Optik 143 (2017) 26–34

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

An extended target tracker based on structural appearance and improved distribution fields for different scenarios

Pei Lu [a], Xiaoyong Liu [b,*], Min Tian [a], Xin Lu [c]

[a] College of Information Science and Technology, Shihezi University, Shihezi, 832000, China
[b] Key Laboratory of Eco-physics and Department of Physics, College of Science, Shihezi University, Shihezi, 832000, China
[c] Key Laboratory of Oasis Eco-agriculture, Xinjiang Production and Construction Group, Shihezi, 832000, China

ARTICLE INFO

Article history:
Received 2 March 2017
Accepted 2 June 2017

Keywords:
Visual tracking
Extended target
Structural appearance
Distribution fields

ABSTRACT

An extended target tracker based on structural appearance and improved distribution fields is proposed in terms of current issues that there is no one general tracking algorithm to different scenarios. For extended target with specific geometric structure and pose variation in simple scenarios, in order to get stable tracking point, we construct a structural appearance model to represent target and utilize triangle and line character to obtain tracking point. In complex scenarios with certain texture information and image blurring, to reduce computational complexity and enhance the adaptivity of tracking box, an improved distribution fields is presented in which non-uniform delamination technology is proposed and the BRISK is used to detect key points in tracking box. Numerical experiments on simulation sequences and public database were provided to demonstrate the good performance of this proposed scheme. The conclusion can be drawn that triangle achieves the best stability among geometric graphics and skeleton can reflect the intrinsic geometric structure of the target in simple scenarios. Moreover, non-uniform delamination technology can reduce computational complexity and the BRISK can achieve affine transformation parameters in complex scenarios, so the tracking box can adapt to scale and rotation changes when matching is performed.

1. Introduction

Despite that numerous algorithms have been proposed in the literatures [1–11], target tracking remains a challenging problem due to appearance change caused by pose, motion, occlusion, and illumination, among others.

In optical imaging system, with the increasing caliber and focus of optics system, the object's size is larger compared with system's small field-of-view. In this case, the object is extended. Extended target tracking has been extensively studied in recent years and there are a few literatures exploiting it, such as Monte Carlo methods [12,13], Bayesian tracking [14], particle filtering [15], random matrix [16–19], ellipse fitting based approach [20] and other methods [21–24]. In general, these methods compute motion trajectory in whole process from a viewpoint of data association. They all consist in both size and state estimation of an extended target mathematically when tracking, so the object was as a whole which is not the case in real application. In addition, the common extended target tracking methods used in practical application are matching [25–27]. Due to the motion, background and illumination disturbances, as well as the minimum computation

* Corresponding author.
E-mail address: llxxyy1017@qq.com (X. Liu).

http://dx.doi.org/10.1016/j.ijleo.2017.06.009
0030-4026/

unit's precision in image processing, the absolute optimal matching location cannot be obtained and tracking may fail under pose variation. As a result, most methods are the same with specific type of object in certain scenario.

The main purpose of this paper is to develop a general visual tracker for extended target according to different scenarios. Firstly, the process of scene selection is performed for input image sequences according to several low level visual features such as gray scale and texture distribution, shape, gratitude etc. If simple scene is to be judged, then a structural appearance model based on skeleton and triangle theorem is constructed to represent target and triangle and line character is utilized to obtain tracking point. The theory behind this is that triangle achieves the best stability among geometric graphics. Moreover, skeleton can reflect the intrinsic geometric structure of the object, especially for extended target with specific geometric structure. Otherwise, in complex scenarios with certain texture information and image blurring, target segmentation has become difficult, which resulted in difficulty in modeling appearance. Generally, visual tracking relies on the assumption that gradient descent of the alignment function will reach the global optimum. As a result, distribution fields (DFs) [28] is developed to smooth the objective function without destroying information about pixel values and also allow the representation of uncertainty about the tracked target. However, the smoothed image is divided into 16 layers uniformly and the high computational complexity and large storage are unavoidable for the target with intensive gray scales, which may not satisfy real-time requirement. So we proposed an improved DFs, where non-uniform delamination technology is utilized according to the image's gray scale statistic information. In addition, to address the problem that tracking box cannot adapt to scale and rotation changes, the BRISK [29] is used to detect key-points in tracking box obtained by improved DFs, and then matching is performed during consecutive frames. Consequently, warp parameters can be calculated via matching and scale and rotation change parameters are obtained and used to provide adaptation.

The contributions in our work are: 1) a structural appearance model based on skeleton and triangle theorem is constructed; 2) improved DFs is proposed to divide the smoothed image adaptively according to its gray scale statistic information, and the tracking box can adapt to scale and rotation changes.

The reminder of the paper is organized as follows. Two parts of the proposed tracker will be presented in detail in section 2. Section 3 shows the experimental results of the proposed tracker. Finally, brief conclusions are drawn in Section 4.

2. The proposed tracker

2.1. Simple scene

2.1.1. Skeleton extraction

Before skeleton extraction, the image frame should be segmented to achieve binary image. In fact, image segmentation is a process to category pixels based on some attributes. The complexity of nature images determines the problem that grouping which pixels into which cluster, thus it is natural to consider image segmentation in the view of fuzzy cluster. So Fuzzy C-Means Cluster (FCM) [30] is utilized, which is a nonlinear iteration optimal algorithm based on objective function in essence, and the objective function is defined as:

$$J(U, c_1, \ldots, c_c) = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m d_{ij}^2 \tag{1}$$

Where U denotes fuzzy subjective matrix with $c \times n$, u_{ij}^m is the element of the matrix and ranges from 0 to 1, m is weighted coefficient, c_i is the center of the i-th cluster, n represents the data number, $d_{ij} = \|c_i - x_j\|$ denotes the Euclidean distance between the i-th cluster center and the j-th data. Through iteration, rational fuzzy subjective matrix and cluster centers are chosen to minimize the objective function. Consequently, the optimal segmentation is obtained.

After segmentation, skeleton extraction can be done using morphological thinning. Because of high computation using Medial Axis Transform, this work adopts iteration thinning algorithm to remove boundary points gradually. After getting skeleton points, linear fitting technique using skeleton points is adopted to compute objects' main axis and the main skeleton of extended target can be obtained finally.

2.1.2. Appearance modeling

Since the main axis can reflect the object's primary geometric structure which remains stable during the whole tracking process, we model the appearance of the object with a structural geometric triangle and two theorems are used as follows.

1. Given the coordinate of one point and the slope of the straight line that passes through this point, the formula of this line can be calculated.

2. Two intersecting lines determine an only point

Assuming that motion model in consecutive frames satisfies affine transformation. The sketch map of the model is shown in Fig. 1. Here, line AB denotes main axis of the object, where A and B represents the intersection point of main axis and edge of the object respectively in current frame. C and C' denotes the tracking point in current frame and next frame respectively. The rotation angle between consecutive frames is θ. Suppose the coordinate of A and B is (x_1, y_1) and (x_2, y_2) respectively, which can be computed according the intersection point of main axis and edge of the object. Thereby, k_c, the slope of AB can be obtained. In the first frame, the tracking point C is labeled as (x_c, y_c). Then, k_b, the slope of AC is obtained too. In addition,