An Improved Mean Shift Object Tracking Algorithm Based on ORB Feature Matching

Yan Yang[1,2], Xiaodong Wang[1,2*], Jiande Wu[1,2], Haitang Chen[1,2], Zhaoyuan Han[1,2]

1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China
E-mail: 605387243@qq.com

2. Engineering Research Center for Mineral Pipeline Transportation. YN, Kunming 650500, China
E-mail: wjiande@foxmail.com

Abstract: It is critical to accurately track objects for video monitoring of intelligent transportation, so an improved Mean Shift object tracking algorithm based on Oriented FAST and Rotated BRIEF (ORB) feature matching was proposed in this paper. The algorithm based on ORB feature matching can be applied to better locate the object in case of great shifts to the tracking window when object is interfered by complex background or rapidly moving. Subsequently, the object location can be accurately tracked through Mean Shift iteration tracking. The experimental results suggested that this algorithm had effectively solved following problems, including poor anti-interference performance and inaccurate tracking of fast moving objects. Meanwhile, it improved the robustness of object tracking algorithms.

Key Words: Mean Shift, Object Tracking, Oriented FAST and Rotated BRIEF, Feature Matching, Robustness

1 INTRODUCTION

Being critical for video monitoring of intelligent transportation, object tracking is the prerequisite for detecting traffic accidents by video, acquiring traffic information, tracking and locating vehicles and so on, so current research focuses on enhancing robustness of object tracking. With the rapid development of research on visual object tracking home and abroad, numerous tracking algorithms have emerged over the past few years, while the most representative were Mean Shift and Particle Filter algorithms [1-4], based on which plenty of algorithms were proposed and great research outcome was thus achieved. With good real-time performances and simple computation, Mean Shift algorithm can be used in the fields where there are high requirements for real-time performances like video monitoring of intelligent transportation, whereas its anti-interference performance is poor. The particle filter algorithm improves its accuracy by massive sampling, thus, its real-time performance is poor.

Concerning Mean Shift algorithm, the object will be easily lost when it is interfered by background or moving fast. Besides, it is impossible to re-locate and further track the object in case of object losses. In view of this, many scholars have put forward some corresponding improved algorithms. H. X. Yang [5] solved the problem on object loss under complicated circumstances and enhanced the robustness of tracking by integrating Kalman filter with Mean Shift, whereas it was still rather difficult to track fast moving objects. In consideration of this, Y. Liao [6] tracked objects by combing particle filter with Mean Shift, proposed a four-direction mechanism to search objects by expanding the search scope and made up the deficiency of unidirectional search for failing to tracking fast moving objects. Nevertheless, objects couldn't be tracked on a real-time basis when they were initially located by particle filter. In literature [7-8], corner features matching was introduced into object tracking. In this way, the anti-interference performances and the adaptability to objects with complicated motions were improved.

However, the detection of SIFT corner feature is completed at low speed because of its complex calculation. Without rotational invariance, FAST isn't effective for matching features although it can detect corners at a high speed. In consideration that Oriented FAST and Rotated BRIEF (ORB) algorithm could effectively match features fast, an improved Mean Shift tracking algorithm based on ORB feature matching was put forward in this paper, to solve the problem concerning unstable tracking and even object loss resulting from objects' interference by background or fast motion. This algorithm increased the accuracy and the robustness of object tracking.

2 FEATURE EXTRACTION

2.1 Extraction of Objects' Features from Color Histograms

Mean Shift is a tracking algorithm based on color feature of object, the preprocessing step of which is convert the image from RGB(red, green, blue) Color Space to HSV(Hue, Saturation, Value) Color Space to reduce the impact caused by the change of light and establish the color histogram of the hue component in object region after statistic.

Suppose there are n pixels within object region which are described as $\{x_i\}_{i=1\cdots n}$, the color histogram is consist of m bins, the eigenvalue of object is described as $u=1...m$, and

*Xiaodong Wang is the corresponding author. This work is supported by Technological Innovation Funds for Technology-based Small and Medium-Sized Enterprises (13C26215305546), Science & Research Program of Yunnan province (No. 2011CI017).

then the probability density of eigenvalue within object region can be described as follows:

$$\hat{q}_u = \frac{\sum_{i=1}^{n} k\left(\left\|\frac{x_0 - x_i}{h}\right\|^2\right)\delta\left[b(x_i) - u\right]}{\sum_{i=1}^{n} k\left(\left\|\frac{x_0 - x_i}{h}\right\|^2\right)} \tag{1}$$

Where x_0 presents the centric position of object region, the bin of eigenvalue of pixel x_i is calculated by function $b(x_i)$, $\delta(\cdot)$ is Delta function, the calculation of the value of the eigenvalue of pixel x_i is limited in the uth bin by function $\delta\left[b(x_i) - u\right]$, $k(\cdot)$ is the contour function of kernel function, h is the width of window.

In the same way, the probability density of candidate region $\hat{p}_u(y)$ can be described as follows:

$$\hat{p}_u(y) = \frac{\sum_{i=1}^{n} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)\delta\left[b(x_i) - u\right]}{\sum_{i=1}^{n} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)} \tag{2}$$

Where, y is the centric position of candidate region.

2.2 ORB Features Detection and Description

ORB [9] (Oriented FAST and Rotated BRIEF) was improved based on FAST [10, 11] feature detection and BRIEF [12] feature descriptor. Compared with feature descriptors such as SIFT and SURF, ORB was greatly improved in speed. It can maintain invariance well in case that images change in terms of rotation, scale and illumination, etc.

2.2.1 Detection of Feature Points

FAST detects feature points based on assuming grayscale intensity of pixels outside feature points. It determines if a feature point is acceptable by checking the grayscale of a circle of peripheral pixels. The pixel will be detected to be the feature point as long as its grayscale greatly differs from the peripheral pixels and peripheral pixels forms a complete arc that is longer than 3/4 of the circumference. It was shown in Figure 1.

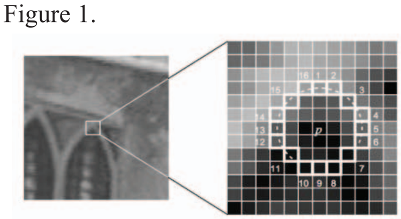

Fig 1.FAST Corner Detection

Based on FAST feature points detection, ORB adopted a scale pyramid, where FAST features were obtained from each layer, so multi-scale features of corners were added. After introducing the grayscale centroid method, shifted vector was considered as the orientation of feature points, so that they could be of rotational invariance. In addition, the feature points extracted by FAST were screened by detecting them with Harris. Thus, corner response functions were determined and the edge response was alleviated.

2.2.2 Feature Description

As regards ORB, the feature points are described with a BRIEF descriptor, whereas BRIEF doesn't have rotational invariance. Hence, BRIEF is oriented along the direction of key points to give steered BRIEF of ORB features. A 2*n-dimension matrix is defined for the n dimensional binary feature set of any features at (x_i, y_i) .

$$S = \begin{pmatrix} x_1, \cdots, x_n \\ y_1, \cdots, y_n \end{pmatrix} \tag{3}$$

Based on the block direction θ and corresponding rotation matrix R_θ , a rotation matrix of S is constructed as follows:

$$S_\theta = R_\theta S \tag{4}$$

Thus, the steer BRIEF operator is obtained:

$$g_n(P, \theta) = f_n(P) \left| (x_i, y_i) \in S_\theta \right. \tag{5}$$

The angles were discretely distributed at (x_i, y_i) , namely a point every 12 degrees, and a look-up table was created to calculate BRIEF. The descriptors of key points would be calculated according to a correct point set S as long as they crossed views in a consistent direction. After obtaining the steered BRIEF, all possible pixel blocks were greedily searched to find at least 256 pairs of slightly correlated pixel blocks. If less than 256 pairs of pixel blocks were obtained, the correlation threshold would be increased to continue the greedy search. At last, required rBRIEF would be obtained.

3 IMPROVED MEAN SHIFT OBJECT TRACKING ALGORITHM BASED ON ORB FEATURE MATCHING

3.1 Similarity Evaluation between Object Model and Candidate Model

Defined by $\hat{q}$ and $\hat{\rho}(y)$ respectively, object model and candidate model were described according to probability density of color histogram. The similarity between the object model and candidate model was measured based on Bhattacharyya coefficient $\hat{\rho}(y)$ as follows:

$$\hat{\rho}(y) = \rho(\hat{p}(y), \hat{q}) = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)\hat{q}_u} \tag{6}$$

Where, $0 \le \hat{\rho}(y) \le 1$. The higher the $\hat{\rho}(y)$, the higher the similarity of probability density between the candidate region and the object region was. On the contrary, there would be greater differences in the probability density between above two regions.

In this paper, the deviation between the object search box and the actual location of object was measured based on the similarity of probability density between object and candidate regions in the process of tracking object. Assuming that $\hat{\rho}(y)$ was lower than ε (a given threshold), it meant that a relatively great shift had occurred to the object tracking.

3.2 ORB Features Matching and Convex Hulls Construction

Great shifts to object tracking will lead to increasingly poorer effects of subsequent tracking and even a loss of objects. Therefore, object tracking was proposed to be improved according to ORB features, so as to track the objects accurately on a non-stop basis. In tracking object, the object should be manually selected at first prior to tracking. Besides, the picture of the object needed to be stored, so that the object could be matched and located in subsequent frames of scenes.

Once ORB features were detected and described for objects and scenes, the objects should be matched by exhaustive nearest neighbor search, while matched pairs needed to be purified. Then, purified matching points were fitted with convex hulls in the scenes, to determine the central coordinate of the hulls z.

3.3 Mean Shift Object Tracking

The initial central location y_0' of candidate objects in current frame was determined according to defined similarity between object and candidate models:

$$y_0' = \begin{cases} y_0 \,, & \hat{\rho}(y) > \varepsilon \\ z \,, & \hat{\rho}(y) < \varepsilon \end{cases} \tag{7}$$

The new location of object y_1 was confirmed by calculating Mean Shift vectors, and the center of object searching window was shifted from y_0' to the new location of object y_1. Mean Shift vectors were calculated by constant iteration, if the Mean Shift vector ($y_1 - y_0$) met following requirement

$$\left\| y_1 - y_0 \right\| < \varepsilon' \tag{8}$$

Then, the iteration converged or stopped in case of maximum iterations. In Formula (8), ε' is a threshold.

3.4 Algorithm Steps

The process of the algorithm proposed in this paper were specifically as follows:

Step 1: The Mean Shift tracking algorithm was initialized. Once object area was selected, the probability distribution of histogram was determined for the object model. The initial location of the object was y_0.

Step 2: The object candidate model were determined for the current frame.

Step 3: Bhattacharyya coefficient $\hat{\rho}(y)$ between object model and candidate model was calculated. In case that $\hat{\rho}(y) > \varepsilon$, ε was experimentally determined to be 0.7 and step 4 was taken. If $\hat{\rho}(y) < \varepsilon$, ORB features should be matched to obtain matching results, construct convex hulls and determine the central position of the hulls z. Meanwhile, the center of object frame was shifted to the centroid position z, y_0' was thus determined to be z. In case of no matching points, step 4 should be taken.

Step 4: Objects were tracked by Mean Shift iteration. The iteration would converge as long as the requirement of Formula (8) was met. Thus, the object's new candidate region y_1 was determined. To avoid great shift to the Mean Shift tracking window under severe interference, if $|y_1 - z| > \varepsilon_1$,then quit the Mean Shift result y_1, y_0 was determined to be z, it was confirmed to be the object's new candidate region and considered as the center of the object's candidate region in the next frame. The next frame was read and step 2 was taken.

The flowchart of the method was shown in Figure 2.

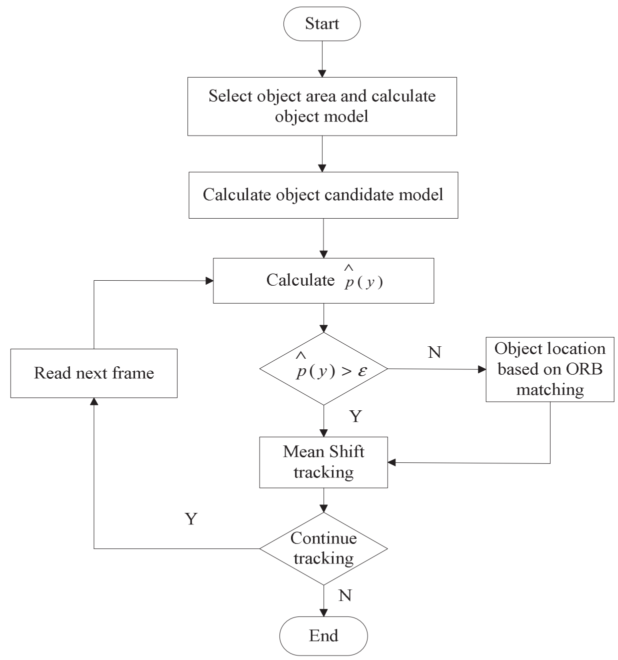

Fig 2.Flowchart of the Algorithm Proposed in this Paper

4 EXPERIMENTAL RESULTS AND ANALYSIS

In this paper, programmes were developed in VS2010 and OpenCV library. To validate the tracking effect and effectiveness of the algorithm proposed in this paper, it was experimentally compared with the Mean Shift algorithm based on color features.

Experiment 1. Video sequences were collected from PETS2000-2007. Figure 3 shows the tracking effect of the Mean Shift algorithm based on color features. No. 869, 880 and 893 frames were extracted from the video sequences. Interference of ground color and surrounding pedestrians' background led to great shift to centroid of object tracking box and poor tracking effect. Figure 4 shows the tracking effect of the improved algorithm proposed in this paper. Tracking objects well on a non-stop basis, the improved algorithm was significantly more effective for tracking objects than the tracking algorithm based on simple color feature. As shown in Figure 5, the location of the object in the scene was determined by ORB feature matching. According to the matching results, ORB-based feature matching was effective for locating objects. Although there were some mismatches with relatively great errors, they would be eliminated and excluded from convex hull construction in the fitted object region, so object location became more accurate.

Fig 3.Object Tracking Effect of the Algorithm Based on Color feature

Fig 4.Tracking Effect of the Algorithm Proposed in this Paper in Experiment 1

Fig 5.Object Location Based on ORB Feature Matching in Experiment 1

Experiment 2. To further verify the robustness of the method proposed in this paper, photographs were taken with unfixed cameras in Experiment 2, so the scenes were somewhat unstable. Furthermore, fast moving objects' motocross sequences were studied in the experiment and mountain riders were tracked. The Mean Shift algorithm wasn't quite effective for tracking objects and could hardly track them due to wide range of fast motions. However, the object tracking method proposed in this paper enhanced object location by ORB feature matching based on Mean Shift tracking. When great shifts occurred to Mean Shift tracking, objects were located by ORB feature matching, as shown in Figure 6. Besides, objects were accurately tracked by Mean Shift algorithm and good effects were achieved in the experiment, as indicated in Figure 7.

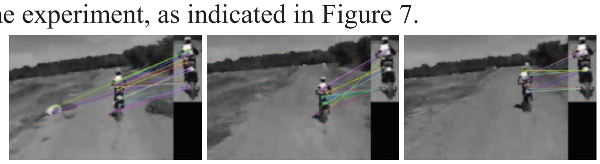

Fig 6.Object Location Based on ORB Feature Matching in Experiment 2

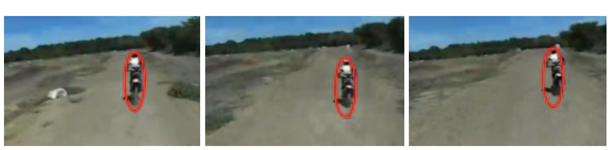

Fig 7.Effect of the Algorithm Proposed in this Paper in Experiment 2

5 CONCLUSION

In this paper, an improved Mean Shift object tracking algorithm based on ORB feature matching was put forward. With the application of this method, objects can be further tracked once they are lost due to interference of complex background or fast motion of object during video monitoring of intelligent transportation. As great shifts occur to object tracking, the object location can be enhanced by ORB feature matching. In addition, ORB meets the requirements for real-time video tracking since it is effective for matching features at a high speed. The experimental results suggested that this algorithm could track and locate objects accurately and enhancing the robustness of object tracking.

REFERENCES

[1] Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using Mean Shift, IEEE International Proceedings on Computer Vision and Pattern Recognition, Stoughton Printing House, 142-149, 2000.

[2] H. X. Chu, Z. Y. Xie, J. X. Wang, Mean shift object tracking with spatiogram corrected background-weighted histogram, Control and Decision, Vol.29, No.3, 528-532, 2014.

[3] W. Yu, X. J. Wu, H. Y. Wang, An anti-occlusion method for object tracking based on adaptive particle filter, Journal of Optoelectronics. Laser, Vol.23, No.11, 2207-2214, 2012.

[4] Arulampalam M S, Maskell S, Gordon N, et al. A tutorial on particle filters for on-line nonlinear/non-Gaussian Bayesian tracking, IEEE Trans on Signal Process, Vol.50, No.2, 174-188, 2002.

[5] H. X. Yang, Y. W. Hang, X. Liu, Algorithm for Video Object Tracking Based on Meanshift and Kalman Filter, Journal of Wuhan University of Technology(Information & Management Engineering), Vol.34, No.2, 147-150, 2012.

[6] Y. Liao, H. Zhou, Z. H. Liang, Y. Zhang, J. H. Liu, L. Su, Tracking Object Based On Particle Filtering And Mean Shift, 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery, 559 - 564, 2011.

[7] R. Dong, B. Li, Q. M. Chen, Multi-degree-of-freedom mean-shift tracking algorithm based on SIFT feature, Control and Decision, Vol.27, No.3, 399-402, 2012.

[8] M. Wang, Y. P. Dai, Q. L. Wang, A Novel FAST-Snake Object Tracking Approach, Acta Automatica Sinica, Vol.40, No.6, 1108-1115, 2014.

[9] RUBLEE E, RABAUD V, KONOLIGE K, ORB: an efficient alternative to SIFT or SURF, IEEE International Conference on Computer Vision, 564-2571, 2011.

[10] ROSTEN E, DRUMMOND T, Fusing points and lines for high performance tracking, IEEE International Conference on Computer Vision, 1508-1515, 2005.

[11] ROSTEN E, DRUMMOND T, Machine learning for high-speed corner detection, In European Conferenee on Computer Vision, 430-443, 2006.

[12] CALONDER M, LEPETIT V, FUA P, Brief: binary robust independent elementary features, European Conference on Computer Vision, 778-792, 2010.