Li Wei
Jincheng College
Nanjing University of Aeronautics and Astronautics
Nanjing, China
e-mail: carry_wei@nuaa.edu.cn

Meng Ding and Xu Zhang
School of Civil Aviation
Nanjing University of Aeronautics and Astronautics
Nanjing, China
e-mail: nuaa_dm@nuaa.edu.cn

—

I. INTRODUCTION

Visual tracking is a remarkable research issue in computer vision and image understanding, with wide-ranging applications including intelligent surveillance, robot and unmanned aerial vehicle. Generally, existing tracking algorithms can be divided into two types as generative algorithm and discriminative algorithm. The former describes target by learning an appearance model and searches for the most similar area in a new frame as the result. The latter extracts the distinguishing feature of target and then uses a discriminative classifier to detect the target from backgrounds. Recently, as one of classic discriminative algorithms, the trackers based on correlation filter (CF) have achieved outstanding performance in different benchmarks. Furthermore, features from pre-training convolution neural network (CNN) are substituted for the hand-crafted features to further improve the performance of correlation filter-based methods [1,2].

This paper proposes a visual tracking method which uses reliability evaluation and feature selection in the framework of CF [3,4]. The main contributions of this paper can be summarized as follows: Firstly, this paper introduces an effective mechanism to evaluate the current tracking nu 58.ap

$$B_t = (1-\eta)B_{t-1} + \eta\sum_{k=1}^{d}\overline{F_t^k}F_t^k \quad (4)$$

Where η is a learning rate. For the following frame t, the feature map z_t is extracted from the image patch with the center of the prediction location in the previous frame. In the frequency domain, the DFT of y_t is calculated by

$$Y_t = \frac{\sum_{l=1}^{d}\overline{A_{t-1}^l Z_t^l}}{B_{t-1}+\lambda} \quad (5)$$

The location of target in current frame can be estimated based on the maximum of the response y_t which is the inverse DFT to the result of Eq. (5) $y_t = \mathscr{F}^{-1}\{Y_t\}$.

B. Tracking State Evaluation

Most trackers update the model directly without considering the reliability of current tracking result. In fact, if the estimated position is inaccurate, updating the model continuously is likely to lead to tracking failure. To solve this problem, the proposed method evaluates the confidence of the current detection result based on the feedback of the response. Based on the evaluation results, the tracker determines whether updating the model and whether using the deep convolutional feature. To some extent, the maximum response and the shape of response map reflect the confidence of the tracking results. When the features extracted from the test sample match the learned model, the correlation response map should have a sharp and single peak, remaining smooth in other regions. The steeper the single peak is, the more reliable the tracking result will be. On the contrary, if there are obviously fluctuations in the response map, it means that the confidence of the tracking results is very low. If we continue using the currently

E. Algorithm implementation

According to the above discussion, the proposed tracking method using reliability evaluation and feature selection is summarized as follows:

The proposed tracking algorithm

```
initial target bounding box x_0
estimated object state x_t = (x̂_t, ŷ_t, ŝ_t) , handcraft feature model H_h and deep feature model H_d

Crop out the searching window in frame t according to (x̂_{t-1}, ŷ_{t-1}) and extract the handcraft features;
// Translation estimation
Compute the correlation response map F_response using H_h and Eq. (5) to estimate the new position (x_t, y_t) , using Eq. (6) and Eq. (7) compute F_max and ARR ;
//Re-detection
  the tracking result is not reliable
    Crop out the searching window in frame t according to (x̂_{t-1}, ŷ_{t-1}) and extract the deep features;
    Compute the correlation response F_response using H_d and Eq. (5) to estimate the location (x_{t_cnn}, y_{t_cnn}) , using Eq. (7) compute ARR;
      the tracking result is reliable (x_t, y_t) = (x_{t_cnn}, y_{t_cnn}) ;


// Scale estimation
Estimate the optimal scale ŝ_t ;
// Model update
  the first frame
  Using Eq. (4) learning model H_h both and H_d ;
        the tracking result is not reliable
  Update the hand-crafted model H_h ;

      End of video sequences
```

III. Experimental Results

In our experiment, PCA-HOG [6] is used as the handcrafted feature for target representation. The feature extracted by VGG-16 network trained on ImageNet is used as the deep feature, which can be obtained from the Matconvnet toolkit. The proposed algorithm is implemented in MATLAB 2016a and all the evaluation algorithms run on a 3.40GHz PC with 16GB RAM. The specific parameters in this paper are set as follows: The regularization parameter of Eq. (1) equals to $\lambda=10^{-2}$, and the learning rate is set to $\eta=0.025$. The size of the search window is set to 2 times of the target size for handcraft model as 2.5 times for deep model generally. The standard deviation of the desired Gaussian function output is set to 1/16 of the target size for handcraft model as 1/5 of the target size for deep model. The number of frames to update the deep feature model is set to $k=3$. The related parameters in the tracking result discriminant mechanism are set to $\theta=0.4$, $\delta=3$, $\mu=0.2$. In the experiments, the parameters of the tracker are fixed.

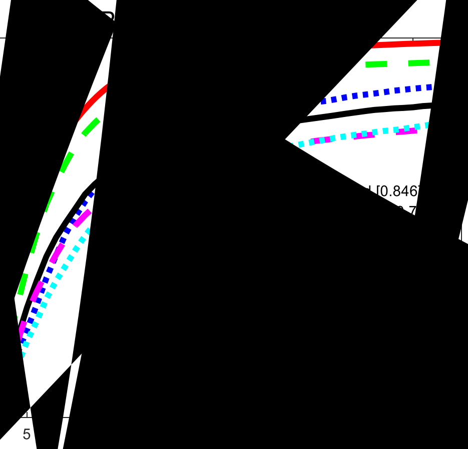
We assess the proposed method on a large benchmark dataset OTB-2013 [7] that contains 50 test sequences. The test method is used to track the target by frame by frame after initializing the initial frame in the sequence. The contrast algorithms used in the experiment are set according to the open source code in the database.

Three criteria are used for quantitative performance evaluation [7]: (1) Center location error (CLE) indicates the average Euclidean distance between the ground-truth and the estimated center location. (2) Precision rate (PR) demonstrates the percentage of frames whose estimated location is within the given threshold distance (20 pixel generally) of the ground truth. (3) Success rate (SR), which is defined as the percentage of frames where the bounding box overlap surpasses a threshold (50% generally).

We evaluate the proposed algorithm using deep convolutional feature and discriminant mechanism (DFDM) on the benchmark with comparisons to five state-of-the-art trackers. These five trackers come from three typical categories of tracking algorithms: (1) tracking with correlation filter (DSST[5], KCF[3]), (2) tracking with multiple online classifiers (TLD[8], SCM[9]) (iii) tracking with convolution neural network (SiamFC[10]). We show the results in one-pass evaluation (OPE) using the precision rate and success rate as shown in Figure. 1, where the legend contains the AUC score for each tracker. The results shown in Figure.1 illustrate that the proposed approach performs well against the existing methods in OPE. Moreover, we present the quantitative comparisons of precision rate where the given threshold distance is equal to 20 pixels, AUC score and tracking speed in Table 1. The speeds are from the original paper as the default processor is CPU. The first and second maximum values are highlighted by bold and underline. The proposed method performs favorably against existing methods in precision rate (PR) and AUC score with a better real-time performance. Figure.2 illustrates the tracking results in several test sequences and shows the proposed algorithm performs favorably against the other five algorithms.

IV. Copyright Forms and Reprint Orders

In this paper, we propose an effective algorithm based on correlation filter. The proposed method can evaluate the current tracking state effectively by analyze the response map. According to the current tracking state, the tracker determines whether the samples are reliable and whether to update model. We further use the deep convolutional feature to track targets in case of unreliable tracking result. Both the quantitative and qualitative experimental results show that the proposed algorithm is superior to the existing methods in terms of accuracy, efficiency and robustness.


[0.846
5

1. Single Target Tracking Using Reliability Evaluation and Feature Selection

Accession number: 20203409061886
Authors: Wei, Li (1); DIng, Meng (2); Zhang, Xu (2)
Author affiliation: (1) Nanjing University of Aeronautics and Astronautics, Jincheng College, Nanjing, China; (2) Nanjing University of Aeronautics and Astronautics, School of Civil Aviation, Nanjing, China
Source title: Proceedings - 2019 12th International Symposium on Computational Intelligence and Design, ISCID 2019
Abbreviated source title: Proc. - Int. Symp. Comput. Intell. Des., ISCID
Volume: 1
Part number: 1 of 2
Issue title: Proceedings - 2019 12th International Symposium on Computational Intelligence and Design, ISCID 2019
Issue date: December 2019
Publication year: 2019
Pages: 228-231
Article number: 9098305
Language: English
ISBN-13: 9781728146522
Document type: Conference article (CA)
Conference name: 12th International Symposium on Computational Intelligence and Design, ISCID 2019
Conference date: December 14, 2019 - December 15, 2019
Conference location: Hangzhou, China
Conference code: 159910
Publisher: Institute of Electrical and Electronics Engineers Inc.
Abstract: This paper proposes a visual tracking algorithm using reliability evaluation and feature selection mechanism in the framework of Correlation Filter (CF). Three additional modules are used to improve the current CF-based tracker. Firstly, a module of reliability evaluation is used to determine whether the current tracking result is reliable. Secondly, an updating module is used to determine whether to update the target model by comparing the reliability of current tracking result with historical average. Thirdly, a feature selection module is presented to select hand-crafted feature or deep convolutional feature according to the current tracking state. Experimental results on a benchmark dataset of fifty challenging test sequences show that the proposed method can reduce the interference of complex factors effectively. © 2019 IEEE.
Number of references: 10
Main heading: Target tracking
Controlled terms: Electric current control - Feature extraction - Intelligent computing - Reliability - Statistical tests
Uncontrolled terms: Benchmark datasets - Complex factors - Correlation filters - Current tracking - Reliability Evaluation - Selection mechanism - Single target tracking - Visual tracking algorithm
Classification code: 723.4 Artificial Intelligence - 731.3 Specific Variables Control - 922.2 Mathematical Statistics
DOI: 10.1109/ISCID.2019.00059
Funding Details: Number: 20155152041, Acronym: -, Sponsor: Aeronautical Science Foundation of China; Number: 18KJB590002, Acronym: -, Sponsor: Natural Science Research of Jiangsu Higher Education Institutions of China; Number: 61203170, Acronym: NSFC, Sponsor: National Natural Science Foundation of China;
Funding text: ACKNOWLEDGMENT This research has been supported by the National Natural Science Foundation of China (No.61673211, No.U1633105, No.61203170), Natural Science Foundation of the Jiangsu Higher Education Institutions of China (No.18KJB590002), and Aeronautical Science Foundation of China (No.20155152041).
Compendex references: YES
Database: Compendex

Data Provider: Engineering Village