Optik 126 (2015) 1830–1837

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Learning based particle filtering object tracking for visible-light systems

Wei Sun [a,*], Chunyu Zhao [a], Long Chen [b], Dajian Li [c], Yicheng Bai [e], Wenyan Jia [e], Mingui Sun [d]

[a] School of Aerospace Science and Technology, Xidian University, No. 2 Tabai Rd., Xi'an 710071, China
[b] Xi'an Research Institute of China Coal Technology & Engineering Group Corp, No. 82 Jinye 1st Road, 710077, China
[c] Northwest Polytechnical University, The 365 Insititute, No. 127 Youyixi Road, Xi'an 710072, China
[d] Departments of Neurosurgery and Electrical & Computer Engineering, University of Pittsburgh
[e] Electrical & Computer Engineering, University of Pittsburgh, PA 15260, USA

ARTICLE INFO

Article history:
Received 17 March 2014
Accepted 3 May 2015

Keywords:
Object tracking
Support-vector machine
Particle filtering
Online learning scheme
eButton

ABSTRACT

We propose a novel object tracking framework based on online learning scheme that can work robustly in challenging scenarios. Firstly, a learning-based particle filter is proposed with color and edge-based features. We train a support vector machine (SVM) classifier with object and background information and map the outputs into probabilities, then the weight of particles in a particle filter can be calculated by the probabilistic outputs to estimate the state of the object. Secondly, the tracking loop starts with Lucas–Kanade (LK) affine template matching and follows by learning-based particle filter tracking. Lucas–Kanade method estimates errors and updates object template in the positive samples dataset, and learning-based particle filter tracker will start if the LK tracker loses the object. Finally, SVM classifier evaluates every tracked appearance to update the training set or restart the tracking loop if necessary. Experimental results show that our method is robust to challenging light, scale and pose changing, and test on eButton image sequence also achieves satisfactory tracking performance.

1. Introduction

Visible light systems are widely employed in vision-based applications because of the low cost. In general, visual object tracking is the process of location estimation for one or multiple objects in videos. It is required by many applications such as visual navigation, human-computer interfaces, video communication, compression and surveillance. Key requirements imposed on a robust tracker can be given as following: (1) being able to track arbitrary targets and tracking accurately through challenging conditions, (2) tracking in real-time and capturing of the object when it reappears in the camera's field of view. So, it is a great challenge to design robust features and tracking methods which can cope with variations in natural scenes, such as changes in illumination, shape, viewpoint or partial occlusion of the target.

In the past decades, various features like points, articulated models, contours, or optical flow [1,2] are used, and many tracking methods based on image patch or color histograms are proposed [1]. In [3], in which color and local binary pattern (LBP) histograms are used to build the objects model, these features result in a system less sensitive to illumination changes and partial occlusion. But these generative model trackers often fail in cluttered background. Recently, particle filter [4] has been successfully applied to solve non-linear and non-Gaussian tracking problems, and some novel particle filter have been proposed in [7,20]. Besides these achievements, it is still a challenge on how to build an adaptive model of the target's appearance which can be generalized to possible future appearances. Several research groups have carried out investigation on this open question and some methods have been proposed. Han et al. [6] introduced a mean-shift based sequential kernel density approximation technique to update a target appearance model online. However, the addition and deletion of a Gaussian component highly depend on the pre-defined threshold values.

So, for handling appearance changes and coping with background clutter [15], a learning scheme with a classifier that represents the decision boundary between the object and its background may be the best choice. Avidan et al. [5] presented an algorithm to adapt the constituent parts and combined an ensemble of classifiers itself to new appearances. Lei et al. [7] tried to adapt an off-line learning ensemble classifier of a particular tracked object

* Tel.: +86 13991805685.
E-mail address: wsun@xidian.edu.cn (W. Sun).

http://dx.doi.org/10.1016/j.ijleo.2015.05.018
0030-4026/

for the changing appearance. Despite their success, these methods are limited by the fact that they cannot accommodate very large changes in appearance.

In recent years, many discriminative trackers [9–12] that are capable to track, learn and detect have been proposed. Grabner et al. [8] demonstrated a semi-supervised on-line learning scheme to tackle the problem of uncertain class assignments of training examples. Babenko et al. [13] suggested a Multiple Instance Learning (MIL) approach for object detection. However, if the tracked location is not precise, the appearance model ends up getting updated with a sub-optimal positive example. Over time this can degrade the model, and can cause drift. Kalal et al. [15] formulated tracking as an online learning problem. Combining with a prior classifier, this method takes all incoming samples as unlabeled and uses them to update the tracker. But it throws away global features of the object. It also suffers from drift and fails if the object leaves the scene for longer than expected.

To deal with the problems discussed above, we present an approach in which template tracking and learning-based particle filtering detection are independent processes that exchange information by learning mechanisms. Firstly, we build the joint color-texture histogram to represent an object reference model, and propose a learning-based particle filter, using the probabilistic output of the SVM classifier to approximate the weight of the particles. Secondly, the classifier of initial target appearance model is quickly learned from the first frame and it will be used to detect the probability distribution of the object by learning-based particle filter in the following frames. Thirdly, to update the object model online and overcome the position drifting problem, we combine template matching and learning-based particle filtering tracking techniques, they work together and supplement each other, and the switching scheme will be discussed in detail in the following sections.

The rest of the paper is organized as follows. Section 2 gives an overview of particle filter based visual tracking system. Section 3 introduces our learning-based object model and describes how it is adapted. Section 4 proposes learning-based tracking framework. Section 5 presents a number of comparative experiments and analysis, which is followed concluding remarks and future work in Section 6.

2. The scheme of classic particle filtering algorithm

Over past few years, condensation algorithm, also known as particle filter, have proved to be a powerful tool for image tracking [3–5]. In [4], an important advantage of particle filtering frameworks has been proposed and the strength of these methods lies in their simplicity, flexibility, and systematic treatment of nonlinearity and non-Gaussianity.

The main idea of the particle filter algorithm is that post probability distribution of the tracked object over state $X = (x, y, w, h)^T$ can be approximated by a set of weighted particles $S_t = \{s_t^j\}, j \in \{1 \cdots J\}$, where (x,j) specifies the center of the tracked object's position in the image, and (w, h) are the dimensions of the target rectangle. Each particle $s_t^j = (x_t^j, \pi_t^j)$ consists of its state vector x_t^j and an importance weight π_t^j. The set of particles is updated from one frame to the next by the following recursive procedure:

Firstly, a new sample set S_t is drawn with replacement from the previous set S_{t-1}, where a sample s_{t-1}^i from the old set is chosen with probability proportional to its weight π_{t-1}^i. Secondly, for each sample x_t^j, a new state is determined by updating the importance weights π_t^j with the likelihood of the observation

$$\pi_t^j = p(Z_t | X_t = x_t^j, Z_0, Z_1 \cdots Z_{t-1}) \tag{1}$$

where the likelihood depends on all frames $Z_0, Z_1 \cdots Z_{t-1}$, because the observation model is adapted over time. The sample set is propagated through a dynamic model given by

$$s_t = A \cdot s_{t-1} + w_{t-1} \tag{2}$$

where A defines the deterministic component of the model and w_{t-1} is a multivariate Gaussian variable. In our application, we use a first order model for A describing a region moving with constant velocities v_x, v_y.

The state of a target is estimated by the weighted average over states of the particles,

$$(\bar{x}, \bar{y}, \bar{w}, \bar{h})^T = \sum_{j=1}^{J} \pi_t^j \cdot (x_t^j, y_t^j, w_t^j, h_t^j)^T \tag{3}$$

So, we need to compute the likelihood $p(Z_t | X_t = x_t^j, Z_0, Z_1 \cdots Z_{t-1})$ to get the weight of the particles. Using color features, the target is tracked by comparing its histogram with the histograms of the sample set using Bhattacharyya distance, for discrete densities of color histograms $p = \{p^{(u)}\}\ u = 1 \ldots m$ and $q = \{q^{(u)}\}\ u = 1 \ldots m$, the Bhattacharyya distance is defined as

$$\rho[p, q] = \sum_{u=1}^{m} \sqrt{p^{(u)} q^{(u)}} \tag{4}$$

The larger $\rho$is, the more similar the distributions are. According to [2], particle filters may perform poorly when the posterior is multimodal as the result of ambiguities or multiple targets. To keep up with the changes in real-world scenarios, it is best not to rely on a fixed target model, but to adapt the model over time. In this way, spatial target properties can be updated.

3. Learning-based particle filter for object tracking

3.1. Analysis of target appearance model

As discussed above, illumination conditions, visual angle, as well as the camera parameters can influence the tracking result of particle filter. We do some statistical research on the TLD dataset of 10 video sequences [15], among these videos, the 'David' video has been used in several recent published papers, since it could present challenging lighting, scale and pose changes. Fig. 1 shows snapshots from 'David' and Fig. 2 lists RGB and LBP histograms of object. Color histogram is calculated in RGB space for each color channel (R, G and B) [4]. LBP is a texture operator with a low computational complexity that describes an object's local structure [6,7]. The object in this sequence is manually annotated with bounding boxes. In Fig. 2(a) and (c), olor and texture histograms of the object in the bounding boxes are various over time. In Fig. 2(b) and (d), Bhattacharyya distance between the initial object histogram and other image frames decreases sharply.

Hence, we get the following conclusions: to cope with appearance variations of the target object during tracking, the model of the object should be adaptive over time. However, perfect target models cannot be build off line [13,14], since the appearance of object is not known in advance. So, we can build a binary classifier that represents the decision boundary between the object and its background to substitute Bhattacharyya distance, and we update the target model by retraining the classifier which is able to incrementally adjust to the changes in the specific tracking environment.

3.2. Learning-based particle filter

In our case, we will propose a leaning-based particle filter. Instead of Bhattacharyya distance, we determine weight value for each particle by a binary classifier. Support vector machines (SVMs,