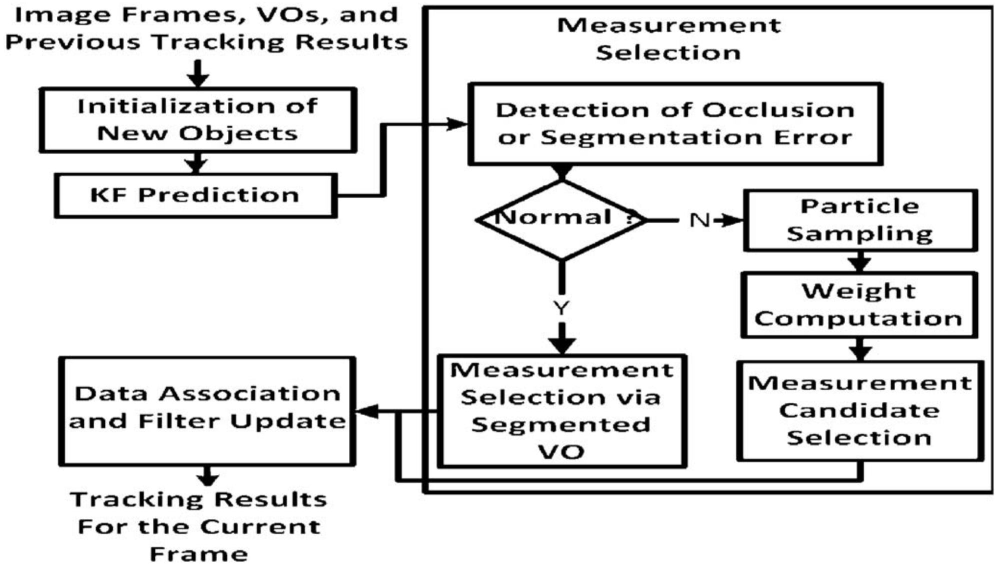

Fig. 1. Proposed tracking framework.

2. Proposed system framework

The proposed tracking system frame work is illustrated in Fig. 1. Before applying the tracking algorithm, a background model is estimated and updated from the video to segment the VOs from the background scene. For a new object entering at time instance k, the system initializes its system state $x_k = [u_k \;\; v_k \;\; \dot{u}_k \;\; \dot{v}_k \;\; a_k \;\; b_k]^T$ and an appearance model ξ_k for it. Commonly used appearance models are colour values of the fitted ellipse (colour matrices), and compact summarization of colour distribution such as histograms or mixture of Gaussians [2]. As shown in Fig. 2, the position (u_k, v_k) is coordinate of the centroid of an object in the image plane. The velocities $\dot{u}_k$ and $\dot{v}_k$ are initialized as zeros. The sizes (a_k, b_k) are the length of the major axis and the minor axis of the ellipse fitted on the VO. The measurement state is defined as $y_k = [u_k \;\; v_k \;\; a_k \;\; b_k]^T$. After the initial states for the objects are obtained, we perform KF prediction. Afterwards, we obtain the measurements using the measurement selection procedure described in the next section and utilize an enhanced version of probabilistic data association (EPDA) [9] to associate the measurements with each target object for filter update. Let β_j denote the probability associated with the jth measurement of a target object in EPDA. We let the prior term of β_j be proportional to the normalized weight of the measurement in the measurement candidate list for $j \neq 0$.

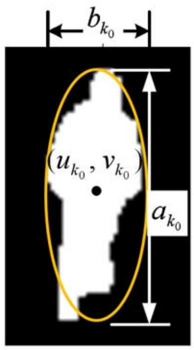

Fig. 2. Object initialization.

3. Measurement selection

In the measurement selection procedure (see Fig. 1), the system detects if an occlusion or a segmentation error occurs. If there is no occlusion or segmentation error, the segmented VOs are reliable and the measurements for a target object can be obtained by referring to the segmented VOs. Otherwise, we do not trust the segmentation results and perform adaptive particle sampling to find reasonable measurement candidates instead.

3.1. Detection of occlusions and segmentation errors

Based on the predicted state and the segmented VOs, we design a reasoning logic to detect occlusions and segmentation errors. Occlusions are discovered when two or more tracked objects start to merge with one another, as illustrated in examples in Fig. 3(a). Occlusion detection can be accomplished by checking the predicted state of every pair of target objects in the tracking list. In order to discover segmentation errors, a validation gate (see [9]) is built upon the predicted system state of the target object. Several segmentation error examples are shown in Fig. 3(b). If we cannot find any measurement within the validation gate of a target object, then we regard it as a segmentation error because either the positions of all the VOs are far away from the predicted position, or the major axes and minor axes of the ellipses fitted on the VOs are