

# A Novel Target Algorithm based on TLD Combining with SLBP

Jitao Zhang<sup>a</sup>, Aili Wang<sup>a,\*</sup>, Mingxiao Wang<sup>a</sup>, Yuji Iwahori<sup>b</sup>

<sup>a</sup>Higher Education Key Lab for Measure & Control Technology and Instrumentations of Heilongjiang,  
Harbin University of Science and Technology, Harbin, 150080, China

<sup>b</sup>Department of Computer Science, Chubu University, Aichi, Japan

---

## Abstract

TLD (Tracking-Learning-Detection) algorithm can be good for a long time to track the target in rotation, occlusion, illumination and other circumstances. However, in the case of uneven illumination, occlusion, tracking target fuzzy and so on, the problem of false tracking or tracking failure often occurs. Aiming at the shortcomings of TLD tracking algorithm, this paper will take TLD as the basic framework of target tracking and improve the detection module. When the tracking target has better texture feature, the SLBP (Semantic Local Binary Pattern) classifier is used to replace the nearest neighbor classifier in the detection module, which converts the image into SLBP texture feature vector to classify the samples using. In this paper, TLD-SLBP, MEEM, SCM, Struck and TLD are compared by using CVPR2013Benchmark test platform. The experiment results show that the TLD-SLBP algorithm has a higher success rate than other algorithms.

*Keywords:* target tracking; TLD; SLBP

(Submitted on February 3, 2017; Revised on April 22, 2017; Accepted on June 16, 2017)

© 2017 Totem Publisher, Inc. All rights reserved.

---

## 1. Introduction

Target tracking is widely used in video surveillance, intelligent transportation, human computer interaction, sports, military and other fields. After many years of research, it has made great progress. However, due to the complexity of the application of tracking tasks, the difficulty of target tracking is greatly increased [1]. Firstly, the pose, shape, texture and scale of the target are constantly changing. Secondly, the object will be affected by the changes of illumination, complex background and occlusion [2]. Good tracking algorithm needs to meet the fast speed, small amount of computation, but also to ensure that the algorithm can maintain good real-time, accuracy and robustness in the case of target change, occlusion, complex background and so on [3].

The following are commonly used tracking algorithm in the past 20 years:

(1) tracking algorithm based on template matching

Use the previous image frames to construct characterization of model target, then conduct the image feature matching to the current frame, such as normalized cross correlation NCC (Normalized Cross Correlation), KLT optical flow (Kanade-Lucas-Tomasi), Kalman filter algorithm and Mean Shift algorithm.

(2) template matching and tracking based on extended appearance model

---

\* Corresponding author. Tel.: +86-0451-86392366; fax: +86-0451-86392366.  
E-mail address: aili925@hrbust.edu.cn.

Use the target appearance or behaviour of the development model to search the best match of the development model in the image, such as increment visual tracking.

### (3) tracking based on discriminant classification

Based on the model of target distinguish between foreground and background, namely detection and tracking, the tracker constructs a classifier to distinguish the target pixels and background pixels, and use the new sample to update the classifier, such as the forward-backward tracking (FBT), the multiple instance learning tracking (MIT), tracking learning detection (TLD) algorithm and the Circulant Structure with Kernels for fast tracking [4].

At present, the target tracking algorithm based on detection of has become the mainstream of research because of its good tracking ability. In 2006, Grabner et al. proposed the supervised and semi supervised learning method which used the adaBoost classifier to classify the target, and the target tracking problem was transferred to a target for detection and classification [5]. In 2009, Babenko Boris used a multi instance learning method to solve the change of object tracking process [6]. In 2012, Kalal proposed TLD tracking algorithm that which combines the traditional tracking algorithm and traditional detection algorithm to solve deformation and partial occlusion problems in the process of target tracking [7]. At the same time, the target model, the related parameters of the feature points, and the detection modules of the online learning mechanism are updated to improve the tracking module, so that the tracking result is more stable, robust and reliable.

## 2.The basic framework of the TLD algorithm

Tracking-Learning-Detection (TLD) is a kind of long time tracking algorithm for a single object in video, which is composed of three parts: the tracker, the learning module and the detector. It can be seen from the composition that the most important innovation of this algorithm is to combine the use of the tracker and the detector, and machine learning is introduced to improve the accuracy of the tracking results [8]. The basic framework of the algorithm is shown in Figure 1.

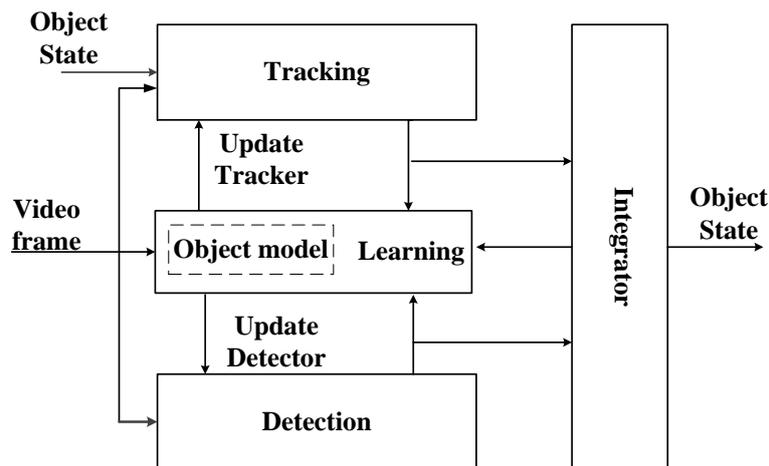


Fig. 1. The basic framework of TLD algorithm

Firstly, in the initial frame, the tracking target tracker is manually selected. And then detector and tracker are integrated tracking through comprehensive tracking output module of the two results, and use the learning algorithm to correct the tracker and detector.

## 3.The working principle of the traditional TLD algorithm

### 3.1. The tracker

The role of the tracker is to track the inter frame motion of the object and use the position of the target in the current frame to estimate its position in the next frame position. In the TLD algorithm, the median optical flow tracking algorithm is widely used by selecting the optical flow feature and 100 feature points. Optical flow tracking algorithm uses L-K (Lucas-Kanade) optical flow algorithm to track, while the use of the forward-backward error and the similarity of the characteristics of the frames to detect and determine [9,10].

The basic principle for determining the tracking point is: firstly, 100 feature points are uniformly produced surrounded the tracking box in the last  $t$ -th frame. Then the L-K tracker is used to forward track these points to the  $(t+1)$ -th frame, and then backward track to the  $t$ -th frame. Through the calculation error of FB, error minimum is select as the optimal half point tracking point. Finally, according to the change of coordinates and distance of these points, the position and size of the bounding box in the  $(t+1)$ -th frame is calculated (the median of the scale of the translation, the median of the scale of the scaling are chosen) [11,12].

The calculation principle of forward-backward error is shown in Figure 2.

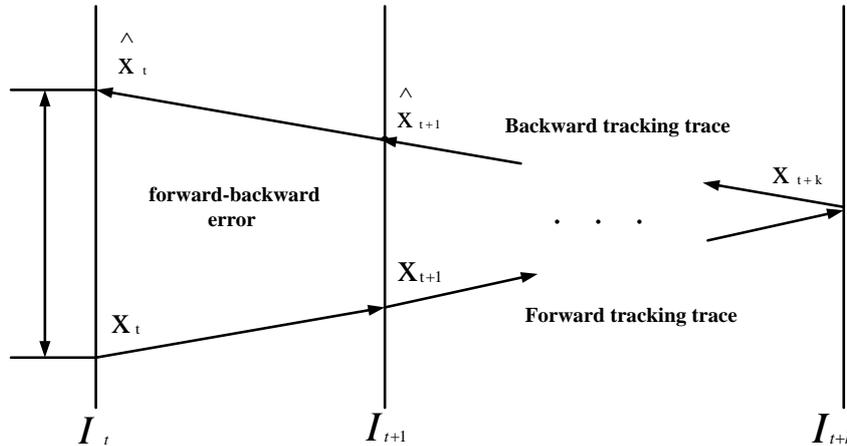


Fig. 2. The Schematic diagram of calculation principle of forward and backward error

In this method, the location of the feature point in the  $t$ -th frame is  $X_t$ , L-K optical flow tracking algorithm is used to carry on the tracking and calculates the position  $X_{t+k}$  of the feature in the  $(t+k)$ -th frame. And then the position is used as the initial feature point tracking coordinates, the reverse flow of light L-K tracking calculates the position  $\hat{X}_t$  in the  $t$ -th frame. The difference between  $X_t$  and  $\hat{X}_t$  (Euclidean distance) is the backward error  $I_t$ .

The TLD algorithm has good characteristic, including small amount of calculation, good real-time and stability in simple scenarios for tracking. However, when the target disappeared and appeared again, it is impossible to continue tracking, the detector is introduced to solve this problem [13,14].

### 3.2. The detector

The function of the detector is to use the detection algorithm, in the current frame of the search target, to identify the target with the target of the maximum similarity as the tracking results of the frame output. In the TLD algorithm, the detector can estimate the tracking error of the tracker and correct the tracking result. The detector consists of three cascaded classifiers: variance classifier, ensemble classifier and nearest neighbor classifier. The three classifiers operated independently. The principle of the detector is shown in Figure 3.

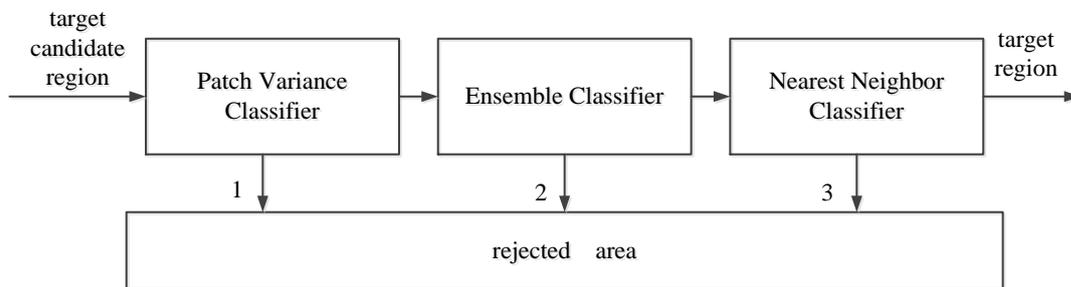


Fig. 3. The schematic diagram of the TLD detector

### 3.2.1. The variance classifier

The variance classifier is the first step of the cascade classifier. Through the comparison between the variance of the target area and the candidate region, the candidate region which are less than half of the variance of target area variance is excluded, so variance filter can eliminate a large number of non-target area. The variance filter needs to calculate the variance of the gray value of the image area, and the formula of the variance is shown in the formula (1).

$$D(x) = E(x^2) - E^2(x) \quad (1)$$

Where  $x$  is image region,  $D(x)$  represents gray variance of the image region,  $E(x)$  is mean value and  $E(x^2)$  is the mean value of image gray square.

### 3.2.2. The ensemble classifier

The ensemble classifier is the second step of the cascade classifier, which is further judged by the image region of the variance filter. The ensemble classifier is mainly composed of  $N$  random forest classifier. When the image area is input to the integrated classifier, the corresponding posterior probability is obtained, and the size of the posterior probability is obtained to determine whether or not to enter the nearest neighbor classifier. The principle of integrated classifier as shown in Figure 4, solid line in the figure represents a possible judgment.

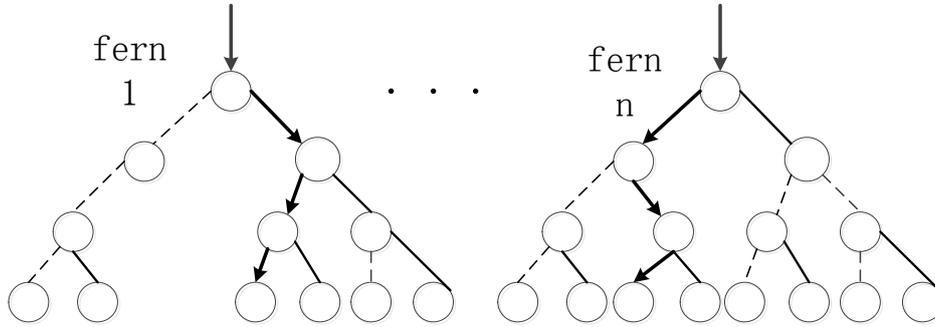


Fig. 4. The principle diagram of ensemble classifier

Each random fern is composed of a tree structure, and a posterior probability is obtained through the judgment of the image area. If the posterior probability of  $n$  random fern classifier is larger than the threshold value, the image region is determined to be positive samples; otherwise the samples are regarded as negative samples.

### 3.2.3. The nearest neighbor classifier

The image is needed to be input to the nearest neighbor classifier for further judgment after the variance filter and ensemble classifier. This stage mainly judges the similarity between the image region and the target model, as formula (2) and formula (3):

$$S^+(p, M) = \max_{p_i^+ \in M} S(p, p_i^+) \quad (2)$$

$$S^-(p, M) = \max_{p_i^- \in M} S(p, p_i^-) \quad (3)$$

Where  $p$  represents the image region,  $M$  represents the target area. The target model is composed of a number of positive and negative samples,  $p_i^+$  represents the positive samples in the target model, and  $p_i^-$  represents the negative ones.  $S^+(p, M)$  represents the maximum value of the similarity of the image region  $P$  and positive samples of the target model  $M$ .  $S^-(p, M)$  represents the maximum value of the image region  $P$  and negative samples of the target model  $M$ . Through these two similarities, we can finally find the final similarity  $S^r$  between the image region and the target model, which is expressed as a formula (4):

$$S^r = \frac{S^+}{S^+ + S^-} \quad (4)$$

Where  $S^+$  is the maximum value  $S^+(p, M)$  of the similarity of the positive samples,  $S^-$  is the maximum value  $S^-(p, M)$  of the similarity of the negative samples. If  $S^r$  is larger than the threshold value, the image area is judged to be a positive sample.

### 3.3. Learning module

According to the positive and negative samples generated by the detector in the TLD learning module, iteratively trained the classifier to evaluate the classification error of the detector, so as to avoid the same classification error and improve the accuracy of the detector. Among them, the application of the algorithm is P-N learning algorithm, which is composed of P experts and N experts in two parts. P experts will be mistakenly classified as negative samples positive samples to join the positive sample set, N experts will be mistakenly classified as positive samples to join the negative sample set. P-N learning can estimate the error of the classifier.

### 3.4. Integrated module

The main function of the integrated module of TLD is to integrate the results of the detector and the tracker, and output the final tracking result. Its main principle is:

a. If the tracker is successful, then determine the status of detection. If the detector detects the target, then determine all candidate samples detected category candidate samples to detect the degree of overlap as the standard cluster analysis of samples, the number of candidate categories are sample level, and calculated the mean samples of each category box, then select the candidate prediction samples and tracker detected samples of the overlap is greater than 0.7 of the candidate samples, accumulated the coordinates, finally the weighted operation sample and a detector to predict candidate samples to predict the tracker coordinates, the tracking results as the final results of the output. If the detector fails, only the tracking results are regarded as a final output, not to learn.

b. When the tracking is a failure and the test is successful, only the results are regarded as the final output, update the scope of the tracker, not to learn.

c. If the tracker and the detector cannot track successfully, then no target frame is output, continue to follow the next frame. In the next frame, the tracker does not work, the detector works until the detector successfully detects the target. The target position is used as the update position of the tracker, and the tracker begins to keep track of the target.

## 4. Tracking algorithm based on TLD combing with SLBP

According to the detector of the TLD algorithm, the three classifiers work independently. After the first two classifiers, the samples to be classified will have a certain deviation; the final test results of third classifiers will directly influence the detector, so improve the third classifier. When the variance of image tracking target is less than a threshold value, texture feature of the image is obvious. The tracking algorithm based on TLD combining with the SLBP classifier (TLD-SLBP) algorithm can effectively improve the tracking effect of light illumination.

The following part gives a detailed introduction of the detection module including the discriminant module and SLBP classifier.

### 4.1 Distinguish module

When the target has good texture features, the classifier based on the SLBP operator is used as a part of the detection module. Therefore, the design of the discriminant module to determine the detection module of the third classifiers is the use of the nearest neighbor classifier or SLBP classifier. When the LBP image variance of the tracking target is less than a certain threshold, the texture features of the image are obvious. At this time, the SLBP classifier is used to combine the detection module of the TLD algorithm, and it can get a good tracking effect. The workflow of the judgment module is shown in Figure 5.

The calculation method of LBP variance is as the following formula (5) and (6):

$$S(LBP)^2 = \frac{(LX_1 - M)^2 + (LX_2 - M)^2 + (LX_3 - M)^2 + \dots + (LX_n - M)^2}{n} \quad (5)$$

$$M(LBP) = \frac{LX_1 + LX_2 + LX_3 + \dots + LX_n}{n} \quad (6)$$

Where  $LX_i$  represents the uniform LBP value of the  $i$ -th pixel,  $n$  is the total number of LBP operators belongs to the tracked target,  $M$  is the mean of LBP image,  $S(LBP)^2$  is the variance.

4.2. The working principle of SLBP classifier

4.2.1. Calculation method of the SLBP feature

First, the color space is processed the two value operation, which the pixel greater than the central local pixel is record as 1; otherwise it is recorded as 0. Then calculate the arc length and the angle of the spindle, and discard the non-uniform model. Finally, we carry out the transformation between the matrix and the vector, each column of the two-dimensional matrix can be linked together to get one-dimensional vector. Figure 6 describes the semantic meaning of SLBP.

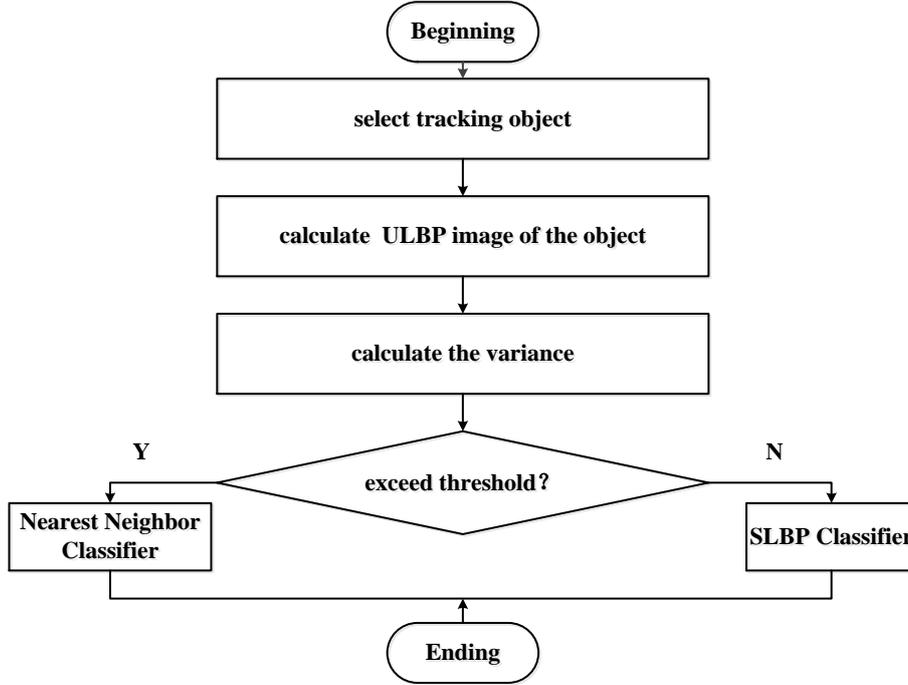


Fig. 5. The flowchart of judging module

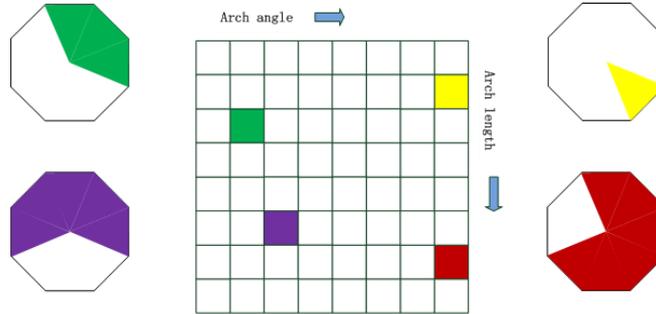


Fig. 6. The semantic meaning of SLBP

4.2.2. The working principle of SLBP classifier

The each input image to be detected are compared with the positive and negative samples of the sample set, and the correlation coefficient is calculated as the formula (7) and (8):

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \bar{H}_1)(H_2(I) - \bar{H}_2)}{\sqrt{\sum_I (H_1(I) - \bar{H}_1)^2 \sum_I (H_2(I) - \bar{H}_2)^2}} \quad (7)$$

$$\bar{H}_K = \frac{1}{N} \sum_J H_K(J) \quad (8)$$

Where  $H_1$  and  $H_2$  respectively represents the normalized histogram of the image to be detected and the positive and negative samples,  $N$  is the number of bin in the histogram,  $d$  is the correlation coefficient, the similarity range is  $[1,0]$ .

When the image to be detected is compared to the positive sample and the correlation coefficient is larger than a certain threshold value (set to 0.95), the target can be detected in the image. When the image to be detected is compared to the negative sample and the correlation coefficient is larger than a certain value, it is determined that there is no object in the image.

### 5.Experimental results and analysis

#### 5.1 Quantitative analysis

In this paper, the experiments are carried out in the window7, 64-bit system, the use of matlabR2012a platform, combined with C++ language mixed programming. The experimental results are shown in Figure 7, using the OPE (one-pass evaluation), SRE (spatial robustness evaluation) and TRE (temporal robustness evaluation) three kinds of evaluation methods in CVPR2013Benchmark to evaluate the tracking algorithm.

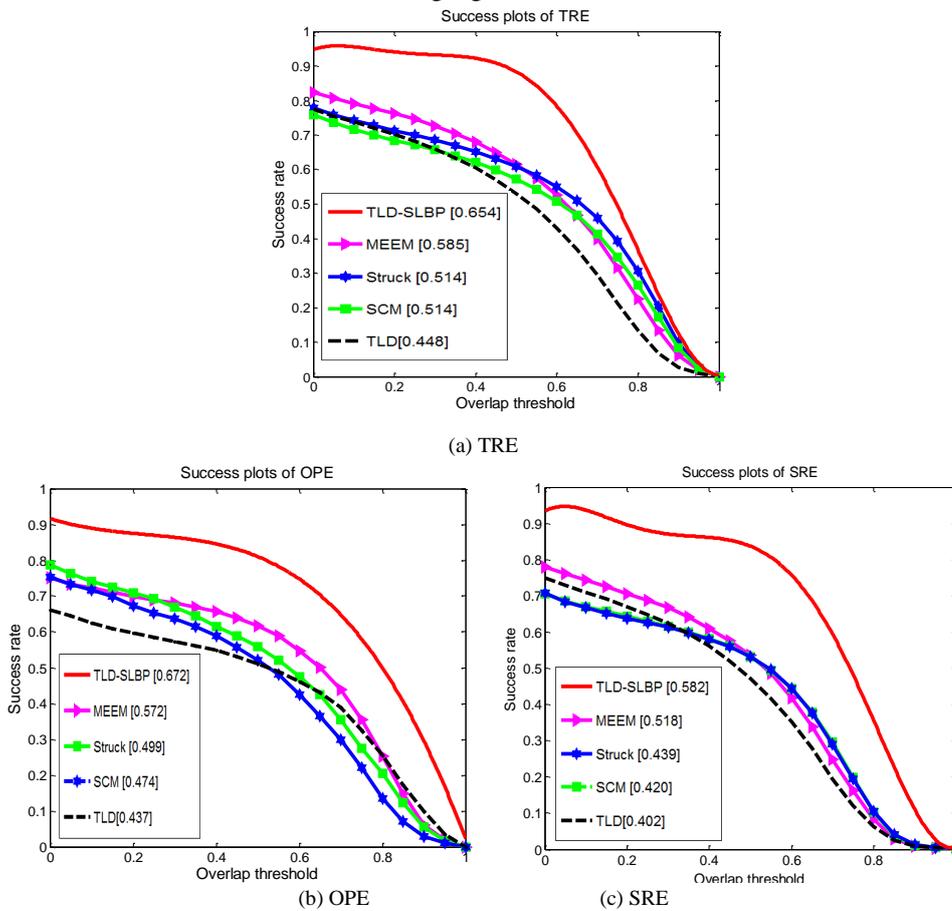


Fig. 7. Comparison of the success rate of tracking algorithms

As shown in Figure 7, TLD-SLBP algorithm can obtain a higher success rate compared with other tracking algorithms, whose AUC (Under Curve Area) score in the OPE, SRE, TRE improved by 0.100, 0.069, 0.064 compared to MEEM algorithm.

The average performance of TRE is better than the performance of OPE, because the test frames of OPE is less than TRE which test frame number from the first to the last fragment. Since the tracker tends to perform better in a shorter sequence, the average value of all the results in TRE may be higher. On the other hand, the average performance of SRE is lower than that of OPE. Initialization error can cause the tracker to update with inaccurate information, leading to gradual drift of tracking frame.

Fig.8 gives the comparison of TLD-SLBP, MEEM, Struck, TLD algorithm in Benchmark in AUC. The Benchmark includes a tracking data set, which contains 50 completely labeled sequences, such as basketball, David, Women series and so on. The data set contains 11 kinds of attribute labeling all sequences to classification, which contains light change, scale

change, motion blur and other attributes. The tracking performance of TLD-SLBP algorithm is largely beyond the other algorithms, like dealing with occlusion and deformation.

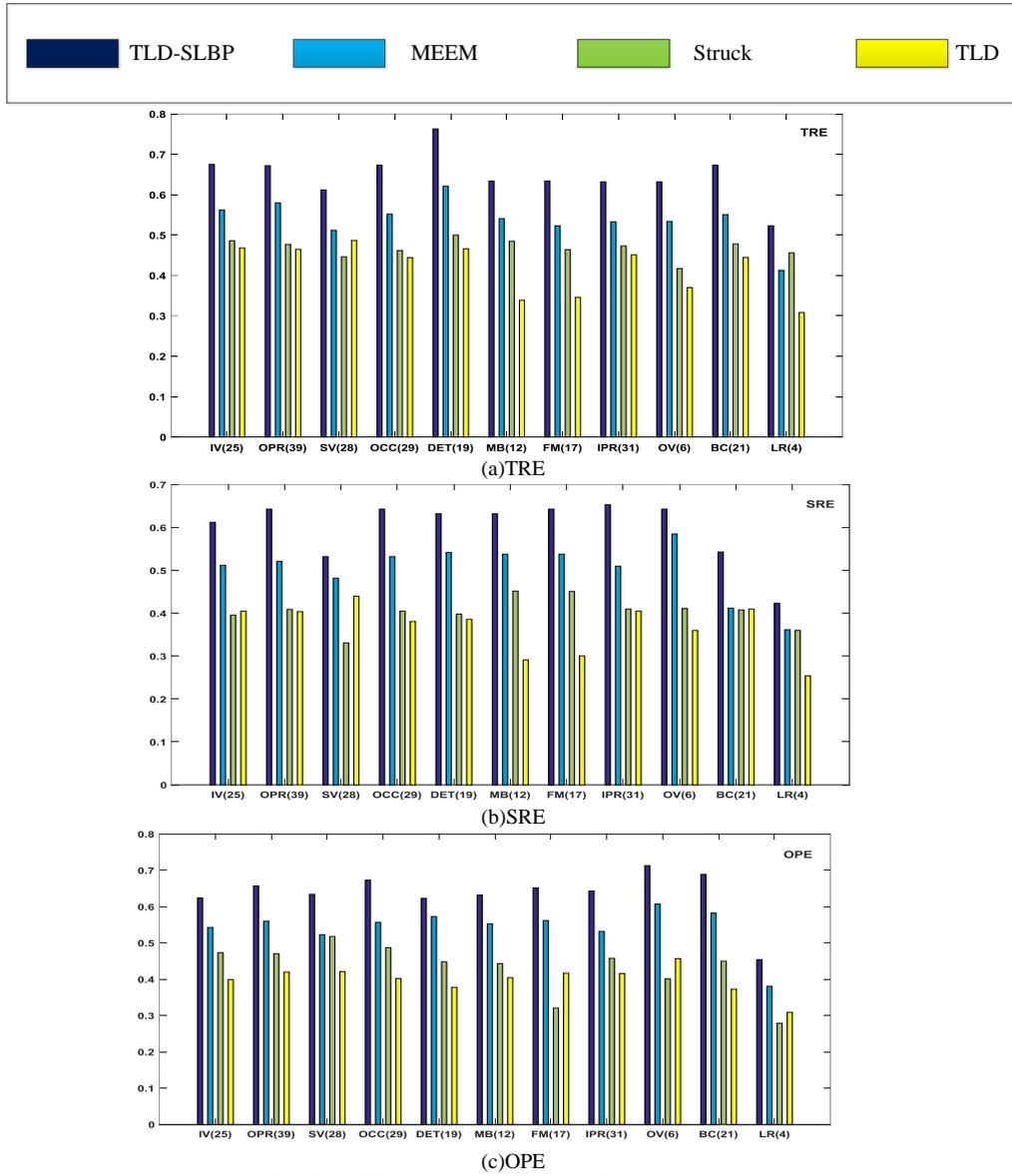


Fig. 8. AUC comparison of four tracking algorithms

### 5.2 Qualitative analysis

The selection of “women”, “car4” and “boxing2” three video sequences from the benchmark website are used as test sequences. A single frame image of “women”, “car4” and “boxing2” sequences are size of  $352 \times 288$ ,  $360 \times 240$  and  $480 \times 360$ , the frame rate are 10 f/s, 5f/s and 10f/s.

Firstly, experiments are carried out on the “women” video sequence, as shown in Figure 9, which are 4, 113, 225, 496, 565, 597 frame respectively. In the process of woman’s walking, human form changes, but is also occluded by the vehicles and trees and foliage, such as covered the upper body or lower body. In addition the background around pedestrians has changed frequently. It can be seen from the Fig.9, in the 113rd frame, the woman has been blocked in vehicle body. In the 565th frame, a pedestrian’s head is covered by the trees, causing the bounding box tracking have deviated from the target. But, after a period of time, the tracker detected the woman successfully, which shows that the algorithm has a good effect on non-rigid objects and occlusion problem.



Fig. 9. The tracking results of women video sequences

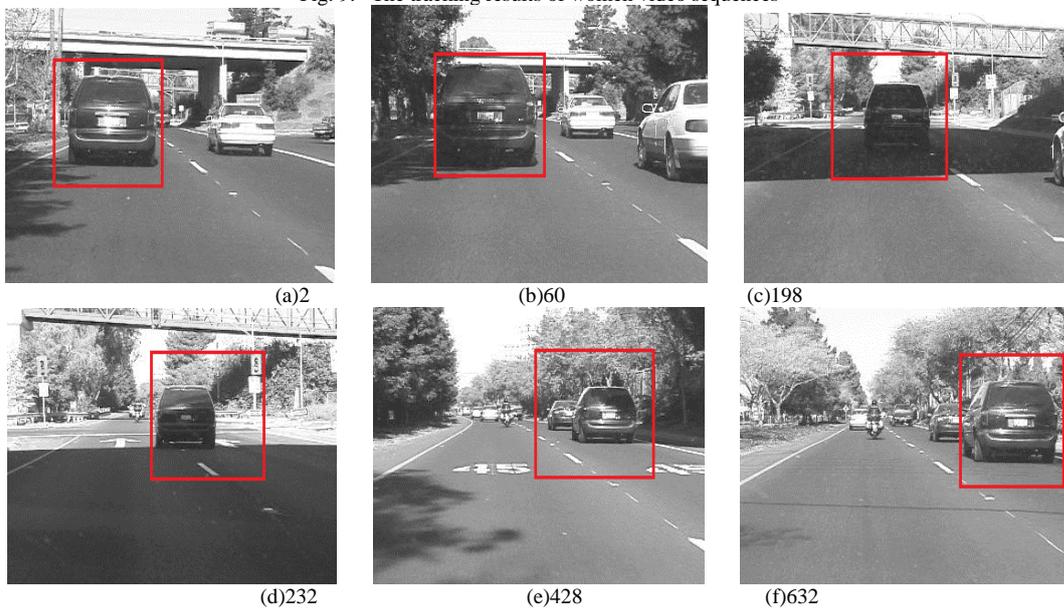


Fig. 10. The tracking results of car4 video sequences

Secondly, experiments are carried out on the “car4” video sequence, as shown in Figure 10, which are 2, 60, 198, 232, 428, 632 frame respectively. In the process of the vehicle’s driving, the scene has changed. In the sixtieth frame, the vehicle is similar in color and easily confused, but this algorithm is able to successfully follow the target. In the 198th frame, when the vehicle enters the aperture, illumination changes. When the vehicle runs out, the bounding box of tracking had an offset, after correction for a period of time, it was gradually close to the vehicle bounding box center. In the 428th frame, by the interference of similar vehicles, the box covered the target and other vehicles, after a period of time, the interference was deleted successfully. The above image frames show that the algorithm can effectively improve the effect of illumination change and complex background.

Finally, experiments are carried out on the boxing2 video sequence, as shown in Figure 11, which are 47, 119, 308, 413, 629, 962 frame respectively. The target for the blue dress is the boxing athlete, who has a larger deformation in the game. The tracking algorithm can accurately follow the target. In the 308th frame, the target is completely covered and the bounding box has lost the goal. But after a period of time, the target has been successfully followed. Therefore, the experimental results show that the algorithm has a good effect on the deformation and occlusion of the target.

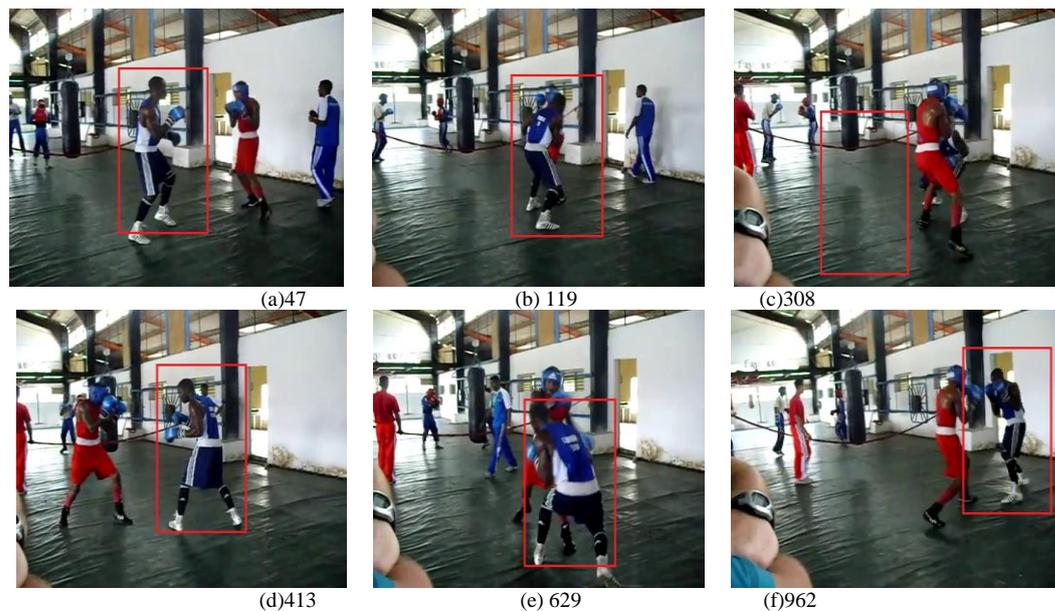


Fig. 11. The tracking results of boxing2 video sequences

## 6. Conclusions

By analyzing the deficiency of TLD algorithm, the object tracking algorithm combining SLBP classifier and TLD is proposed in this paper. By adding the discrimination device into the detection module TLD tracking framework to make the success rate of target tracking improved. By using the video from Benchmark test, the experimental results show that the proposed algorithm can achieve better results than other tracking algorithms in processing illumination change, scale change and occlusion etc. This target tracking algorithm can be applied to a single target for a long time tracking, which greatly limits the scope of application, and further improvement is needed to increase the function of automatic identification and detection of targets.

## References

1. Maliha A, Yusof R, Madani A, "Online sequential-extreme learning machine based detector on training-learning-detection framework", *Control Conference, IEEE*, pp. 1-5, 2015
2. Jia C, Wang Z, Wu X, et al, "A Tracking-Learning-Detection (TLD) method with local binary pattern improved", *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1625-1630, 2015
3. Zhou C, Dong Q, Ma W, et al, "PE-TLD: Parallel Extended Tracking-Learning-Detection for Multi-target Tracking", *Algorithms and Architectures for Parallel Processing*, pp. 665-677
4. Park E, Ju H, Jeong Y M, et al, "Tracking-Learning-Detection Adopted Unsupervised Learning Algorithm", *Seventh International Conference on KnowledgeError! Reference source not found. and Systems Engineering, IEEE*, pp. 234-237, 2015
5. H Grabner, J Sochman, H bischof, et al, "Training Sequential On-line Boosting Classifier for Visual Tracking", *International Conference on Pattern Recognition*, pp.1-4, 2008
6. B Babenko, "Multiple Instance Learning With Manifold Bags", *Journal of Clinical Microbiology*, pp. 2609-2611, 2011
7. Kalal Z, Mikolajczk K, et al, "Tracking Learning Detection". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(7), pp.:1409-1422, 2012
8. Hu J, Hu S, Sun Z, "A Real Time Dual-Camera Surveillance System Based On Tracking-Learning-Detection Algorithm", *Control and Decision Conference, IEEE*, pp. 886-891, 2013
9. Chen L, Zhu S, Li X, et al, "Target tracking via improved TLD algorithm", *Control and Decision Conference, IEEE*, pp. 5942-5947, 2015
10. Hongjun S, Tao Y, "Crane tracking and monitoring system based on TLD algorithm", *IEEE International Instrumentation and Measurement Technology Conference Proceedings*, pp. 1-5, 2016

11. Tsai S Y, Chen C Y, Lee J S, et al, "Evaluation of Effective Dose using TLDs With Different Weighted PMMA Phantoms Undergoing Coronary Artery Calcium Computed Tomography Examination", *IEEE Transactions on Nuclear Science*, Vol.60, No.3, pp. 2147-2154, 2013
12. Z. Ping, S. Youngqi, W. Yali, and Z. Rui, "A Parallel Implementation of TLD Algorithm using CUDA", *ICWMMN2013 Proceedings*, pp. 220-224, 2013
13. Min W P, Jung S K, "TLD based vehicle tracking system for AR-HUD using HOG and online SVM in EHMI", *IEEE International Conference on Consumer Electronics*, pp. 289 – 290, 2015
14. Kong W, Zhang D, Zhao S, et al, "Autonomous track and land a MAV using a modified tracking-learning-detection framework", *Control Conference, IEEE*, pp. 5359-5366, 2015