

Robust observation model for visual tracking in particle filter

Anping Li^{a,*}, Zhongliang Jing^{a,b}, Shiqiang Hu^{a,b}

^a*Institute of Aerospace Information and Control, School of Electronic, Information and Electrical Engineering, Shanghai Jiao Tong University, Hua Shan Road 1954, Shanghai 200030, PR China*

^b*Institute of Aerospace Science and Technology, Shanghai Jiao tong University, Hua Shan Road 1954, Shanghai 200030, PR China*

Received 2 December 2005

Abstract

A robust observation model for visual tracking is proposed in this paper. The model consists of three appearance models: fixed appearance model, adaptive appearance model, and two-frame appearance model. The three appearance models are used, respectively, for catching unchanged components, slow changes, and rapid changes in object appearance. During tracking, the robust observation model is incorporated in a particle filter, and the particle filter can automatically select proper appearance models to track object according to the current tracking situation. Occlusion analysis is implemented using the *M*-estimation technique. Experimental results demonstrate that the proposed algorithm can track objects well under many challenging tracking situations.

© 2006 Elsevier GmbH. All rights reserved.

Keywords: Robust observation model; Occlusion handling; *M*-estimation technique; Particle filter

1. Introduction

Visual tracking has very important applications in many fields, such as vision-based control, intelligence robotics, smart surveillance, and video recognition. During tracking, object appearance may change due to illumination changes, pose variations, object deformations, or occlusions, etc. The potential variability of object appearance makes tracking difficult. Thus, one principle challenge for visual tracking is to develop an effective representation model which can adapt to such appearance variations.

In most existing tracking algorithms, the appearance model is either fixed, slowly changing, or rapidly changing. These tracking algorithms can be divided into three categories: tracking with a fixed appearance model, tracking

with an adaptive appearance model, and tracking with a two-frame appearance model. In the first tracking algorithm, the appearance model is extracted from the first frame and kept unchanged, such as color models in [1,2], and gray-level template in [3]. In the second, the appearance model is obtained by slowly updating the object template, such as color model in [4]. In the third, the appearance model is updated every frame (or every *n* frames) using the tracking results from the previous frames, such as gray-level templates in [5,6]. Although the three kinds of the tracking algorithms can track object well in some cases, the tracking robustness is still greatly limited. If the appearance model is fixed during tracking, the tracker easily loses object in the presence of illumination changes, pose variations, or occlusions. If the appearance model is slowly updated, the appearance model can adapt to slow changes in object appearance, but it cannot deal with the sudden or rapid changes in object appearance, moreover, once the appearance model is improperly updated, this kind of tracker may completely fail. However, if the

* Corresponding author. Tel.: +86 21 62932119x203;
fax: +86 21 62932120.

E-mail address: lapjt@sjtu.edu.cn (A. Li).

appearance model is rapidly updated, using the tracking results from the previous frames, the tracker may drift away from object due to occlusions or the accumulation tracking errors from the previous tracking steps. Therefore, it is very important to design a robust representation model which can adapt to slow or rapid changes in object appearance.

In this paper, we propose a robust observation model. This model combines the above three appearance models together. In tracking objects, the robust observation model is incorporated in a particle filter. The particle filter automatically selects proper appearance models for the current tracking situation. Occlusion is handled by the M -estimation technique. Experimental results demonstrate the effectiveness and robustness of our proposed algorithm.

The remainder of this paper is organized as follows. In Section 2, we detail the object model, with a special focus on the robust observation model. Section 3 shows occlusion handling. Particle filter tracker with the robust observation model is given in Section 4. Experimental results and conclusions are presented in Sections 5 and 6.

2. Object model

2.1. Motion model

The motion of an object between two consecutive frames can be represented by an image warping of a similar transformation. We use four parameters of a similar transformation to model the state transition from X_{t-1} to X_t . Denote X_t by $X_t = [u_t, v_t, s_t, \theta_t]^T$, where u_t , v_t , s_t and θ_t correspond to x , y translation, scale, and rotation angle at time t , respectively. Each parameter in X_t is modeled independently by a Gaussian distribution around its previous value in X_{t-1} , namely

$$P(X_t|X_{t-1}) = N(X_t; X_{t-1}, \Sigma), \quad (1)$$

where Σ is a diagonal covariance matrix whose elements are the corresponding variances of the similar transformation parameters, i.e. σ_u^2 , σ_v^2 , σ_s^2 , σ_θ^2 .

2.2. Robust observation model

The observation model consists of the three appearance models. We first explain how observations are obtained under the given states.

2.2.1. Observation representation

Given the object state X_t , then the object region at time t can be determined by the following similar transformation:

$$\begin{bmatrix} P_{X,t} \\ P_{Y,t} \end{bmatrix} = s_t \begin{bmatrix} \cos(\theta_t) & -\sin(\theta_t) \\ \sin(\theta_t) & \cos(\theta_t) \end{bmatrix} \begin{bmatrix} P_{X,t-1} \\ P_{Y,t-1} \end{bmatrix} + \begin{bmatrix} u_t \\ v_t \end{bmatrix}, \quad (2)$$

where $[P_{X,t} P_{Y,t}]^T$ is the pixel point in the current object region, $[P_{X,t-1} P_{Y,t-1}]^T$ is the pixel point in the object region at time $t-1$. If let φ denote the similar transformation function and $P_{t-1} = [P_{X,t-1} P_{Y,t-1}]^T$, the observation of the object region at time t can be represented as

$$Z_t = I_t(\varphi(P_{t-1}; X_t)), \quad (3)$$

where $I_t(\varphi(P_{t-1}; X_t))$ is the pixel gray value observed in the current image at the pixel point $\varphi(P_{t-1}, X_t)$.

2.2.2. Robust observation likelihood

The observation likelihood is designed based on the three appearance models. We denote the three appearance models at time $t-1$ by $\{F_{t-1}, A_{t-1}, D_{t-1}\}$, where F_{t-1} , A_{t-1} , and D_{t-1} , respectively, denote the fixed appearance model, the adaptive appearance model, and the two-frame appearance model. Assume that the pixels in object appearance are independent of each other, the observation likelihood function over the state X_t is designed as

$$p(Y_t|X_t) = p(Z_t|X_t, M_t) = \begin{cases} \prod_{i=1}^d N(Z_t(i); F_{t-1}(i), R_{F_{t-1}}(i)) & \text{if } M_t = F_{t-1}, \\ \prod_{i=1}^d N(Z_t(i); A_{t-1}(i), R_{A_{t-1}}(i)) & \text{if } M_t = A_{t-1}, \\ \prod_{i=1}^d N(Z_t(i); D_{t-1}(i), R_{D_{t-1}}(i)) & \text{if } M_t = D_{t-1}, \end{cases} \quad (4)$$

where M_t represents the selected appearance model at time t , which is determined by particle filter, i is the i th pixel, d is the number of pixels in the object appearance. $R_{F_{t-1}}(i)$, $R_{A_{t-1}}(i)$, and $R_{D_{t-1}}(i)$ are the observation noise covariance matrixes of the i th pixel in F_{t-1} , A_{t-1} , and D_{t-1} , respectively. They are adapted over time. We will discuss them later. $N(x; \mu, R)$ is a Gaussian density function

$$N(x; \mu, R) = (2\pi\|R\|)^{-1/2} \exp \left\{ -g \left(\sqrt{(x - \mu)^T R^{-1} (x - \mu)} \right) \right\}, \quad (5)$$

where $g(x) = x^2/2$, $\|R\|$ is the determinant of the covariance matrix R . In most existing algorithms, the above observation likelihood function is designed based on single appearance model, i.e. F_{t-1} , A_{t-1} , or D_{t-1} . The disadvantages of using single appearance model have already been discussed in Section 1. To increase the robustness of the observation model, we design the observation likelihood function above, in which the three appearance models are used. Among the three appearance models, the adaptive appearance model is obtained by smoothing the object appearance using the Kalman filter, which is similar to the method used

Download English Version:

<https://daneshyari.com/en/article/447962>

Download Persian Version:

<https://daneshyari.com/article/447962>

[Daneshyari.com](https://daneshyari.com)