

Anti-Occlusion Tracking Method Based on Kernel Correlation Filter

Weichuang Jiang ^{a)}, Fuxing Zhang

Zhejiang Sci-Tech University, Hangzhou, China

^{a)} WeichuangJiang@163.com

Abstract. In the real scene, the target itself and the background will be unpredictable changes, such as the occlusion and other factors will bring greater challenges to the object tracking. In this paper, we proposed a method that handle the occlusion based on Kernel Correlation Filters (KCFFR, KCF with forward response). An anti-occlusion model is established based on the distribution feature of the forward classifier in response to the maximum value of the forward classifier. The target is searched by the block region helix search method after the target is occluded, and the response of the sliding box is calculated in the target search process to determine whether the target is found. The algorithm is tested on the OTB (Object Tracking Benchmark) sequence and compared with the four tracking algorithms. The results show the effectiveness and accuracy of the KCFFR algorithm.

Key words: kernelized correlation filters; target tracking; classifier response; occlusion detection; target search.

INTRODUCTION

Target tracking is a basic branch in the field of computer vision, which is widely used, such as intelligent camera, autopilot, video surveillance and so on. Grabner H[1][2] et al. proposed a method using online Boosting to update the target appearance model while minimizing error product tracking method, this method of online Boosting out of the visual field for target is more robust, but the classifier tracking results depends too much on the first frame of the training. Z Kalal[3] proposed tracking learning detection (TLD) algorithm, and the algorithm of parallel tracker detector processing result to the learning process, and then update the target scale information and processing block, this tracking method is effecter than the long-term tracking algorithm in [4] performance, but for the complex background of the heavy occlusion detection the process does not meet the requirements of real time tracking, the effect is not ideal.

Hare S[5] et al. considered the sample distribution in the search space to build joint structured output tracking algorithm (Struck), the method reduces the noise problem of online model update samples, but the similarity is sensitive to multi-objective problems, the situation with the wrong target is easy to appear in the block. Zhong; W[6] et al. proposed a criterion model and generation model based on sparse combination tracking method (SCM), target occlusion problem using histogram method of image sub block spatial information, but the tracking process has high computational complexity, it is difficult to achieve the real-time tracking for the actual requirements. Henriques J[7] et al. proposed a nuclear cycle structure (CSK) tracking algorithm, the algorithm of correlation filters is proposed in the literature [8] tracking mechanism based on introducing the kernel function and use the gray feature, the tracking speed is greatly improved, on the basis of CSK, Henriques, J[9] et al proposed a correlation filter (KCF- Kernelized Correlation Filters) tracking algorithm, the algorithm improves the kernel function and HOG[10] features, although the tracking robustness is improved but for occlusion and multiscale problems have not been solved.

KCF ALGORITHM

KCF is a discriminative tracking method. This method trains a target detector in tracking process, uses target detector to detect whether the prediction position of the next frame is target, and updates the training set through the new detection result, and then updates the target detector. When the target detector is trained, the target area is selected as the positive sample, and the negative sample is obtained from the circle of the target area. Sample training process is ridge regression problem. The goal of training is to get a classifier by solving a set of weights:

$$f(z) = w^T z \tag{1}$$

In order to improve the performance of classifier, using the properties of the circulant matrix to avoid computing the inverse of the weight vector of W solution to Fourier transform domain, kernel function is introduced to improve the nonlinear regression function for better performance, resulting in time-consuming calculation, KCF algorithm uses the properties of the cyclic matrix and kernel matrix to solve this problem. When converting to the DFT domain, the (1) transformation is as follows:

$$\hat{f}(z) = (\hat{k}^{xz})^* \square \hat{a} \tag{2}$$

x is the training samples, z is the candidate region and a is the dual parameter of w, k^{xz} is the kernel function of x and z, the first line of circulant matrix composed of vector can be calculated at the same time, all the test samples in response to cyclic displacement test samples based on the components of the z, in response to find the maximum value in the sequence the area, corresponding to the target area. The aim of the tracking model is to distinguish the background and target from the image data of each frame. The KCF algorithm transforms the solution of the training model into the DFT domain by using the Fourier transform and the cyclic matrix, thus greatly reducing the computational complexity of the algorithm.

TARGET OCCLUSION DETECTION AND TARGET SEARCH

When the target is blocked in the process of motion, the maximum response of the classifier is greatly reduced because of the interference of the target area. Figure 1 is showing that KCF tracker in Jogging set of test sequences on the distribution of the maximum response of classifier, target began to be blocked in 64th frame, the distribution of the maximum response classifier is in the range of 0.3-0.6 between 1-64th frames, targets in the 64-73th frames to be completely obscured by partial occlusion, tracking area gradually reduced until it disappeared the classifier, the maximum response of a sharp decline to below 0.2 in the 102th frame even if the target again appeared on the horizon, the classifier also is unable to distinguish the target lead to the loss of target. In the three conditions, the maximum response of classifier to three intervals, the maximum response of the classifier depends on the size of the target feature point tracking by region the degree of interference, the target is occluded part of the maximum response of the classifier more.

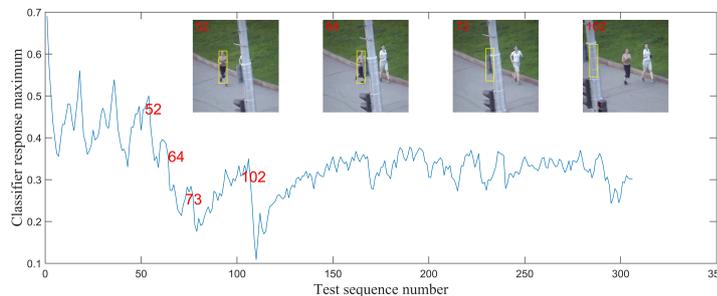


FIGURE 1. Jogging test sequence.

Based on the variation characteristics of the maximum response of the classifier, this paper predicts the occlusion of the target by the feedback information of the forward classifier response. The discriminant mechanism is as follows:

$$\max[F_{i+1}(z)] - rf_{(i+1-N,i)}(z) < 0 \quad (3)$$

i is the frame number of any frame, N is the frame interval length, and $F_{i+j}(z)$ is the classifier response sequence of the $i+1$ frame. $F_{(i+1-N,i)}(z)$ is the expectation of the maximum response of the forward classifier, r is the occlusion discriminant threshold, and the response maximum of the forward classifier is expected to be as follows:

$$f_{(i+1-N,i)}(z) = \sum_{j=i+1-N}^i \theta_j f_j(z) \quad (4)$$

$j = \{i+1-N, i+2-N, \dots, i\}$, $f_j(z)$ is the maximum of the response sequence of the j frame classifier, and θ_j is the weight coefficient, and its update strategy is:

$$\theta_j = \begin{cases} \frac{e^{-|f_j(z) - f_{(i+1-N,i)}(z)|}}{\sum_{j=i+2-N}^{i+1} e^{-|f_j(z) - f_{(i+1-N,i)}(z)|}} & f_j(z) \geq rf_{(i+1-N,i)}(z) \\ 0 & f_j(z) < rf_{(i+1-N,i)}(z) \end{cases} \quad (5)$$

And $j = \{i+2-N, i+3-N, \dots, i+1\}$, we can get an updated N rights value vector $\theta_{(i+2-N,i+1)}$, in the occlusion discrimination threshold range, the maximum distance to the maximum response of classifier expectations closer to the j frame classifier response, its weight is larger. When the occlusion discrimination threshold range, the target is occluded, tracking model stops updating, and gives zero to θ_j . The forward classifier responds to the update of the maximum expected value. It needs to provide an initial value. In this paper, the mean of the response maximum of the first frame to the N frames is taken as the initial value.

$$f_{init} = \frac{1}{N} \sum_{j=1}^N f_j(z) \quad (6)$$

When the occlusion detection algorithm detects the occlusion behavior, the tracking algorithm has lost the target. For the target that appears again in the camera's field of vision, we need to search the target first and resume the tracking. In this paper, a block region spiral search method is proposed to restore target tracking. We extend the structure of a block area based on the target area as a center. When the target is reappearing in the camera's field of view after being occluded, it is closer to the original target area than the block area. Therefore, compared with the top-down search and the left to right search, the target can quickly find the target. In the center of the target area to 4 sliding search, KCF algorithm with improved HOG feature, from the literature shows that the sliding step 8 pixels can achieve the best effect, the response of FS sliding block corresponding to the sub image and classifier calculation, then according to equation (7) is found again that target discrimination:

$$f_s - rf_{(i+1-N,i)}(z) \geq 0 \quad (7)$$

EXPERIMENTAL RESULTS

The KCFR tracking algorithm is completed in the MATLAB 2014b experimental environment. The experimental platform is Intel (R) Core (TM) i5 6300 (HQ) CPU, quad core 2.3GHz, memory (CPU) is x, the operation system is 10 64 bit operation system. In order to test the performance of KCFR tracking algorithm, the experimental data OTB (Object Tracking Benchmark) 51 sets of video sequences in [19], firstly, occlusion handling algorithm was verified, and the OTB code base on KCFR algorithm and 4 algorithms (Staple, DSST, KCF, SST)

were analyzed and compared. In this paper, the parameter tracking algorithm are set as follows: R block decision threshold, the range is $0 < r \leq 1$. During the experiment the small may result in occlusion error, $r=0.8$ experiment. N is the frame interval length, and its range is $N > 1$. In order to improve the generalization ability of the forward classifier's response to the maximum expected value, we need to increase the value of N. If the N value is too large, it may lead to occlusion detection and $N=30$.

In this paper, one pass evaluation method (OPE) is used in the OTB library for the five algorithms. The representative accuracy and the representative success rate of the five tracking algorithms have been tagged in the graph. In tracking accuracy and tracking success rate, the algorithm is better than the other four algorithms. In the tracking accuracy, this algorithm compared to Staple, DSST, KCF, SST were increased by 6.1%, 7.1%, 10.4%, 23%, the average increase of 11.7%; the success rate of tracking, this algorithm compared to Staple, DSST, KCF, SST were increased by 1.5%, 10.9%, 18.7%, 25.8%, increased the average 14.2%. It is known from literature [13] that the results of the tracking success rate are more accurate than the results of tracking accuracy evaluation.

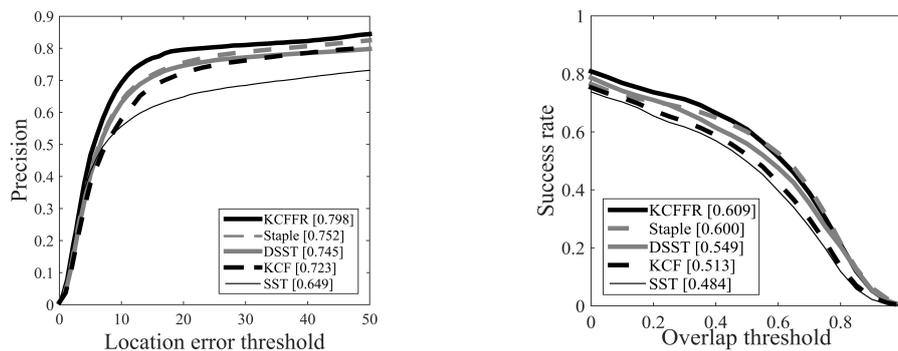


FIGURE 2. Comparison results of five algorithms.

REFERENCES

1. Grabner H, Grabner M, Bischof H. Real-Time Tracking via On-line Boosting[C]// British Machine Vision Conference 2006, Edinburgh, Uk, September. DBLP, 2006:47-56.
2. Grabner H, Leistner C, Bischof H. Semi-supervised On-Line Boosting for Robust Tracking[C]// European Conference on Computer Vision. Springer-Verlag, 2008:234-247.
3. Kalal Z, Mikolajczyk K, Matas J. Tracking-Learning-Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 34(7):1409-1422.
4. Pernici F. FaceHugger: The ALIEN Tracker Applied to Faces[J]. Lecture Notes in Computer Science, 2012, 7585:597-601.
5. Hare S, Saffari A, Torr P H S. Struck: Structured output tracking with kernels[C]// International Conference on Computer Vision. IEEE, 2011:263-270.
6. Zhong W. Robust object tracking via sparsity-based collaborative model[C]// Computer Vision and Pattern Recognition. IEEE, 2012:1838-1845.
7. Henriques J, o F, Caseiro R, et al. Exploiting the Circulant Structure of Tracking-by-Detection with Kernels[M]// Computer Vision – ECCV 2012. Springer Berlin Heidelberg, 2012:702-715.
8. Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters[C]// Computer Vision and Pattern Recognition. IEEE, 2010:2544-2550.
9. Henriques J F, Rui C, Martins P, et al. High-Speed Tracking with Kernelized Correlation Filters[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(3):583-596.
10. Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2005:886-893.