Neurocomputing 204 (2016) 97–105

Contents lists available at ScienceDirect

Neurocomputing

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

Adaptive structured sub-blocks tracking ☆

Jing-Wen Liu, Wei-Ping Sun *, Tao Xia

School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China

ARTICLE INFO

Article history:
Received 1 March 2015
Received in revised form
30 September 2015
Accepted 1 October 2015
Available online 8 April 2016

Keywords:
Visual object tracking
Structured sub-blocks
Particle filter

ABSTRACT

Visual object tracking algorithms based on middle level appearance have been widely studied for their effective representation to non-rigid appearance variation and partial occlusion. Sub-blocks are often adopted as local feature in mid-level based tracking algorithms. How to select representative sub-blocks to reveal the spatial structure of objects and retain the flexibility to model non-rigid deformation has not been adequately addressed. Exploiting discrimination, uniqueness and historical prediction accuracy of sub-blocks of a target, we propose a local feature selection method which includes rough initial subblock selection and refined subblock-sample particle bi-directional selection under particle filter tracking framework. A quantitative evaluation is conducted on 10 sequences. Experimental results show the robustness of our proposed algorithm in tackling with non-rigid deformation and partial occlusion.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Visual object tracking is a fundamental task in computer vision, applied in a wide range of domains such as intelligent surveillance, human computer interaction, video compression, and so on. Generally a visual object tracker includes four modules: object initialization, appearance modeling, motion estimation and object localization, among which robust adaptive appearance modeling could play an important role [1].

Recently tracking algorithms based on middle level appearance cues have been widely studied. Compared with global visual representations, local visual representations are less sensitive to non-rigid appearance variation and partial occlusion. Among these outstanding trackers, Adam et al. [2] design an appearance model composed of several image fragments, where fragments are obtained in a predefined way with no selection and every fragment votes on the possible positions and scales of the target. Kwon et al. [3] propose an appearance model which is comprised of multiple local patches and the topology between those patches. The appearance model also need no specific object model and is more flexible. Researchers are interested in finding novel ways to choose the most informative local features with respect to tracking task and local features are selected for specific purposes. Nejhum et al. [4] use multiple rectangular blocks to model the constantly changing foreground shape. Yang et al. [5] propose an attentional tracking method AVT based on attentional patches. In AVT, the early selection and the late selection processes are adopted to obtain more discriminating local patches gradually. Yang et al. [6] conduct a discriminative appearance model based on segmentation from the perspective of mid-level cues and superpixels are adopted for structural information. Based on the multiple instance learning (MIL), Zhang [7] proposes an online feature selection algorithm with prior information of instance labels. Moo Yi et al. [8] describe object with feature points, select local features with motion saliency and descriptor saliency and suppress degradation caused by inaccurate initializations. These feature selection algorithms always focus on information got from historical frames and in some application appearance in current frame could provide more accurate and effective information for local features.

Motivated by the above observation, under particle filter tracking framework we propose a novel local feature selection method taking current sampling state into account and an adaptive structured sub-blocks tracking (ASST) algorithm is implemented. A large amount of particles in current frame naturally generated in particle filter tracker cannot only be used as target candidates, but also can be viewed as important reference to select local features because the "quality" of samples in current frame exhibits the good or bad of the tracking result in previous frame in some degree. What is more, when calculate the similarity between each sample and the target model we need only take part of local features into account since some local features in samples maybe disturbed by noise or occlusion. So in our proposed tracker, rough initial sub-block selection and refined sub-block-sample particle bi-directional selection are utilized. The method can select more

☆This article is supported by the National Natural Science Foundation of China (No. 61300140), the Natural Science Foundation of Hubei Province (No. 2015CFB565) and the Fundamental Research Funds for the Central Universities, HUST: No. 2014QN008.

* Corresponding author.
E-mail addresses: M201272672@hust.edu.cn (J.-W. Liu), wpsun@hust.edu.cn (W.-P. Sun), xiatao@hust.edu.cn (T. Xia).

http://dx.doi.org/10.1016/j.neucom.2015.10.133
0925-2312/© 2016 Elsevier B.V. All rights reserved.

representative sub-blocks to reveal the current spatial structure of tracking target and model non-rigid deformation flexibly. Several measures are adopted in our local feature selection processes.

2. Structured sub-blocks tracking framework

Given the position of the target in one frame, a tracker should be able to track the target in subsequence frames. Tracking based sub-blocks can be viewed as matching based on similarity measuring. In order to narrow the gap between low level features and high level semantic of the images, those local feature (i.e. sub-blocks) and sample particles which will take part in final decision should be considered seriously. We model the decision-making in tracking as local feature selection in current frame. Fig. 1 illustrates the basic pipeline of our tracker.

The target for tracking is firstly divided into overlapped sub-blocks. Some of them are chosen to add to a candidate subblock set through initial selection process. Two measures, *discrimination* and coverage, are taken into consider during initial selection. Sub-blocks in candidate set then are ordered by *historical prediction accuracy*.

When a new frame (frame t in Fig. 1) arrives, a two-stage selection process, one is on sub-blocks and another is on sample particles, is performed. In first stage, sub-blocks in candidate set are re-chosen according their "quality" distribution among sample particles and local feature which are more representative to reveal the current spatial structure of target under current appearance are chosen as templates of frame t. In second stage, "good" particles are filtered to join final decision-making. In this refined bi-directional selection *discrimination* and *uniqueness* of sub-blocks and *confidence* of particles are adopted as measures to help to adapt non-rigid deformation flexibly.

After "good" local feature and "good" particles are obtained, localization of frame t is conducted by a simple similarity measuring. The particle with high similarity with templates of frame t is output as the tracking result of frame t.

As mentioned above, some measure will be adopted in our tracking and we list them below.

(1) Coverage ratio
The coverage ratio of sub-block i and j (which denoted by p_i, p_j respectively) is defined as

$$CR(p_i, p_j) = \frac{Area_{p_i} \cap Area_{p_j}}{Area_{p_i} \cup Area_{p_j}} \tag{1}$$

Here $Area_{p_n}$ is the area of the rectangularity of sub-block p_n.

(2) Discrimination and uniqueness
Discrimination which measured by the maximum margin between positive and negative training samples shows the ability one sub-block could distinguish target from background. Uniqueness indicates whether a sub-block is unique to target object, that is to say, if the sub-block can distinguish the tracking target from other similar objects.
Let D_i^t and U_i^t be the discrimination and uniqueness respectively for sub-block i ($i \in \{1, 2, ..., N\}$) at time t. Denote $\tilde{X}_i^t$ and $\overline{X}_i^t$ the positive and negative training samples of sub-block i at time t respectively, $X_{i,j}^t$ the jth sample particle at time t. D_i^t and U_i^t are defined as follows.

$$D_i^t = f\left(\tilde{X}_i^{t-1}, \overline{X}_i^{t-1}, X_{i,j}^t\right) = \max_j \left(P\left(\tilde{X}_i^{t-1}, X_{i,j}^t\right) - P\left(\overline{X}_i^{t-1}, X_{i,j}^t\right)\right) \tag{2}$$

$$U_i^t = 1 - \frac{1}{M}\sum_{j=1}^{M} g\left(\tilde{X}_i^{t-1}, \overline{X}_i^{t-1}, X_{i,j}^t, \tau\right). \tag{3}$$

where $j = 1, 2, ..., M$ and M is the number of sampling particles in particle filter tracking. $P(\cdot)$ is histogram intersection coefficient and $g(\cdot)$ is defined as

$$g\left(\tilde{X}_i^{t-1}, \overline{X}_i^{t-1}, X_{i,j}^t, \tau\right) = \begin{cases} 1, P\left(\tilde{X}_i^{t-1}, X_{i,j}^t\right) - P\left(\overline{X}_i^{t-1}, X_{i,j}^t\right) \geq \tau \\ 0, P\left(\tilde{X}_i^{t-1}, X_{i,j}^t\right) - P\left(\overline{X}_i^{t-1}, X_{i,j}^t\right) < \tau \end{cases}, \tag{4}$$

where $\tau = D_i^t - \Delta\tau$. $\Delta\tau$ is a parameter to measure the similarity of confusing particles (which refer to sampling particles having similar appearance with target object) and optimum particle. Larger confusing particle number is, more sampling particles have similar appearance within sub-block area, which indicates this sub-block will be unhelpful in subsequent localization.

(3) Historical prediction accuracy
In two-stage selection sub-blocks are operated one by one to get template for frame t. For convenience we will order sub-blocks by their historical prediction accuracy before the selection and only those sub-blocks in candidate set with higher historical prediction accuracy will be take into account. Historical prediction accuracy HP_i indicates historical

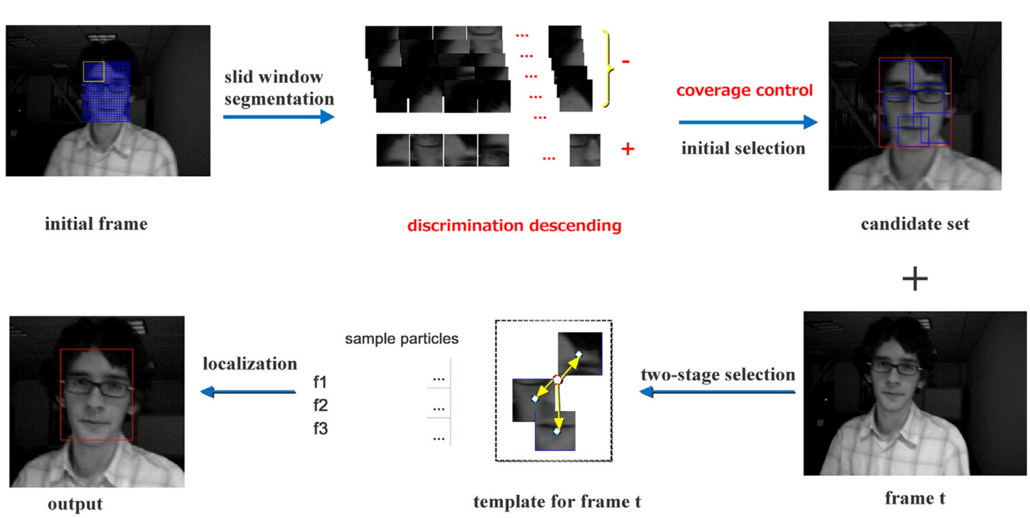

Fig. 1. Pipeline of our tracking algorithm ASST.

Download English Version:

https://daneshyari.com/en/article/405788

Download Persian Version:

https://daneshyari.com/article/405788

Daneshyari.com