Neurocomputing ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Neurocomputing

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

Parallel search strategy in kernel feature space to track FLIR target

Zhen Shi, Chang'an Wei *, Junbao Li, Ping Fu, Shouda Jiang

Department of Automatic Testing and Control, Harbin Institute of Technology, Harbin 150080, China

ARTICLE INFO

Article history:
Received 18 September 2015
Received in revised form
31 May 2016
Accepted 17 June 2016
Communicated by Yue Gao

Keywords:
Sparse representation
Kernel method
FLIR target tracking

ABSTRACT

Robust tracking in the forward looking infrared (FLIR) sequences is still a challenging problem in the field of computer vision. Because images acquired by the infrared sensors are characterized by low signal-to-clutter ratios (SCR) and targets of interest may exhibit profound appearance variations due to ego-motion of the sensor platform and complex maneuvers. Though many efforts have been delivered, there are still some issues to be addressed. First, intensity features are not enough to deal with complex appearance variations for the challenging sequence. Second, to obtain satisfying estimation of target state, a plenty of particles have to be employed to approximate its probability density function (pdf). To deal with the two problems, a parallel search strategy based on kernel sparse representation (KL1PS tracker) is proposed to perform the tracking task in the FLIR sequences. With the ability of capturing the nonlinear features, kernel method is introduced to deal with complex appearance variations. After the kernel function is constructed based on the histogram features, both the target templates and candidates are mapped into the kernel feature space. Then efficient state particles are selected based on sparse representation which can be used to estimate the target state. The proposed method is tested on the AMCOM database, and the experimental results demonstrate its excellent performance in tracking accuracy compared with some state-of-the-art trackers.

1. Introduction

Recently, the production cost of infrared sensors is reduced obviously due to technological advancements, thus it can be employed in more application fields. For infrared images, the intensity of an object is determined by its temperature and radiated heat which is not influenced by illumination conditions and surface features. Specially, attention has been specifically devoted to infrared sensors which are sensitive to long-wave infrared (LWIR) light in the range 8–12 nm. It is very useful in many practical applications because they can see heat sources at night or through smoke, fog, haze, etc. [1].

Object tracking is a fundamental task which has been well studied in the last decades and numerous approaches have been proposed [2]. Generally speaking, it can be categorized into two types. The first category is free tracking, where initialization of the target on the first frame is required, and tracking is carried out based on information in current and past frames [3]. The second category is specific tracking, in which a pre-trained object detector is needed to obtain the desired target, and information in both past and future frames can be used for tracking [4]. For practical application, the target in FLIR sequence is desired to be locked since it has been detected, thus it belongs to the first category obviously. In fact, target tracking has remained a challenging problem in the field of computer vision and special treatments are necessary to design a tracker for FLIR image sequences [5]. In this paper, two issues are discussed about tracking a target in the FLIR sequences which are stated as follows.

One important issue in the tracking system is target representation. For a target of interest in the complex scenes, its appearance may exhibit profound variations over relatively short time scales. Intrinsic factors include pose change and shape deformation while extrinsic factors include illumination change, ego-motion of camera and partial occlusions. In fact, it is still a challenging problem to deal with the unexpected appearance variations. Obviously, it is not reliable to track a target in FLIR sequences only with intensity information, and more sophisticated techniques are required for representing the target appearance to deal with various challenging problems. It is desired to be robust to pose change, shape deformation, illumination change, partial occlusion, etc.

The other issue is target state estimation in the particle filter framework. At present, there are two ways to estimate the target state after the particles with associated weights are available. For the first method, the most similar state particle is taken as the result directly. In fact, the target state is a continuous variable whose pdf is approximated with discrete particles. Obviously, a reliable estimation can be obtained only when the particle number

* Corresponding author.
E-mail address: weichangan@hit.edu.cn (C. Wei).

http://dx.doi.org/10.1016/j.neucom.2016.06.052
0925-2312/

Please cite this article as: Z. Shi, et al., Parallel search strategy in kernel feature space to track FLIR target, Neurocomputing (2016), http://dx.doi.org/10.1016/j.neucom.2016.06.052

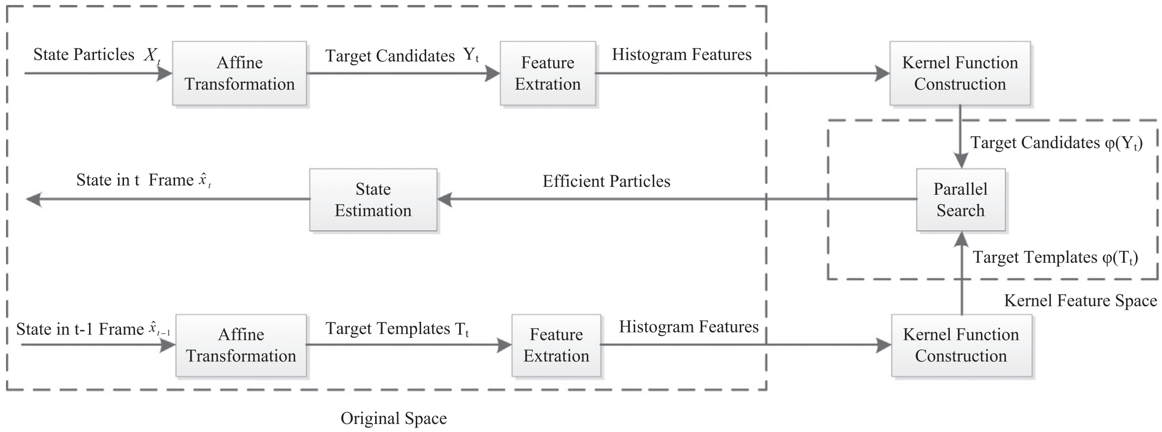

Fig. 1. Block diagram of KL1PS tracker.

is large enough. For the other method, the target state is estimated by all the weighted particles. However, not each candidate contains the target while background is dominate in the false candidates. In fact, only those candidates containing the target are corresponding to the efficient state particles. When the target state is estimated with all the weighted particles, it is inevitable to introduce error which will cause drift problem.

To solve the two issues mentioned above, a parallel search strategy is proposed in the kernel feature space. First, kernel method is introduced to deal with complex appearance variations. After the kernel function is constructed based on the histogram features, both the target templates and candidates are mapped into the kernel feature space. Then efficient state particles are selected based on sparse representation which can be used to estimate the target state. The block diagram of proposed method is shown in Fig. 1.

Compared to exciting literature, the distinctive contributions of this paper are summarized as follows.

1. For a target with profound variations, we propose to model its appearance with kernel method which is efficient to capture the nonlinear similarity of features. Considering the target characteristic in infrared images, the kernel function is constructed with histogram features which are not sensitive to part occlusion, illumination variation and shape deformation.

2. When the target state pdf is approximated with weighted particles, it is not reliable to estimate its state only with the most similar one. Moreover, it is inevitable to introduce error when all the particles are employed. Considering the present methods, we propose to parallel search for the efficient particles based on sparse representation which are fused to obtain the target state. In this way, the estimation accuracy can be improved and satisfying tracking performance can be obtained.

The rest of this paper is organized as follows. The related works is reviewed briefly in Section 2. After the principle of particle filter is introduced in Section 3, a parallel search strategy in the kernel feature space is proposed in Section 4. In Section 5, the solution to kernel sparse representation is obtained based on constructed kernel matrixes and vectors. Experiments results are given in Section 6 then the paper is concludes in Section 7.

2. Related work

There are a plenty of methods to represent the target appearance in the previous work. The sum of squared difference (SSD) is used as a cost function in the tracking problem [6] and the mixture model is used to model the appearance variation [7]. An appearance-adaptive model is incorporated in the particle filter framework to realize robust visual tracking and classification algorithms [8]. The target appearance can also be represented with an eigenspace [9], affine warps of learned linear subspaces [10], or an adaptive low dimensional subspace [11]. Besides, the appearance is modeled as a structured combination of the subspaces learned by partial least squares analysis [12]. Due to the powerful feature learning characteristic of deep learning, target appearance can be observed by multiple sub-regions and each region is observed by a deep learning model [13]. Histogram of the pixel intensities has been adopted in the appearance models of several recent mean-shift trackers [14]. Besides, histogram of the local standard deviation has also been used for mean-shift tracking of infrared targets [15]. The histogram features are widely used due to their simplicity and efficiency as well as the scale and rotation invariance properties. For the target representations based on histogram, appearance learning is generally accomplished by iteratively updating a reference histogram. An extension of the simple histogram-based appearance learning strategies maintains explicit appearance models for both the target and the surrounding background which has been used to combat the drifting problem [16,17].

Generally speaking, there are two categories of methods to search for the target. In the first category, tracking is formulated in the optimization framework with an objective function, and gradient descent methods can be used for locating the target. For example, the first order Taylor expansion is used to linearize the nonlinear cost function, and the motion parameters are estimated iteratively [18]. Further, mean shift method is used for searching the target by using the Bhattacharyya coefficient as the similarity metric with kernel-regularized histograms [14]. A discriminative approach is adopted to identify spatial attentional regions with a gradient-based formulation to locate objects [19]. A tracking algorithm based on distribution fields is proposed which allow smoothing the objective function without blurring the image, and the target is located by searching for the local minimum by using a coarse-to-fine strategy [20]. However, objective functions for tracking are usually nonlinear with many local minima. To alleviate this problem, dense sampling methods have been adopted at the expense of a high computational load [21–23]. Thus stochastic search algorithms such as particle filter have been widely used since they are relatively insensitive to the local minimum and are computationally efficient [24,25]. Recently, methods based on particle filter have been developed