

Improvement and Realization of CamShift Algorithm Based Motion Image Tracking

Wang Yubian*

Department of Railway Transportation Control
Belarusian State University of Transport,
34, Kirova street, Gomel, 246653, Republic of Belarus

*is the communication author.

e-mail: alika_wang@mail.ru

Yuri Shebzukhov

Department of the International Relations
Belarusian State University of Transport,
Republic of Belarus

34, Kirova street, Gomel, 246653, Republic of Belarus
e-mail: oms@bsut.by

Abstract—The detection and tracking technology of moving object image is one of the key technologies of computer vision, and is widely used in many fields such as industry, transportation, and military. The detection and tracing of moving object in the motion scenes based on the UAV platform is a technical difficulty in this field. In practical applications, the characteristics of complicated environment, small target object and moving platform require higher real-time performance and reliability of the algorithm. If it is possible to add some other features of the target to the tracking process, it will be possible to improve the shortcomings of Camshift itself. Based on the Camshift tracking algorithm, this paper integrates SURF feature detection into the algorithm, which greatly improves the tracking accuracy of the target object, and also has better real-time performance which can achieve better tracking performance of the object.

Keywords—*Motion Image; CamShift Algorithm; SURF; Feature Detection*

I. INTRODUCTION

Video tracking is the key technology for motion target object detection in dynamic scenes on UAV platform. It can be realized by two methods: one is based on target recognition technology, of which the core concept is frame-by-frame recognition algorithm for motion video to identify the target object and determine the target matching. The other is based on the detection technology of the moving object, of which the core concept is the active detection of the moving object, and the position of the

moving object is determined in accordance with the detection result to realize the tracking. This method can achieve the tracking of any moving object without the need of complicated priori information for detection, such as characteristics of object shape, object sizes. However, the tracking effect of various tracking algorithms also depends on the background migration of the object, the unpredictable tracking path, the unpredictable target motion path and mode, the scene switch, the target object movement does not have an analyzable pattern, the change of camera model, the camera shift, and the change of illumination condition. And the causes of the changes in the color and shape of the moving object are very different [1]. The current mainstream motion object tracking methods with good tracking performance mainly include feature-based tracking methods, region-based tracking methods, model-based tracking methods, motion estimation-based tracking methods, and contour-based tracking methods.

The detection and tracking algorithm of the conventional motion object is only suitable for the scene with static or almost static background, but it is not suitable for the detection and tracking of moving objects in the UAV video. Therefore, the digital sequence information of the video images acquired by the drone should meet the real-time, Accurate, robust and other requirements [2]. The current research shows that the target object tracking method based on Camshift algorithm can meet the requirements of drone video target tracking.

The CamShift algorithm performs target object recognition and tracking by analyzing the Hue component

information of the target region in the HSV color space. The target has low deformation sensitivity. The algorithm has good real-time performance, little computation and low complexity. Therefore it has been extensively studied.

In the video target tracking by drone, the Camshift algorithm [3], comparing to other tracking methods, has many advantages and disadvantages. Under the condition such as the similar color setting to the target object in the background or in the complicated scene, the Camshift algorithm may have tracking error or fail to track because of the tracking characteristics that the algorithm is based on the color information of the moving object. If other auxiliary features of the object can be obtained during the search process and input conditions of algorithm are specified, it is possible to make up for the problems caused by Camshift in such scenes. Because the SURF algorithm has the advantage of good object recognition, the implementation of the SpeedUp Robust Features (SURF) in the Camshift algorithm will greatly improve the tracking accuracy and the tracking reliability of the object. The improved method ensures the good real-time performance of the tracking, greatly improves the tracking accuracy of the object by UAV, and eventually achieves a better tracking effect of the moving object.

II. THE CHARACTERISTICS OF SURF

The tracking of the characteristics of the target object is based on the tracking of point of interest, which is often used in engineering applications and it works well. The difficulty of this method lies in the selection and extraction of features [4]. The selection of features should be able to fully cover the key features of the target object in different scene settings, and such features shall also be convenient for extraction. In general, if the number of sampling points is insufficient in the process of extracting features, it is easy to lose tracking of the object and the tracking effect would deteriorate. On the contrary, the calculation amount and complexity will be greatly increased, and unable to satisfy the actual application.

Although the Harris corner detection method is a traditional point of interest detection method, due to its

fixed scale characteristics, it is becoming difficult to determine the change of position of the target object between image frames when the target object is deformed or there is change in dimensions. Prof. David G. Lowe of the University of British Columbia in Canada first proposed the Scale Invariant Feature Transform (SIFT) [5]. The SIFT algorithm for querying feature key points is to use feature detection within the constructed scale. The orientational selection of point of interest is calculated according to the gradient of its neighborhood, thus achieving the goal that the feature in the scale does not change in accordance with the orientation. However, the algorithm requires high computational complexity, and has high requirements for hardware devices, and a shortcoming of poor real-time computation. On the basis of this algorithm, Yan Ke et al. introduced the Principal Components Analysis (PCA) in the SIFT system and proposed the PCA-SIFT algorithm, which profoundly improved the matching inefficiency of the common SIFT algorithm. However, this method will lead to the failure of feature extraction in the later stage as well as the deterioration of the characteristic distinctiveness.

On this basis, some people have proposed a point of interest algorithm based on SURF [6]. The extracted features are used as the key local features of the images. The speed of the fast calculation is improved by the integral image method. Then the point of interest obtained by Haar wavelet transform are used to obtain the main orientation and its feature vector [7]. Finally, the Euclidean metric is calculated to verify the matching effect between the images. The SURF feature is invariant in the presence of changes in brightness, translation, rotation, and scale. In addition, the method does not show any negative effect to its robustness under noise interference conditions even if the viewing angle changes. This method not only realizes the accurate feature recognition, but also reduces the computational complexity, which greatly improves the inefficiency of the method in use, and has broad application.

III. THE SURF ALGORITHM

A. Constructing Hessian matrix

SURF relies on approximation of the determinant images of Hessian matrix. The Hessian matrix is the core of the SURF algorithm [8]. First, the Hessian matrix of each pixel is calculated according to the equation, and then the Hessian matrix discriminant is used to determine whether the point is an extremum. Gaussian filtering is used to construct the Gaussian pyramid [9]. Compared with the Gaussian pyramid construction process of SIFT algorithm, the speed of SURF algorithm is improved. In the SURF algorithm, the image space occupancy of each set is different. The previous set of images is downsampled by 1/4 to obtain the next scale. At the same time, the same set of images has the same size, and the difference is that different scales σ are used. In addition, in the process of blurring, the Gaussian template size is always constant, and only the scale changes σ . For the SURF algorithm, the image size remains the same, only the Gaussian blur template size and scale σ need to be changed.

B. Preliminary determination of point of interest using non-maximum suppression

Each pixel processed by the hessian matrix is compared with its 26 points in the three-dimensional domain. If it is the extremum of these points, it shall be selected as a preliminary point of interest for the next process [10]. The detection process uses a filter corresponding to the size of the scale image to detect. In this paper, a 3×3 filter is used as an example for detection analysis. The point of interest for detection which is one of the 9 pixel points in the scale and the remaining 8 points in the scale is compared with the 9 pixel points in each of the adjacent two scale layers above and below it to complete the calculation of 26 points in the three-dimensional domain.

C. Precisely locate extremum

The three-dimensional linear interpolation method removes the points whose values are less than a certain threshold, but the sub-pixel point of interest can be obtained, and the extremum be increased, the number of detected point

of interest is therefore reduced, and finally several highly-featured points are detected and the amount of work is reduced [11].

D. Select the main orientation of the point of interest

In order to ensure the rotational invariance, the gradient histogram is not counted in the SURF, but the Haar wavelet response around the point of interest region is computed [12]. That is, centering on the point of interest, within the circular neighborhood of a radius of $6s$ (s is the scale of the point of interest), the Haar wavelet side length is $4s$, and the Haar wavelet response of all points in the 60 degree fan in both x - and y -directions is computed. Sum, and assign a Gaussian weight coefficient to the Haar wavelet responses, to improve the contribution of the response close to the point of interest, suppress the contribution of the response away from the point of interest, and then add the response within the range of 60 degrees, eventually yielding a new vector, traversing the whole circular region. The main orientation of the point of interest was defined by the orientation of the longest vector [13].

E. Construct descriptor of the SURF point of interest [14]

Extract a square region around the point of interest. The size of the window is $20s$. The orientation of the region is the main orientation detected in step 2.4. The region is then divided into 16 sub-regions, each of which the Haar wavelet responses of the x - and y -direction of 25 pixels are summed, where the x - and y -directions are relative to the main orientation. The Haar wavelet response is the sum of the horizontal direction, the sum of the absolute values in the horizontal direction, the sum of the vertical directions, and the sum of the absolute values in the vertical direction.

IV. EXTRACT AND MATCH THE POINT OF INTEREST

(1) Select a frame from the drone tracking video and extract the points of interest using SURF detection method, as shown in Figure 1.



Figure 1. Extract points of interest

(2) Matching the target region. After selecting the target region, extract the target regions in the adjacent two frames and match the points of interest of the two frames. In Figure 2, it can be seen that 6 points of interest are successfully matched.



Figure 2. Matching the points of interest

V. VERIFICATION

After manually selecting the target window, the feedback mechanism is used to calculate the color similarity between the Camshift tracking window and the initial window, and the feature similarity between the SURF tracking window and the initial window. Suppress the major displacement, and assign the weight of displacement dynamically to the two tracking algorithms in accordance with Bhattacharyya distance [15]. The Camshift tracking algorithm is preferred when the motion tracking is stable, and otherwise, the SURF tracking method [16] is preferred.

The experiments show that the tracking method is ideal for tracking, and it can solve the problem of tracking interference caused by background changes and object similarity. Takes a picture of tracking object every 15 seconds, and an example is shown in Figures 3, 4, 5, and 6.

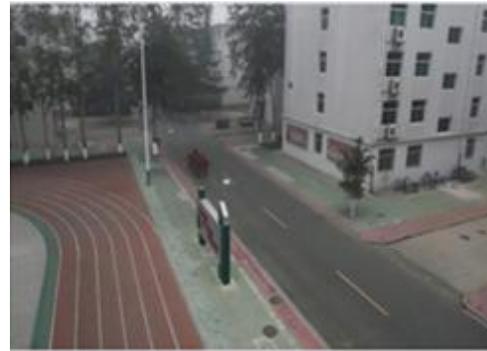


Figure 3. Tracking image 1

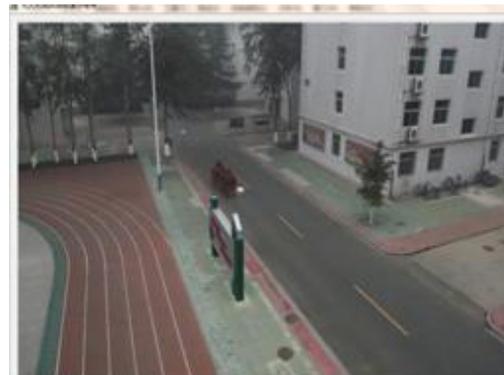


Figure 4. Tracking image 2

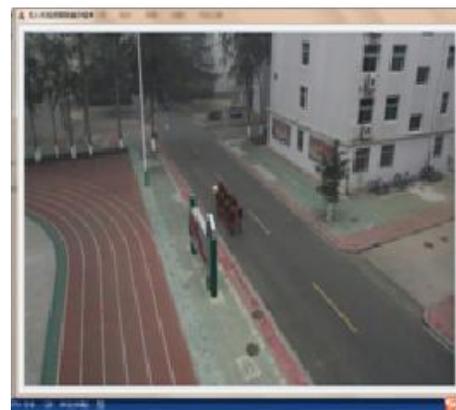


Figure 5. Tracking image 3

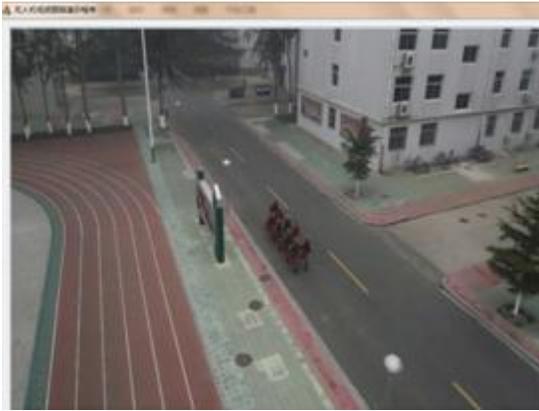


Figure 6. Tracking image 4

As shown in the above figures 3, 4, 5, and 6, the drone has achieved good results in tracking the moving object.

VI. CONCLUSION

Based on the classic tracking algorithm Camshift and focused on the deficiencies of the algorithm in the motion object tracking of drone video, this paper proposes the combination of Camshift algorithm and SURF feature detection to realize the tracking of motion object in UAV video. The experimental results show that the proposed method can effectively track and locate the target object in the more complex aerial photography background. The experiment has achieved good results, basically realizes the tracking of the moving object, and the tracking speed is fast, thus the real-time performance is satisfactory and the timeconsumption is little.

However, in practical applications, the environment of the object tracking is more complicated and diverse. The further study of this paper can be carried out in the following respects: This paper only studied the tracking of a single object, which does not involve the tracking of multiple similar objects. The multi-object tracking has practical significance in video surveillance, intelligent traffic detection, air formation, and geographic monitoring. Therefore, it is very necessary to further study the tracking of multiple objects; the object tracking system in this paper needs to be improved, including optimization of the logic

system and add the feature of parallel processing to improve the tracking efficiency of the system. The Camshift algorithm in this paper is still based on color histogram, which is sensitive to changes in illuminance and the change in color of objects. When the resolution of the camera is not high and the ambient light is insufficient, the tracking effect is not good, therefore the study should focus on the tracking of physical features that are insensitive to illumination.

REFERENCES

- [1] Liu yanli, tang xianqi, Chen yuedong. Application research of moving target tracking algorithm based on improved Camshift [J]. Journal of Anhui Polytechnic University, 2012, 27(2):74-77.
- [2] Xiong TAN, Xuchu YU, Jingzheng LIU, Weijie HUANG. Object Fast Tracking Based on Unmanned Aerial Vehicle Video[C]//Proceedings of PACIA, IEEE Press, 2010:244-248.
- [3] C. Harris, M. J. Stephens. A Combined Corner and Edge Detector[C]. Proco of the 4th Alvey Vision Conf, 2013: 147-152.
- [4] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2014, 60(2): 91-110.
- [5] C. Harris, M. J. Stephens. A Combined Corner and Edge Detector[C]. Proco of the 4th Alvey Vision Conf, 2015: 147-152.
- [6] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2014, 60(2): 91-110.
- [7] Liu yawei. Review of target detection and tracking methods in uav aerial photography video [J]. Airborne missile, 2016.(9):53-56.
- [8] Yan K, Sukthankar: a more distinctive representation for local image descriptors[C]//Proceedings of CVPR, Los Alamitos, IEEE press, 2014:268-235.
- [9] Bay H, Ess A, Tuytelaars T. Speeded up robust features (SURF). [J] Computer Vision and Image Understanding, 2007, 110(3):346-359.
- [10] Leutenegger S, Chli M, Siegwart R. BRISK: binary robust invariant scalable keypoints. [C]//Proceedings of ICCV, IEEE Press, 2013:326-329.
- [11] Cui zhe. Image feature point extraction and matching based on SIFT algorithm [D]. Xi'an: XiDian: university. 2016:38-46.
- [12] Yu huai, Yang wen. A fast feature extraction and matching algorithm for uav aerial images [J]. Journal of electronics and information technology, 2016, 38(3):509-516.
- [13] Wang jianxiang. Research on key technologies of low altitude photogrammetry of unmanned airship and practice of large scale map formation [D]. Xi'an: chang'an university, 2011:36-48.
- [14] Li yifei, research on PID control in four-rotor aircraft, [J] technology and market, 2016.07:90-91.
- [15] Li xiang, wang yongjun, li zhi, misalignment error and correction of attitude system vector sensor, [J] journal of sensor technology, 2017.02:266-271.
- [16] Wang donghua, yue dawei. Design and implementation of large remote sensing image correction effect detection system, [J] computer programming skills and maintenance, 2015.12:118-120.