

# A Tracking Drift Correction Method Based on Feature Points Matching

Xinming Zhou

The People's Liberation Army Unit 91550, Dalian, 116023, China

Email: ZhouXMDL@163.com

**Keywords:** Tracking Drift; Cross-Correlation Template Matching; Feature Point Matching

**Abstract:** In order to improve the efficiency and stability of tracking drift correction, a tracking drift correction method based on feature point matching is proposed in this paper. Tracking drift detection is judged by the similarity between real-time image target template and reference image target template. If the similarity degree is low, the tracking drift phenomenon occurs and the current tracking results need to be corrected. The affine transformation parameters of the middle target are calibrated and then matched with the reference template to determine the correction results of the tracking points.

## 1. Introduction

Influenced by the target motion, background change and camera motion, the shape of the target in the sequence image often changes greatly. Tracking directly using the gray features of the target often leads to the tracking point gradually deviating from the actual position of the target, and the tracking drift phenomenon occurs, which seriously affects the tracking of the target, especially in the image sequences [1]. It is particularly important to track drift correction in the field of tracking measurement that requires accurate solution.

## 2. The Basic Principle of Tracking Drift Correction

Tracking drift usually occurs in image sequences with large changes in the contour of the target. When the contour of the target changes little, the gray features of the target can be considered to be basically stable. A feasible stable tracking method is to update the template frame by frame within the allowable range of tracking accuracy [2]. Once the template changes beyond the set threshold, the tracking point is corrected accurately, which ensures the tracking speed and satisfies the tracking stability. Tracking drift detection can be judged by the similarity between the target template  $T_{ref}$  of the reference image and the real-time template  $T_{cur}$ . For the classical cross-correlation template matching tracking method, if the value of cross-correlation function between  $T_{ref}$  and  $T_{cur}$  is less than a certain threshold, it is considered that drift occurs and the current target tracking results need to be corrected. In the drift correction step, the affine transform parameters of the target are calculated by the correspondence between the target region  $R_{ref}$  of the reference image and the feature points of the real-time image  $I_{cur}$ , and the tracking results are corrected accordingly. Template matching tracking method, the whole target stable tracking method can be divided into three parts: template matching tracking target, tracking drift detection and drift correction.

## 3. Steps of Algorithm Implementation

Step 1: Template matching tracking target

In this paper, the classical cross-correlation template matching method is used. The template is selected as the local region of the target tracking point in the image. In order to improve the computational efficiency, only the translation motion of the target in the image is considered. The value of the cross-correlation function is calculated pixel by pixel in the real-time image search area. The position of the peak value of the correlation function is the target tracking node. The cross-correlation function in this paper takes the form of standardized covariance correlation [3]. As

shown in (1),  $Q$  is the template region, and  $Cor$  represents the standardized covariance correlation function. During the tracking process, the target template is updated by frame.

$$Cor(x, y) = \frac{\sum_{(i,j) \in Q} [f(x+i, y+j) - f_m][g(i, j) - g_m]}{\sqrt{\sum_{(i,j) \in Q} [f(x+i, y+j) - f_m]^2 \sum_{(i,j) \in Q} [g(i, j) - g_m]^2}} \quad (1)$$

### Step 2: Tracking drift detection

Tracking drift detection is determined by the similarity between the real-time template and the target template of the reference image, such as Eq. (2). In order to simplify the calculation, the similarity evaluation index here can directly use the standard covariance correlation function value.

The threshold  $T_h$  represents the tolerance of the algorithm to drift, such as  $T_h=0.95$ , indicating that the tolerance is small.

$$drift = \{cor(T_{cur}, T_{ref}) > Th : True, False\} \quad (2)$$

### Step 3: Drift correction

The drift correction consists of three steps: feature point matching, affine parameter calculation and target relocation.

Feature point matching is one of the hotspots in computer vision. It is widely used in image matching, 3D reconstruction and image retrieval. It mainly involves three key links: feature point extraction, feature point description and feature matching. Krystian Mikolajczyk [4] has made a comprehensive summary of this, and has proved that "MSER + SIFT" is an optimal feature extraction and description method, that is, extracting the stable feature points from the image by MSER (Maximal Stable Extremal Region) operator, and then using SIFT (Scale Invariant Feature Trans) Form) the operator describes the feature point. MSER is a common feature point extraction method. A watershed-like segmentation algorithm is used to extract the maximum stable extremum region with good affine invariance. The ellipse is used to fit the shape of the region and the center of the ellipse is taken as the feature point. The specific implementation process can be referred to in [5]. SIFT descriptor is the most classical feature point description method in image matching. It has the advantages of scale invariance and rotation invariance. It also has strong adaptability to affine transformation.

Generally speaking, the field of view and depth of field of field of view of the target in the image are small, and the change of the target area can be described by affine transformation approximation [6]. The affine transformation parameters of the target can be calculated by the correspondence between the target region of the reference image and the feature points of the real-time image, such as Eq. (3):

$$X' = A \cdot X + B \quad (3)$$

In the above formula,  $X$  and  $X'$  represent the coordinates of the matching feature points of the target region of the reference image and the real-time image respectively,  $A$  and  $B$  are affine transformation parameters.

Object relocation usually uses affine transformation parameters to transform real-time image  $I_{cur}$  to get  $I'_{cur}$ , then the shape of the target in the reference image region is basically the same as that in  $I'_{cur}$ . In  $I'_{cur}$ , the standard cross-correlation matching method can be directly used to relocate the tracking points, and then determine the location of the target in the real-time image.

In order to cope with the continuous change of target morphology, the target template  $T_{ref}$  and the target region  $R_{ref}$  of the reference image need to be updated according to the real-time image and the current tracking results. Considering the change of target scale, the affine parameters can be decomposed into the product of rotation transform and scale transform by SVD decomposition method, such as Eq. (4):

$$A = R(\theta)R(-\phi) \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R(\phi) \quad (4)$$

$R(\theta)$  and  $R(\phi)$  represent the rotation transformation,  $\lambda_1$  and  $\lambda_2$  represent the scale change factor, reflecting the scale change ratio of the target in the X and Y direction, according to which the size of  $R_{ref}$  and  $T_{ref}$  can be adjusted accordingly.

#### 4. Experiment and Results

Fig. 1 describes the processing of the algorithm for the aircraft target sequence image. The target tracking position is shown as a cross wire. The lower right corner of the image is a magnified image of the target local area. In the sequence images, the shape and position of the aircraft changed greatly.

The initial position of the target is shown in Fig. 1(a), and the subsequent frames are processed by the cross-correlation template matching tracking method which updates the template frame by frame. The target tracking result of the 36th frame image is shown in Fig. 1 (b). At this time, the correlation coefficient between the real-time template and the target template of the reference image is lower than the set threshold, and the drift correction step is turned into. The affine transform parameters of the target are estimated by the feature point matching method, and the real-time image is transformed accordingly. As shown in Figure 1 (c), it can be seen that the shape of the target in Figure 1 (a) is basically the same as that in Figure 1 (c). Therefore, the corrected target tracking result is determined as shown in Figure 1 (d).

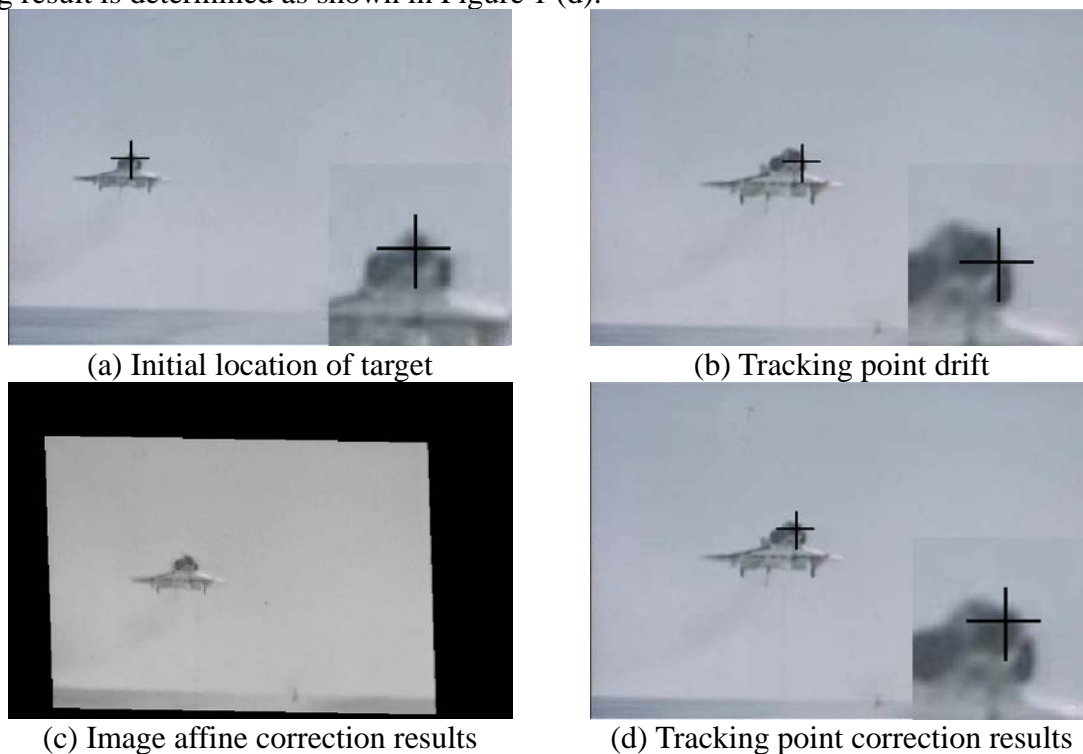


Fig.1. Correction of aircraft target image drift

Comparing the proposed algorithm with the frame-by-frame template updating cross-correlation template matching tracking method, the target rotation occurs in the sequence image. The tracking results of each algorithm are shown in Figures 2 and 3. It can be seen that compared with the traditional template matching tracking method, the proposed algorithm can better adapt to the appearance of the target. The drastic changes of life effectively inhibit the phenomenon of tracking drift.



(a) First frame                      (b) Thirtieth frame                      (c) Seventy-fifth frame

Fig.2. Target tracking results based on traditional template matching method



(a) First frame                      (b) Thirtieth frame                      (c) Seventy-fifth frame

Fig.3. Target tracking results of algorithm in this paper

## 5. Conclusion

The key of the drift correction algorithm is to calculate the affine transformation parameters of the target. The traditional drift correction method directly uses the method of real-time template registration and reference template registration. The adaptability of the target morphology and background illumination changes is poor. This algorithm uses the method of feature point matching to calculate, and can reliably. The parameters of target affine transformation are solved, which have strong adaptability to target movement and background changes.

## References

- [1] M. Elad, M. Aharon. Image denoising via sparse and redundant representations overlearned dictionaries. *Image Processing, IEEE Transactions on*, 2016, 15(12):3736-3745.
- [2] K. Mikolajczyk, C. Schmid. An Affine Invariant Interest Point Detector[C]. *Proc. Seventh European Conf. Computer Vision*, 2012: 128~142.
- [3] Matas, J., Chum, O., Urban, M.. Robust wide-baseline stereo from maximally stable extremal regions [J]. *Image and Vision Computing*, 2014, 22(10): 761~767.
- [4] David G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints[J]. *International Journal of Computer Vision*, 2014, 60(2): 91~110.
- [5] Richard Hartley. *Multiple View Geometry in Computer Vision*, 2nd ed.[M], Cambridge University Press,Cambridge, 2013:168-170.
- [6] Lang R, Warwick K. The plastic self organising map[C]. *Proceedings of the 2012 International Joint Conference on Neural Networks*. Honolulu, HI, USA, 2012, 727-732.