

Improved TLD Target Algorithm Based On Feature Fusion

Wenjun Dai, Tianqing Chang, Kuifeng Su, Quandong Wang

Department of Control Engineering, Academy of Armored Force Engineering, Beijing, 100072, China

805126841@qq.com

Keyword: TLD, Feature fusion, Rotating target.

Abstract. A modified TLD tracking algorithm based on fusion of color model and ORB feature is proposed for the poor tracking performance of TLD algorithm for rotating target. The algorithm uses a feature fusion tracking algorithm instead of the optical flow method in the TLD tracking module, updating target tracking window and target template in the tracking module to improve the target tracking capability of rotating target, and compare the Bhattacharyya coefficient of the target model and the candidate model of with threshold to determine whether the tracking is failure or success, the detection module will reinitialize tracking if the tracking module failure. The experimental results show that, compared with the original TLD tracking algorithm, the improved algorithm could improve the ability of the rotating target tracking, and can meet the real-time requirements.

1. Introduction

Visual tracking as an important direction of computer vision research and has extensive attention of the researchers, and has been widely used in many fields, such as civil and military [1, 2]. But for the complex background and the target tracking, especially takes a long time to track the target, inevitable will target shape change, such as light illumination condition changes and the object occlusion, which brings some challenges and difficulties to the target tracking. For the traditional long-time tracking algorithm, the tracking module and the detection module does not cooperate closely, when the tracking algorithm detected target tracking and entered the tracking module, and the detection module will not involve in the process of tracking. This method has a major flaw: when the object exists shape change or occlusion, it is easy to fail, and when the target reappears in the field of view, the tracking algorithm unable to detect the target and start tracking. Due to the change of the target shape and other factors can affect the tracking effect, the online learning mechanism is introduced into the tracking algorithm and considering tracking and detection algorithm comprehensively. Paper [3] introduced the TLD tracking algorithm is a neoteric long-time tracking algorithm, it combined target tracking algorithm with detection algorithm and integrated into the online learning mechanism to solves the problem that when the target disappeared again appear, the algorithm cannot continue to track, which make the overall target tracking more stable. But because the TLD tracking algorithm based on median flow tracker algorithm, it cannot achieve effective tracking for rotating targets. Aiming at this problem, this paper proposes a tracking algorithm based on fusion of color models and ORB feature to instead the median algorithm, and solve the this problem that original TLD tracking algorithm cannot achieve effective tracking on rotating target.

2. Tracking algorithm based on Feature Fusion

2.1 TLD tracking algorithm

TLD tracking algorithm is proposed by Zdenek kalal et al in 2010, and the algorithm presents a new tracing framework. It divide the main task into three modules: tracking module is based on pyramid LK optical flow and detection module is builded by cascade classifier and a P-N online

learning module. The Structure of TLD is shown in Figure 1.

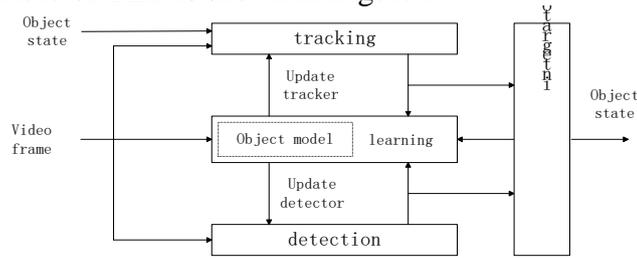


Fig.1 Structure of TLD

The tracking module is based on median flow method, which is based on the Pyramid LK optical flow method, using the forward and backward tracking error for target tracking. For a point in the image, assuming forward tracking trajectory and backward tracking trajectory is consistent, on the basis of this hypothesis, marking the target with a rectangular box and calculate the forward and backward error of the point of uniform distribution in a rectangular frames and the correlation of those points, and elect the point which less than the median of the forward error and more than the value of the correlation value to determine new location of the target. Tracking failure strategy is as follows: if the prediction of the front and back to the median error is greater than the threshold or beyond the image boundary, then determine the target tracking failure and do not return to the target rectangle. This method can be used to automatically find the tracking failures and errors in the case of object occlusion.

The detection module scans the current frame through the sliding window, and uses the constructed cascade classifier to detect the sliding window which contains the target. By using the target rectangle frame of the first frame, the sliding window is built according to the following rules: Sliding window size scaling ratio of 1.2, a total of 21 times, and the smallest bounding box size is 20 pixels, the horizontal sliding step coefficient and the vertical sliding step coefficient for the target frame width and height of 10%. For an image with size of 320*240, about 50000 image blocks can be obtained after the sliding window scan images. All the images should through the cascade classifier which consist of variance classifier, ensemble classifier and nearest neighbor classifier, as shown in figure2.

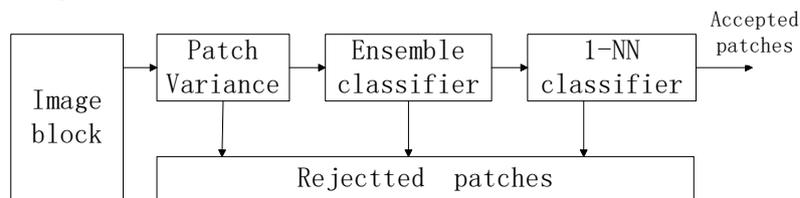


Fig. 2 detection module

The learning module in TLD is a semi supervised learning method which based on P-N learning algorithm and its main function is to initialize the target detector in the first frame of image sequence, and using P experts and N experts to update detection module in subsequent image sequence. The main steps are as follows: firstly, initialize the classifier through the image blocks are marked as positive samples or negative samples and constraint conditions; Then, using classifier to label the unmarked image block and record image blocks which are not consistent with the constraints; Lastly, modify the sample labels which is not consistent with the constraints and added to the training set to update the parameter of classifier.

In the original TLD algorithm, the tracking module use the optical flow method which has the character of large computation and poor tracking effect for the rotating target. So it can choose other tracking algorithm to instead of LK algorithm. Considering the Mean Shift algorithm which has less calculation and it can effectively solve the problem that the large computation of the optical flow method in the TLD tracking algorithm , but the Mean Shift algorithm only consider the color information, as a result, the tracking effect is poor when background is compelled. It can improve the tracking ability in complex background through the fusion of color and ORB feature.

2.2 Feature Fusion

The Hue component of HSV color space is not sensitive to the light intensity, so it is suitable to use for the target tracking algorithm in the complex environment, and the Mean-Shift algorithm uses the Hue component as the target model. The distribution range of the Hue component is 0-360, in order to simplify the calculation, the Hue component was divided to 0-255 quantization. Assume $\{x_i^*\}_{i=1,2,\dots,n}$ represents the location of the target area which center location is y_0 , and the color histogram of the target is $q = \{q_u\}_{u=1,2,\dots,m}$

$$q_u = C \sum_{i=1}^n K \left(\left\| \frac{y_0 - x_i^*}{h} \right\| \right) \delta[b_f(x_i^*) - u] \quad (1)$$

Among them, $b_f: R^2 \rightarrow \{1, 2, \dots, m\}$ is the mapping function of the color features to the histogram, δ is the Kronecker function, $K(\cdot)$ represents kernel section function, and h is the radius of the nuclear section, C is the normalized coefficient.

Mean-shift visual tracking algorithm[4] firstly select the tracking target and set its position as the initial target position, making the greatest similarity of the tracking window as the target position by compare the similarity between the tracking window calculation region of color histogram and the color histogram template. Target tracking steps are as follows:

- 1) Select the target template;
- 2) Determine the search window;
- 3) Put the target area of the current frame as the next frame input until the end of the track is tracked.

Orb feature [5] is a kind of better local invariant feature which based on the improvement and optimization of brief detection and fast feature descriptor. ORB feature gets key points through the FAST feature algorithm [6]. Because the FAST feature points only compare the gray value, the advantage is that the calculation speed is fast, but it does not have the rotation invariance and scale invariance. In order to make the ORB feature obtain rotation invariant, a direction is needed for each FAST feature point. ORB uses the gray centroid method to select the direction of the FAST feature points. The gray centroid method assumes that there is an offset between the gray and the centroid of the corner, then the offset vector can be expressed in a direction. Taking the feature point as the center and the origin of the coordinate, the position of the center of mass is calculated in the field of S . Taking characteristic points as the starting point and the center of mass as the terminal to structure direction vector and the direction of this vector is the direction of the feature points. Defining the moment of the area S :

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y) \quad (2)$$

The angle between the feature points and the center of mass is defined as the direction of the FAST feature points:

$$\theta = \alpha \tan 2(m_{01}, m_{10}) \quad (3)$$

The ORB algorithm using binary brief descriptor described the detected feature points, and the brief descriptor with rotation invariance and sensitive to the noise problem has been improved. For a given image, from the image to select a size of $S \times S$ of the local area P and take two values test in the area. Two values of test τ criteria are as follows:

$$\tau(p; x, y) = \begin{cases} 1, & p(x) < p(y) \\ 0, & \text{others} \end{cases} \quad (4)$$

Among them, X and Y is the location of the two pixel points, that is, two-dimensional coordinates. $P(X)$ and $P(Y)$ are the gray values of X and Y at the pixel points. Image block P CPC selected pixels of two value tested (N generally select 256), so the N dimensional binary bit string of brief feature descriptors of image block P is:

$$f_{nd}(p) = \sum_{1 \leq i \leq n_d} 2^{i-1} \tau(p; x_i, y_i) \quad (5)$$

At the same time, the ORB algorithm is used to select the sub window of 5×5 in the 31×31 pixel region of the feature point neighborhood. It can eliminate the noise by comparing the gray integral of the sub window instead of the pixel value.

Select the minimum Hamming distance of two pairs of matching points $(d_i, d_i^*), (d_j, d_j^*)$ after

obtain the optimal matching points and the Hamming distance $D_i \leq D_j$, Euclidean distance is V_i, V_j . If meet $\frac{V_i}{V_j} \leq 0.8$, the target in the current frame center position is equal to the target center position of previous frame plus the minimum Hamming distance matching points offset. If do not, the center position of the target in the current frame is equal to the target center position in previous frame plus the average offset of two pairs of matching points with the minimum Hamming distance.

In the algorithm, the s coefficients of the two color histogram are usually used to measure the similarity. The process of target tracking is using MS vector gradient in the current frame target area to search probability distribution of maximum similarity with target template. Selected the tracking target in the initial frame and extract its ORB features and color model to locate the target area through the ORB feature matching and Mean-Shift algorithm; Determine the target's ORB feature and color histogram and output the tracking target location through the fusion and correction of target position of two target area. The fusion rules of the target candidate region are as follows:

1) Calculate the coincidence degree of the candidate regions of the two target and the formula of coincidence degree is as follows:

$$\text{ratio} = \frac{\text{area } 1 \cap \text{area } 2}{\text{area } 1 \cup \text{area } 2} \quad (6)$$

Among them, *area 1* is target area obtained by the mean – shift algorithm and *area 2* is the target area obtained by ORB algorithm. If the coincidence degree is greater than 0.9, the center of the coincident area is used as the center of the target and modify the ORB feature template and color histogram template. If this condition is not satisfied, continue the next step.

2) Calculate the Euclidean distance between the center of the target candidate region and the regional center of the upper frame image. If the two European distance is less than the threshold value A , then enter the next step; if one of the European distance is greater than the threshold, then the distance to the smaller of the output as the final output; If the distance between the two is greater than the threshold value, then continue the next step.

3) Calculate the similarity of the normalized model histogram and color histogram of the tracking target, the similarity is expressed by Bhattacharyya coefficient. If the two Bhattacharyya coefficients are less than the threshold B , then enter the next step, the B value in this paper 0.85; if the two has one less than the threshold value, then the output of the larger coefficient is the final output; if the Bhattacharyya coefficients is less than the threshold value, then continue the next step.

4) If the above conditions do not meet, the tracking is failure. When the tracking failure, the detector will reinitialize the tracking.

Compared with the original TLD algorithm, this paper only replace optical flow method by tracking algorithm based on feature fusion, and the detection module and the learning module have not changed.

3. Experimental Simulation

In order to test the effectiveness of the algorithm, this paper selects a video for the target tracking experiment in complex environment, and compared with the original TLD algorithm. The target in the video should be able to turn rapidly and the hardware environment of the experiment is a computer with Windows 7 operating system and 4GB memory and Intel(R) Core(TM) i3-2100@3.10GHz and 3.09GHzCPU. The Software environment is studio Visual 2010 and Opencv2.43 database. Test video resolution is 1080*720.



(a)



(b)



(c)

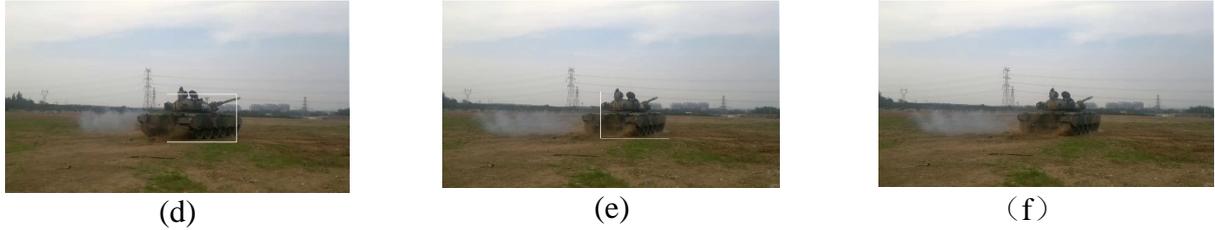


Fig 3 experiment result comparison

The tank target in the video do fast rotational motion with a lot of dust, it certainly bring much influence to the target tracking. In the first picture from left to right: fourth frame, Fifteenth frame and twenty-fourth frame. (a) ~ (c) express the tracking effect using this method and (d) ~ (f) for the original TLD algorithm tracking effect. The original TLD algorithm tracking failure in the twenty-fourth frame, while this algorithm can track the target stably, and the previous frame tracking effect is better than the TLD algorithm

The comparison of the tracking error and tracking rate is in Table 1. From table 1, when the tracking fast rotating targets, the original TLD algorithm slightly better than the improved algorithm TLD algorithm in this paper in the tracking rate, but have large tracking error. Comprehensive comparison, the improved TLD algorithm in this paper have better tracking effect than the original TLD algorithm in tracking fast rotating targets.

Table 1 tracking error and average FPS comparison		
Algorithm Type	TLD Algorithm	Present Algorithm
Average FPS / Hz	21.2	18.7
Error Rate / %	24	3
Algorithm Type	TLD Algorithm	Present Algorithm
Average FPS / Hz	21.2	18.7
Error Rate / %	24	3

4. Summary

TLD tracking algorithm is a novel online learning tracking algorithm, based on the original TLD tracking algorithm to track the problem of poor ability of rapid rotation of the target, an optimized TLD tracking algorithm based on feature fusion is proposed. The algorithm use fusion tracking algorithm based on ORB features and color model to replace the original optical flow in the TLD and optimizing the original TLD algorithm, computing Bhattacharyya coefficient between the target model and candidate model and compared with a threshold to the tracking failure self-detection. Determine the validity of tracking by computing the similarity between the target frame which tracking algorithm returned and the target frame the previous frame output, the detector will reinitialize tracking when the tracking failure occurs. Finally, a simulation experiment that tracking the fast rotation of the tank target show that compared with the original TLD algorithm, this algorithm in this paper can realize the fast rotating target tracking with robustness and accuracy.

References

- [1]. Chen Zhimin, Bo Yuming, Wu Pan-long, et al. Novel particle filter algorithm based on adaptive particle swarm optimization and its application to radar target tracking[J]. Control and Decision, 2013, 28 (2):193-200.
- [2]. Rublee E. Rabaud V. Konolige K, et al. An Efficient Alternative to SIFT or SURF//Proc of the IEEE International Conference on Computer Vision. Barcelona.Spain.2011:2564-2571.

- [3]. Kalal Z, Mikolajczyk K, Maters J. Tracking-learning-detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7):1409 -1422.
- [4]. Beyan C. Temizel A. Adaptive Mean Shift for Automated Multi Object Tracking. IET Computer Vision, 2012, 6(1):1-12.
- [5]. Rosten E. Porter R. Drummond T. Faster and Better: A Machine Learning Approach to Corner Detection. IEEE Trans on Pattern Analysis and Machine Intelligence. 2010.32(1):105-119.
- [6]. Alahi A, Ortiz R, Vandergheynst P. FREAK: Fast retina Keypoint [C]. Computer Vision and Pattern Recognition, 2011: 510-517.