

## Research Article

# A Novel Vision-Based PRPL Multistage Image Processing Algorithm for Autonomous Aerial Refueling

Ling Wu <sup>1,2</sup>, Yongrong Sun <sup>1</sup>, Kedong Zhao <sup>1</sup> and Xiyu Fu <sup>1</sup>

<sup>1</sup>College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

<sup>2</sup>School of Mechatronics Engineering and Automation of Nanhang Jincheng College, Nanjing 211156, China

Correspondence should be addressed to Ling Wu; [wuling@nuaa.edu.cn](mailto:wuling@nuaa.edu.cn)

Received 28 September 2021; Revised 23 October 2021; Accepted 27 October 2021; Published 10 November 2021

Academic Editor: Deepak Kumar Jain

Copyright © 2021 Ling Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Autonomous aerial refueling (AAR) technology can increase the flight endurance of unmanned air vehicles (UAVs) effectively. Drogue detection and target tracking method are significant for probe-drogue refueling system in the docking stage. This paper proposes a novel vision-based multistage image processing algorithm of drogue detection and target tracking for AAR. This algorithm divides the whole task into four stages: preprocessor, recognizer, predictor, and locker (PRPL). The adaptive threshold segmentation (ATS) algorithm and support vector machine (SVM) classifier are utilized in preprocessor and recognizer for drogue detection. An improved kernelized correlation filter (IKCF) tracking algorithm and scale adaptive method by window position as well as image resolution adjusted are adopted in predictor and locker for target tracking in complex dynamic environments. Finally, the proposed PRPL multistage image processing strategy is tested using an autonomous aerial refueling testbed. The results indicate that the proposed algorithm achieves high precision, good reliability, and real-time capability compared with conventional algorithms. The average processing time is within 11 ms in various environments, which can meet the requirement for drogue detection and tracking in AAR.

## 1. Introduction

The AAR technology can greatly improve the flight duration and increase the payload of UAVs. It has always been favored by the experts in the military and national defense fields. Compared with manned aerial refueling technology, which has a higher risk factor and greater difficulty in operation, the UAV technology has greatly reduced operational costs and risks [1–3]. There are two major methods for the AAR: the boom-and-receptacle refueling system and the probe-and-drogue refueling system. The operation procedure is different from each other. Firstly, in the boom-and-receptacle refueling method, the tanker is an active part that steers a rigid retractable boom stretched from the rear of it to a socket installed on the UAV. The advantages of this method are that the oil delivery speed is fast and it is not sensitive to air turbulence, while the disadvantages are structure complex and it only fuels one aircraft at a time [4]. Secondly, in the probe-and-drogue refueling method, the tanker drags a refueling drogue linked in the end of the flexible hose, and

then, the refueling probe attached on the UAV is controlled to dock into the drogue. The advantages of this method are that two or more aircrafts can be refueled at the same time and it has simple structure. And the disadvantage is the refueling drogue susceptible to atmospheric turbulence [5, 6]. At present, the probe-and-drogue refueling method is widely used at home and abroad.

There are two key technologies in the docking of autonomous probe-and-drogue refueling system: vision-based drogue detection and target tracking and relative pose measurement. The former is to obtain the exact location of drogue in the visual images [7], and the latter is to establish the relative pose between the UAV and drogue [8]. Nowadays, a lot of research work has carried out on the vision system for probe-and-drogue refueling tasks. It mainly contains two aspects. One is the active vision system and the other is passive vision system [9, 10]. The active vision system introduces obvious features, such as installation of special light sources or marking points on the drogue to assist visual navigation, while the passive vision system makes no changes

to the drogue, only relying on the characteristics of itself for detection and positioning [11–13]. Vision-based Navigation (VisNav) is a commonly utilized active vision system, which has been adopted in several AAR projects. A set of light-emitting diode (LED) beacons are distributed on the profile of drogue, whose light spots are detected by sensing cell installed on the UAV. The beacons are lightened in order, and the sequence of detected beacons is ensured by a communication link [14]. Pollini et al. used an inexpensive charge-coupled device (CCD) camera with an infrared filter to detect LED beacons installed on the drogue [15]. Wang et al. placed a layer of red marking tape on the ring of the drogue to highlight the circular feature [16]. In active vision system, drogue target can be detected in less image information with faster processing speed and higher reliability. However, the main disadvantage is that slight modification of the tanker should be made to afford electrical power for beacons [17].

The passive vision system needs no supernumerary hardware equipment to install on drogue or the tanker. Martinez et al. developed a vision-based position estimation merely relying on the characteristic of drogue itself and employed a Sobel edge template matching strategy for drogue detection [18]. Yin et al. proposed a new drogue detection and target tracking method based on the approximate circle and inner dark feature of refueling port of drogue, in which the edge information is used to obtain image feature [19]. Gao et al. developed a drogue detection method utilizing low rank and sparse matrix decomposition for local multifeatures, in which the sequences of drogue images are disintegrated into the low rank background and sparse maneuvering targets [20]. Qin et al. proposed a coarse and fine two-level detection structure based on the characteristics of the inner refueling port, and the row and column scanning method for the local region of image was adopted to obtain high-precision shape parameters [21]. Huang et al. developed a multidirection closest point searching to extract all possible image areas and row-column scanning method to obtain the inside circle edge points for detecting and tracking drogue [22]. Xu et al. presented a novel algorithm with adaptive boosting and convolutional neural network classifier to detect the drogue in complex environments [23].

The challenge of passive vision system still focuses on improving the reliability and real-time capability. Therefore, it is essential to develop an economic and efficient image processing algorithm with strong robustness and high success rate for the AAR docking task [24].

In this paper, an image processing algorithm is developed specifically for probe-drogue refueling system. A PRPL multistage drogue detection and tracking strategy based on monocular passive vision system is proposed. A high fidelity test platform using full-size drogue and a micro six-rotor aircraft for AAR docking task is employed to evaluate the proposed image processing algorithm.

The rest of this paper is organized as follows. Section 2 presents the PRPL multistage detection and tracking strategy for drogue, including strategy overview, drogue detection method, and target tracking method in detail. In Section 3,

three various experiment scenes are conducted on the imitation test platform for AAR, and results are analyzed in details. Section 4 summarizes the conclusions of this paper and future work.

## 2. PRPL Multistage Detection and Tracking Strategy

*2.1. Strategy Overview.* In the docking stage of the probe-drogue refueling system, the position information of the drogue is calculated through the visual acquisition equipment and the corresponding image processing algorithm. The drogue is connected to the tail of tanker with a flexible hose. It moves irregularly in the wake vortex of tanker and atmospheric turbulence. Therefore, it puts forward higher requirement for the detection and tracking algorithm of drogue. In this paper, the detection and tracking of drogue target are divided into four phases: detection preprocessing phase, detection recognition phase, tracking prediction phase, and tracking locking phase. Accordingly, the preprocessor, recognizer, predictor, and locker are designed to solve the difficulty of detection and tracking of drogue. The PRPL multistage strategy diagram of drogue detection and tracking is shown in Figure 1.

During docking phase, the real-time video data is obtained by camera installed on receiver. The video data is converted into image sequences as  $\mathbf{I} = \{I_1, I_2, \dots, I_n\}$ . Then, the preprocessor is used to deal with the first frame  $I_1$  of image sequence, and the candidate region set  $\mathbf{RS}_p$  of drogue target is generated. The recognizer is used to extract the normalized feature of  $\mathbf{RS}_p$ , and the candidate regions are judged by SVM classifier. Afterwards, the information of position  $\mathbf{P}_1$  and size  $\mathbf{S}_1$  of the drogue area in the image are obtained. According to the detected characteristics of the target region, the target tracker is initialized, and the predictor is adopted to locate the target area quickly. Finally, the locker is adopted to discriminate and update the model of the target position for the subsequent images of  $I_2, I_3, \dots, I_n$ . If the discrimination fails, that is, the target cannot be tracked, the drogue detection process is performed again on the next frame of image and the previous operation is repeated.

As seen in Figure 1, the preprocessor and recognizer are designed in drogue detection stage; and the predictor and locker are designed in drogue tracking stage. Afterwards, the drogue position in image can be obtained reliably in the locker, which can be used to solve the relative pose information between drogue and UAV in the next step for docking task.

*2.2. Drogue Detection Method.* The main objective of drogue detection is to scan the global image and select the target area in the image by recognition. The traditional detection method is to use windows with different scales to scan the entire image. Then, the recognition area library is achieved and the target area is identified. It is time-consuming, and the database information is huge, which is high redundancy of recognition area. Therefore, the detection preprocessor and the recognizer are designed, respectively, for the detection task. The preprocessor is responsible for reducing the

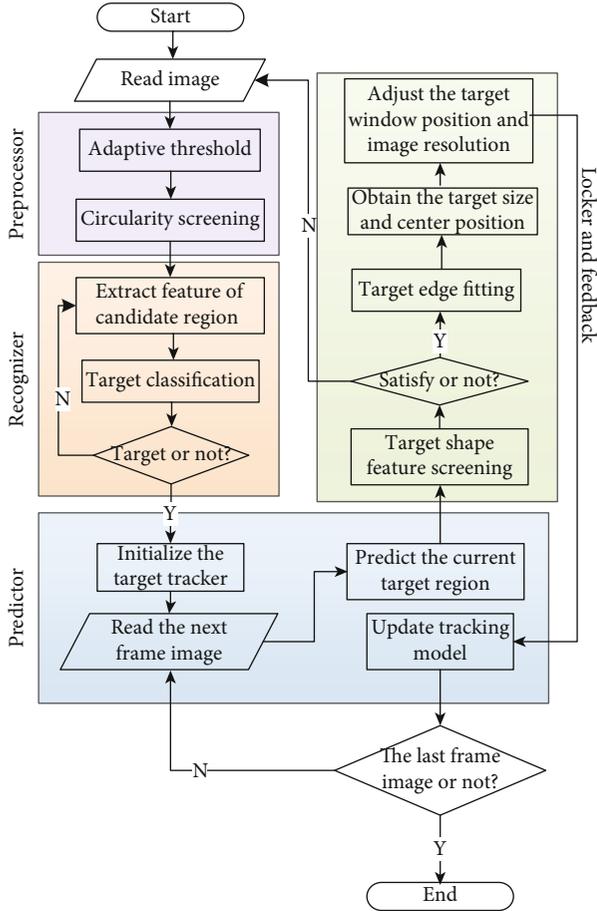


FIGURE 1: The PRPL multistage strategy diagram of drogue detection and tracking.

dimensionality and redundancy of the recognition area database. The recognizer is responsible for determining the target area from recognition database.

**2.2.1. Adaptive Threshold-Based Preprocessor.** Based on the color and shape features of drogue, a cascade classifier is built to generate a detection and recognition library. The detection preprocessing cascade classifier is shown in Figure 2.

The outline diagram of drogue is shown in Figure 3. Due to the black feature of the inner refueling port of drogue, the candidate target area can be obtained by threshold segmentation of the grayscale image, and then, the color feature classifier can be established. Because the change of environmental light will cause the gray value of the target to change greatly, it is important to choose the threshold. An adaptive threshold selection is used.

The image grayscale histogram shows the overall distribution of grayscale values in the image. The histogram changes as light changes. Therefore, a selected threshold based on the grayscale histogram can partly eliminate the impact of lighting change on target detection. The histogram of the grayscale image can be regarded as a one-dimensional matrix  $\mathbf{H} = (h_0, h_1, \dots, h_{255})$ . Among them,  $h_i (i = 0, 1, \dots, 255)$  is the number of pixels whose gray value is  $i$ . The local

minimum of the matrix  $\mathbf{H}$  is used for image segmentation. It can segment the target better from the background.

Firstly, through comparing the current value  $h_i$  with the previous one  $h_{i-1}$ , matrix  $\mathbf{F} = (f_1, f_2, \dots, f_{255})$  is obtained as follows.

$$f_i = \begin{cases} 1, & \text{if } (h_i - h_{i-1}) > 0, \\ -1, & \text{if } (h_i - h_{i-1}) < 0, \\ 0, & \text{if } (h_i - h_{i-1}) = 0, \end{cases} \quad (i = 1, 2, \dots, 255). \quad (1)$$

Then, comparing the new current value  $f_i$  with the next one  $f_{i+1}$ , matrix  $\mathbf{B} = (b_1, b_2, \dots, b_{254})$  is calculated as well. Finally, the local minimum is determined.

$$b_i = \begin{cases} 1, & \text{if } (f_{i+1} - f_i) > 0, \\ -1, & \text{if } (f_{i+1} - f_i) < 0 (i = 1, 2, \dots, 254), \\ 0, & \text{if } (f_{i+1} - f_i) = 0, \end{cases} \quad (2)$$

where  $b_i = 1$  represents the local minimum of the histogram threshold  $i$ .

The adaptive threshold set is obtained as follows.

$$\mathbf{T} = \{b_i \mid b_i = 1, i = 1, 2, \dots, 254\} = \{t_1, t_2, \dots, t_m\}. \quad (3)$$

Then,  $t_1, t_2, \dots, t_m$  are used to judge the validity of the target region in turn until the effective target fitting information is obtained.

Based on the color feature of target set obtained in the previous step, the shape feature classifier is used to classify the regions in the target set. Firstly, the segmented regions are processed with morphological method to remove the influence of noise points. Then, the contour information of remaining effective region is extracted, and the boundary and area information of corresponding region is determined by the contour information. According to the shape feature of a relatively standard circular of the inner fueling port of drogue, the shape feature classifier criteria are established as follows.

$$\text{Length and width ratio} : \frac{l}{w} > t_r, \quad (4)$$

$$\text{Area ratio} : t_a < \frac{4 \cdot s}{\pi \cdot l \cdot w} < t_b, \quad (5)$$

$$\text{Circularity} : \frac{4 \cdot \pi \cdot s}{c^2} > t_c, \quad (6)$$

where  $l$ ,  $w$ ,  $c$ , and  $s$  represent the length, width, perimeter, and area of the target region, respectively.  $t_r$ ,  $t_a$ ,  $t_b$ , and  $t_c$  represent the threshold of aspect proportion, area ratio, and circularity, respectively.

**2.2.2. SVM Classifier-Based Recognizer.** The preprocessor divides the original image into a set of candidate regions for detection. These regions contain several different scale areas.

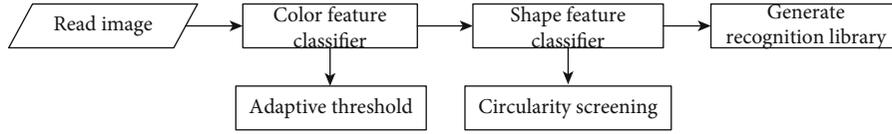


FIGURE 2: Detection preprocessing cascade classifier.

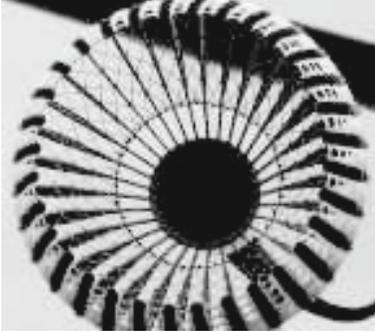


FIGURE 3: Outline diagram of drogue.

A SVM classifier-based recognizer is designed to determine the target region from the candidate regions set quickly and accurately.

According to the drogue image characteristics, which have obvious edge structure, the histogram of gradient (HOG) feature is utilized to extract the feature vector, which usually involves the following steps: gamma color correction, gradient calculation, gradient feature statistics, eigenvector normalization, and eigenvector concatenation. Since HOG features are sensitive to the scale changing, the region scale normalization is used to normalize all candidate regions to  $128 * 128$  pixel. Then, an SVM classifier based on HOG features is designed [25].

Define the sample set to be classified as

$$\mathbf{Z} = \{(\mathbf{x}_k, y_k) | \mathbf{x}_k \in R^m, y_k \in \{-1, +1\}, k = 1, \dots, n\}, \quad (7)$$

where  $\mathbf{x}_k$  is the HOG feature vector,  $y_k$  is the sample classification label, and  $n$  is the number of sample.

The linear SVM classifier is utilized to obtain the maximum interface, that is,

$$g(\mathbf{x}) = \boldsymbol{\omega}^T \mathbf{x} + b = 0, \quad (8)$$

$$\max \frac{1}{\|\boldsymbol{\omega}\|} \quad (9)$$

$$\text{s.t. } y_i(\boldsymbol{\omega}^T \mathbf{x}_i + b) \geq 1, \quad i = 1, 2, \dots, n,$$

where  $\boldsymbol{\omega}$  and  $b$  are the two parameters of the SVM, which can be solved by secondary optimization method [26].

Based on the detection preprocessor and recognizer, the detection process of drogue is shown in Figure 4 when  $t_r = 0.7$ ,  $t_a = 0.8$ ,  $t_b = 1.2$ , and  $t_c = 0.65$ .

**2.3. Drogue Tracking Method.** Drogue target tracking is to model of the detected region feature and predict the position of the target in the next image frame. It is important to improve the robustness, processing speed, and accuracy of the target tracking model. Therefore, the tracking predictor and the locker are designed, respectively, for the tracking task. The predictor is responsible for establishing and updating of the tracking model and predicting the region of the target in the next image frame based on the tracking model. The locker is responsible for judging failure of tracking, extracting the target edge of the predicted region, and calculating target scale position accurately. The target position calculated by the locker will be feedback to the predictor for model update.

**2.3.1. Improved KCF-Based Predictor.** Target tracking is a technology with extremely high requirements for real time and reliability. This paper develops an improved KCF tracking (IKCFT) algorithm based on a highly reliable detection results. Feature extraction and modeling of the target are the key technologies for target tracking, whose speed will influence the real-time performance of tracking. The HOG feature with faster feature extraction and higher robustness and classification model with simple and efficient ridge regression model are adopted. Meanwhile, the training of the model is performed with the positive and negative samples based on the characteristics of fast inversion of the circulant matrix.

The training samples of the tracking model are generated by the cyclic shift of the target region as follows:

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ x_n & x_1 & \cdots & x_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & \cdots & x_1 \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}, \quad (10)$$

where  $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]^T$  is the target sample.

Suppose the training sample set is  $(\mathbf{x}_i, y_i)$ , and its regression function is  $f(\mathbf{x}_i) = \boldsymbol{\omega}^T \mathbf{x}_i$ . The objection of the ridge regression model is represented as follows:

$$\min_{\boldsymbol{\omega}} (\|\mathbf{X}\boldsymbol{\omega} - \mathbf{y}\|^2 + \lambda\|\boldsymbol{\omega}\|^2), \quad (11)$$

where  $\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_n]^T$  is the label value corresponding to each row of feature  $\mathbf{X}$ ,  $\lambda$  is the regularization coefficient, and  $\boldsymbol{\omega}$  is the tracking model parameter.

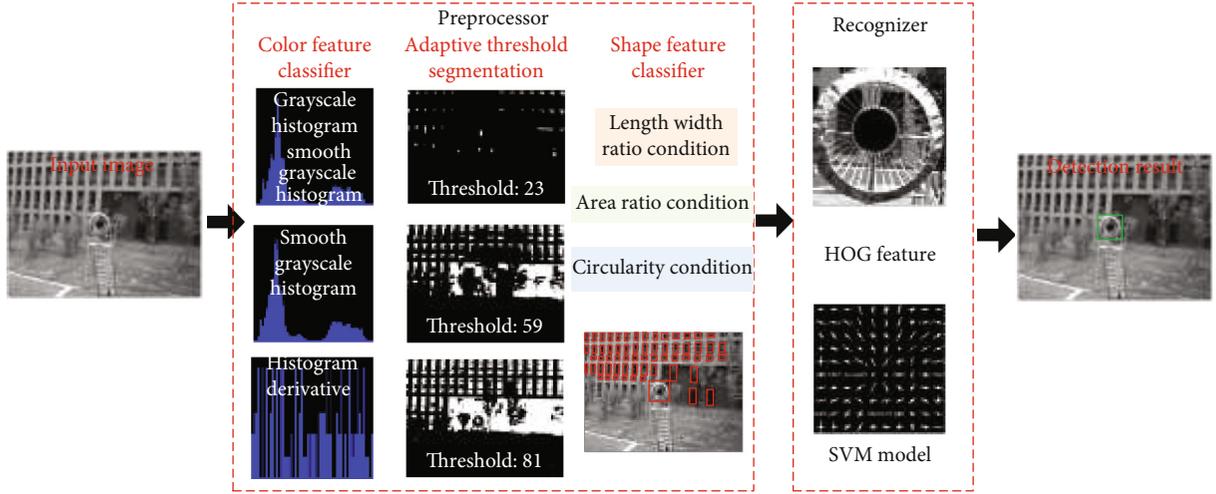


FIGURE 4: Process diagram of drogue detection algorithm.

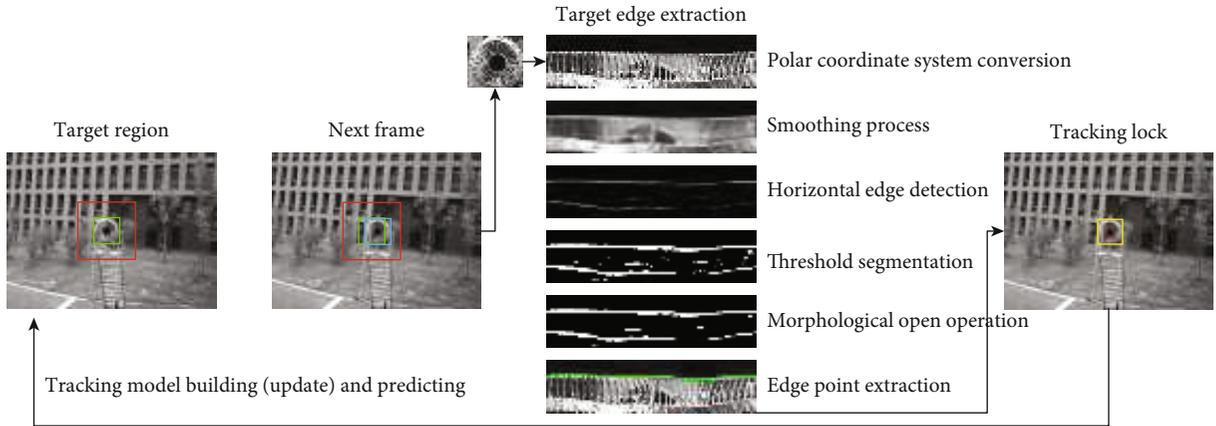


FIGURE 5: The tracking lock phase strategy diagram.

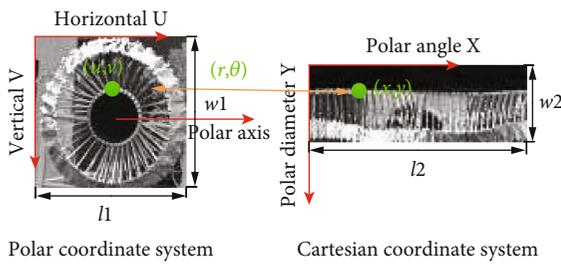


FIGURE 6: Polar coordinate transformation diagram.

By the nature of the circulant matrix,

$$\mathbf{X} = \mathbf{F} \text{diag}(\hat{\mathbf{x}}) \mathbf{F}^H, \quad (12)$$

where  $\mathbf{F}$  is the discrete Fourier matrix. Then,  $\omega$  can be calculated.

$$\omega = (\mathbf{X}^H \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^H \mathbf{y} = \mathbf{F} \text{diag} \left( \frac{\mathbf{x} \wedge^*}{\mathbf{x} \wedge^* \odot \hat{\mathbf{x}} + \lambda} \right) \mathbf{F}^H \mathbf{y}. \quad (13)$$

That is,

$$\omega = \frac{\hat{\mathbf{x}} \odot \hat{\mathbf{y}}}{\mathbf{x} \wedge^* \odot \hat{\mathbf{x}} + \lambda}, \quad (14)$$

where  $\odot$  represents the product of corresponding vector element and  $\mathbf{x} \wedge^*$  is the conjugate vectors of  $\hat{\mathbf{x}}$ .

Because of the nonlinear characteristic of target classification, a kernel space function [27] is introduced to map the feature vectors  $\mathbf{x}$  of nonlinear space to the linear space  $\Phi(\mathbf{x})$ , and the linear regression model is converted to  $f(\mathbf{x}_i) = \omega^T \Phi(\mathbf{x}_i)$ . There are two main advantages of using circulant matrix for tracking. On the one hand, a large number of negative samples can be obtained from the circulant matrix instead of extracting from background. On the other hand, the solution of the ridge regression model parameters becomes a simple and efficient positive sample vector element product instead of a complex matrix inversion operation. It greatly improves the speed of training of model.

The sample set  $\mathbf{Z} = [\mathbf{z}_1 \ \mathbf{z}_2 \ \dots \ \mathbf{z}_n]$  to be detected is obtained from the predicting regions and their displacement. And the largest sample from the final predicted target region

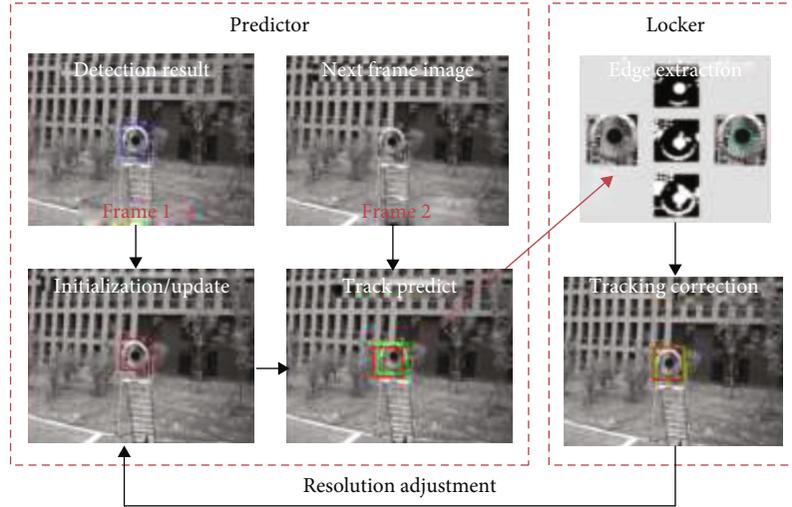


FIGURE 7: Process diagram of drogue tracking algorithm.

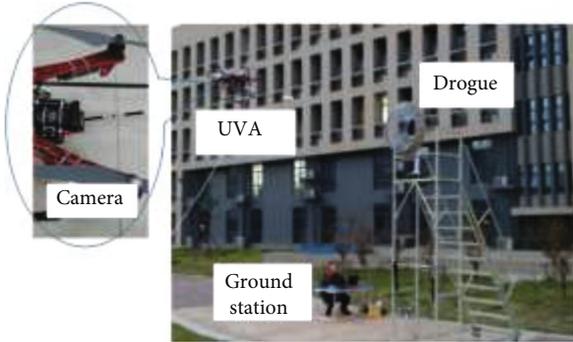


FIGURE 8: Testbed for AAR docking task.

$f(\mathbf{z}_j) = \omega^T \Phi(\mathbf{z}_j)$  is selected as the new detected target region.  $\mathbf{z}_j$  is the offset of the target movement.

**2.3.2. Scale Adaptive-Based Locker.** Due to fixed tracking window size of the KCF tracking algorithm, it is ineffective for large-scale variation in the tracking target. In addition, it is prone to emerge tracking deviation because of tracking accumulated error in long time target tracking. Furthermore, it will track wrong target when the target is out of sight. In the light of the problems above, a tracking and locking feedback process is increased based on the KCF algorithm.

In view of the tracking target will not change large scale in short time, the method of detecting and correcting frame by frame is adopted. This paper adopted the ratio of the inner circle size of the drogue to the tracking window size as the correction criterion. The tracking window size remains uniform during the entire tracking process. The size of the target relative to the tracking window keeps invariable by changing the resolution of the overall image so as to solve the problem of size change of the target. The tracking phase strategy diagram is shown in Figure 5.

As shown in Figure 5, after adjusting the resolution of the input image and predicting by tracker, the tracking region is obtained. Polar coordinate system conversion is

used to change the boundary of the target from a circle to a line to simplify the edge extraction of a circular target. Then, the tracking effectiveness is verified through the extracted target edge and tracking area size information. If the target tracking is successful, the tracking area is corrected and the tracking model is updated.

In this paper, the method of polar coordinate conversion is used to detect the edge of the circular target. The circular edge detection is converted to the line edge detection, and the wrong edge generated by the occlusion can be effectively filtered out. Assuming that the center of the circular target basically coincides with the center of the image area, after the target area is obtained, a polar coordinate system is established with the center of the target area as the pole and the horizontal right as the polar axis. The polar coordinate system can be converted into a Cartesian coordinate system with the  $x$ -axis as the polar angle and the  $y$ -axis as the polar diameter as follows. Polar coordinate transformation diagram is shown in Figure 6.

$$\begin{cases} r = y \cdot \text{delta}_r, \\ \theta = x \cdot \text{delta}_t, \end{cases} \quad (15)$$

$$\begin{aligned} u &= \text{center}_x + r \cdot \cos \theta, \\ v &= \text{center}_y - r \cdot \sin \theta, \end{aligned} \quad (16)$$

where

$$\begin{aligned} \text{delta}_r &= \frac{\min \{l1, w1\}}{2 \cdot w2}, \\ \text{delta}_t &= \frac{2 \cdot \pi}{l2}, \\ \text{center}_x &= \frac{l1}{2}, \\ \text{center}_y &= \frac{w1}{2}. \end{aligned} \quad (17)$$

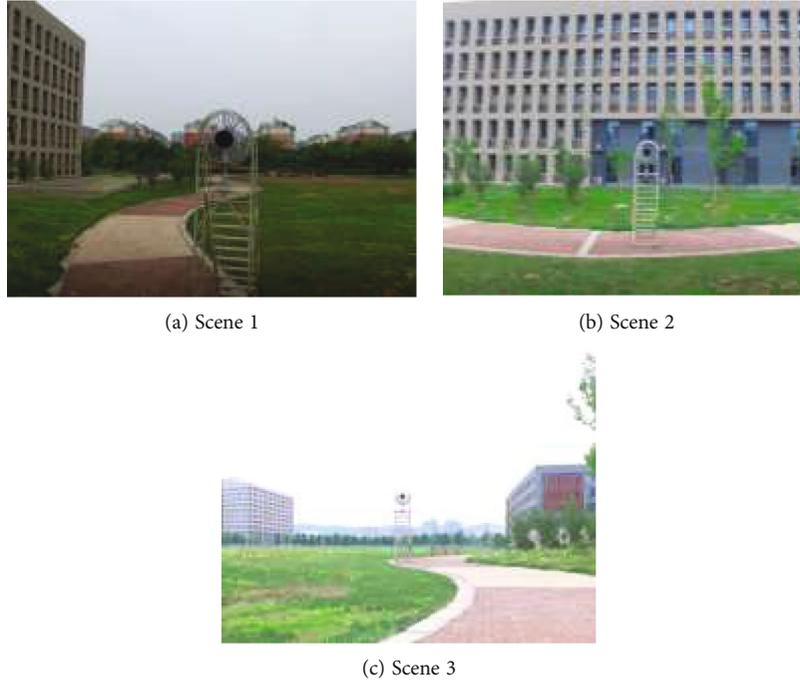


FIGURE 9: Experiment scene diagram from one to three.

TABLE 1: Comparison of performance between various detection algorithms at different distances.

Options			Close	Middle	Far
Sliding window detection	Original image	Success rate	100%	74.3%	0%
		Consuming time (ms)	580.7	578.4	576.2
	1.5 times image	Success rate	100%	97.6%	31.2%
		Consuming time (ms)	1596.4	1588.4	1610.6
	2 times image	Success rate	100%	100%	67.4%
		Consuming time (ms)	2398.2	2410.3	2450.6
Adapted threshold segmentation detection	Original image	Success rate	100%	100%	100%
		Consuming time (ms)	41.6	45.4	47.5

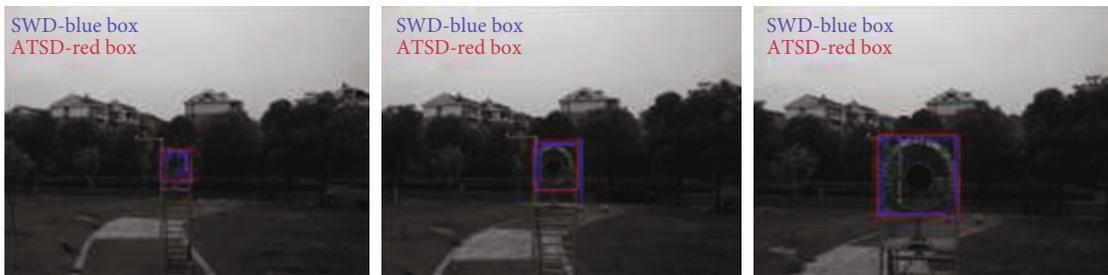


FIGURE 10: Partial detection pictures of various detection algorithms.

Afterwards, the converted image is smoothing processed, horizontal edge detected, threshold segmented, morphological open operated, and edge point extracted, which is as shown in Figure 5. Finally, the edge points of the Cartesian coordinate system are mapped back to the original image for ellipse parameter fitting.

The iterative least squares ellipse method is used to fit all the effective edge points [22]. And then, the accurate target size and center position  $(a, b, x_0, y_0)$  are obtained. With the movement of the camera or drogue, its size and position in the image will change accordingly. When the size of tracking window remains the same, increasing

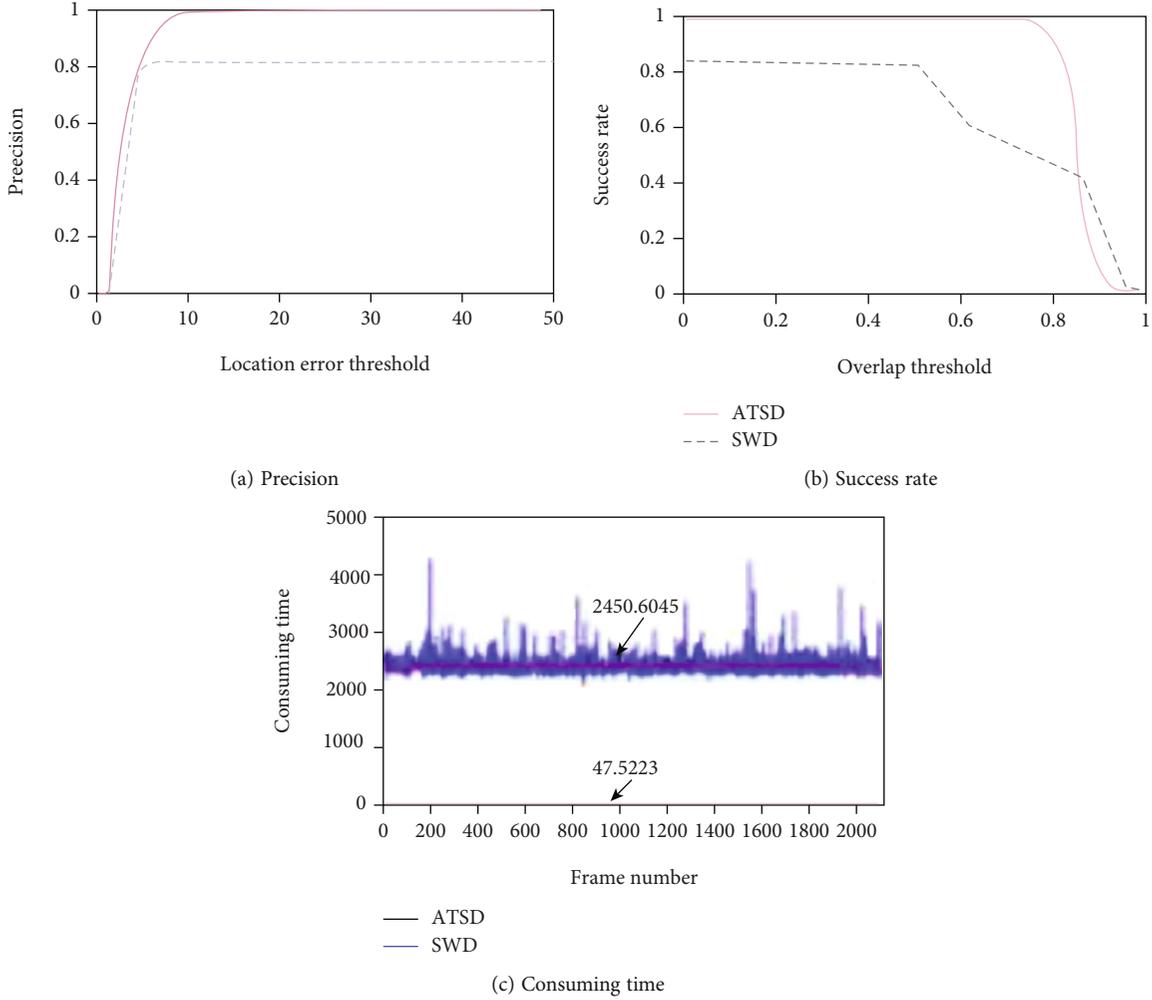


FIGURE 11: The performance of various algorithms in far distance.



FIGURE 12: Partial tracking pictures of various tracking algorithms.

or decreasing the size of the target will cause the tracking model to lose some target features or introduce wrong background features. Meanwhile, due to the change of position, it needs to predict the moving position of the target at the next moment. Because there will always be errors in the predicted position and as the errors gradually accumulate, tracking drift will occur, and it will lead to

track failed. In this paper, the tracking error is corrected by the real-time correction of the inner circle center of the target, and the image resolution is adjusted in real time to remain the size of the drogue target in the image relative to the tracking window unchanged. This method can ensure the correctness of the tracking model and prevent the accumulation of predicting errors.

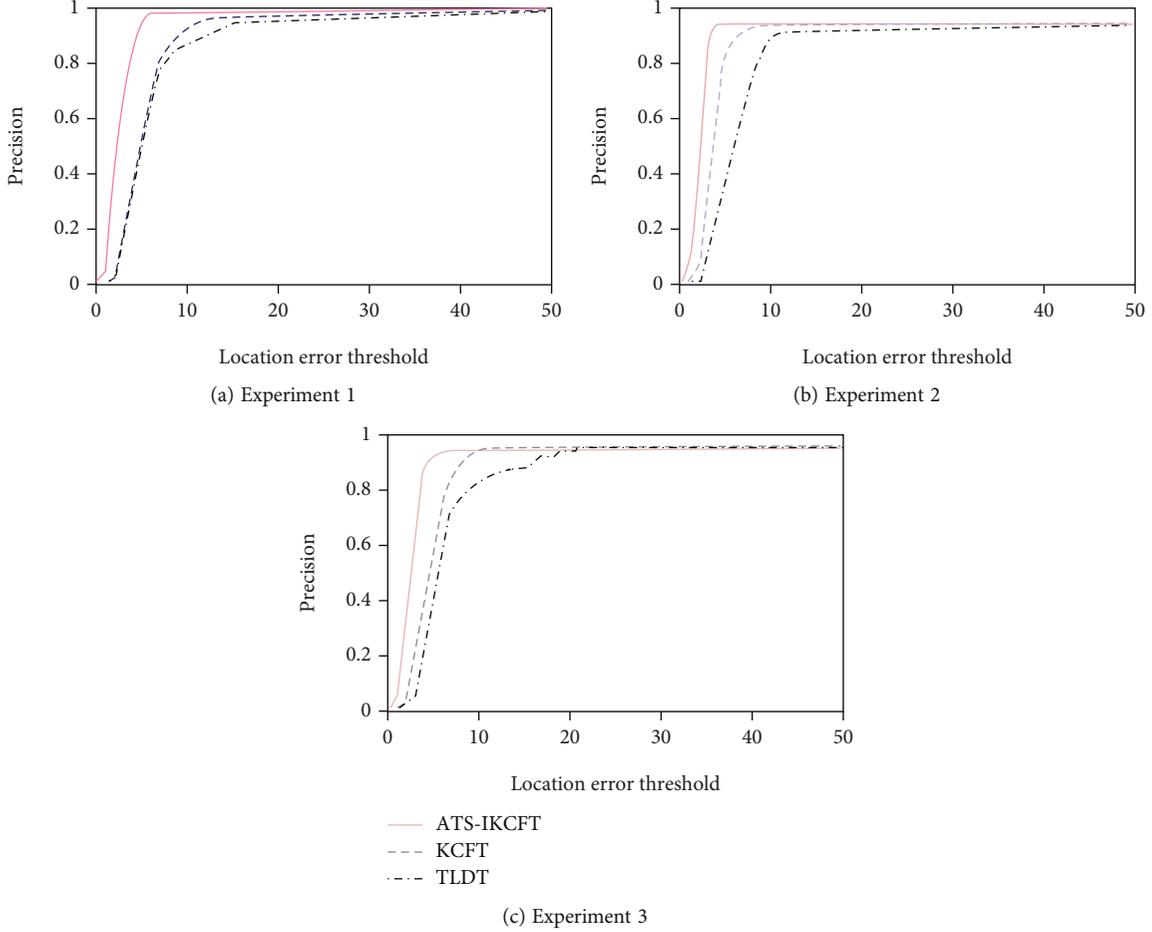


FIGURE 13: The tracking accuracy of various tracking algorithms.

The image resolution is adjusted according to the fitted target size and position information; the adjustment criteria are as follows.

$$\begin{aligned} \text{rate}_x &= \frac{4 \cdot a}{l}, \\ \text{rate}_y &= \frac{4 \cdot b}{w}, \end{aligned} \quad (18)$$

where  $l$  and  $w$  are the length and width of tracking window. After the resolution adjustment above, the new image  $I'$  is obtained, and the center position of prediction window  $W$  is corrected. Here,  $W = (\text{win\_cx}, \text{win\_cy})$ .

$$\begin{aligned} \text{win\_cx} &= x_0 \cdot \text{rate}_x, \\ \text{win\_cy} &= y_0 \cdot \text{rate}_y. \end{aligned} \quad (19)$$

The tracking model is updated by the region  $I'(W)$ , where  $I'(W)$  represents the window  $W$  corresponding to the image  $I'$ . The process diagram of drogue tracking algorithm is shown in Figure 7.

### 3. Results and Discussion

**3.1. Testbed for AAR Docking Task.** The PRPL multistage detection and tracking algorithm of drogue proposed in this paper is tested experimentally with a testbed for AAR docking task shown in Figure 8. There are three parts in the testbed, including drogue, UVA, and ground station. A full-scale refueling drogue model is mounted on a 2D mobile platform with 3.5-meter height. A micro six-rotor aircraft is used as the UVA, and an industry camera is installed on it to capture visual images of drogue motion. The whole system operation is monitored by ground station, which guarantees the controllability and safety of the experimentation.

An industry camera of FL3-U3-20E4M-C is installed on UVA shown in Figure 8, which is able to catch moving images at 110 Hz with resolution of  $1280 \times 720$ . In this paper, the images are processed at resolution of  $800 \times 600$  to reduce image processing time. In addition, the camera is linked to a full-fledged computer with operation system of Ubuntu 16.04 for the visual system running. Moreover, the proposed image processing algorithm is exploited in C++ language, and the image data is managed with OpenCV libraries.

In order to analyze the algorithm performance proposed in this paper, three experimental scenarios are designed to

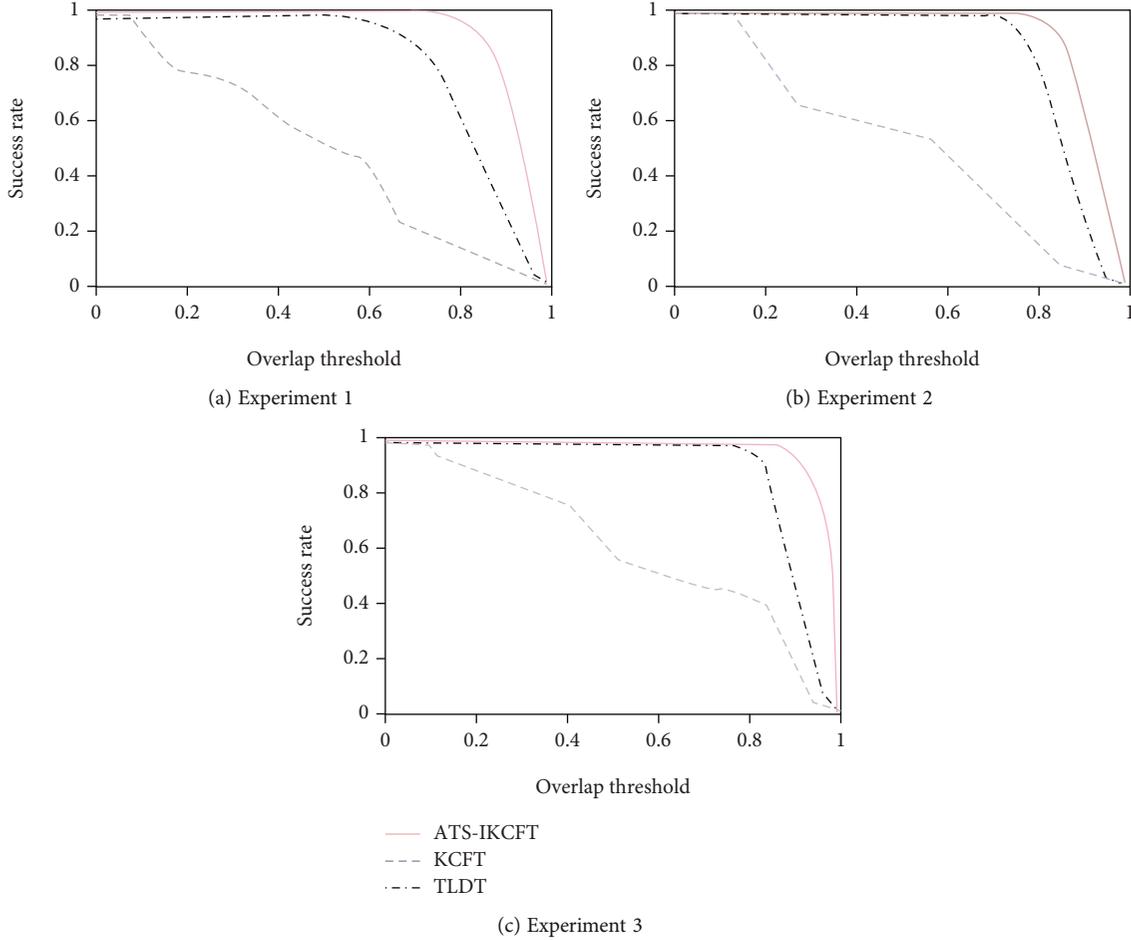


FIGURE 14: The tracking success rate of various tracking algorithms.

test the precision, success rate, and consuming time. As shown in Figure 9, scene 1 is weak light condition with complicated background. Scene 2 is moderate light condition with complicated background. Scene 3 is bright light condition with simple background.

**3.2. Precision, Real-Time Capability, and Reliability of the Drogue Detection Method.** For target detection, the real-time capability and reliability are particularly important. In this paper, precision, success rate, and consuming time of adaptive threshold segmentation detection (ATSD) algorithm are compared and analyzed with sliding window detection (SWD) method. Here, precision refers to the number of frames below the given center error as the percentage of the total frames when the tracking center error changes from 0 to 50 pixels. Success rate refers to the number of frames higher than the given overlap rate as the percentage of the total frames when the overlap rate of the tracking region to calibration region changes from 0 to 1.

The experimental video data is selected in scene one at close (2~5 m), middle (5~10 m), and far distance (10~20 m), respectively. The precision, success rate, and consuming time under different distances are compared in Table 1. Here, the success rate indicator is the value when the overlap

threshold is 0.6. As shown in Table 1, the distance of the target causes a large change in the size of the target in the image, which causes a decrease in the detection success rate. If the resolution of the image is changed to adapt the change of size, the consuming time of target detection is greatly increased, and the real-time performance is decreased. As for adaptive threshold segmentation detection algorithm, the detection success rate does not change significantly when the distance changes. It is highly robust to the target size and consumes less time. Partial detection effect pictures of the two algorithms are shown in Figure 10. The performance of various algorithms in far distance is shown in Figure 11.

**3.3. Precision, Real-Time Capability, and Reliability of the Drogue Tracking Method.** In order to analyze the adaptive threshold segmentation detection and improved KCF tracking (ATSIKCF) algorithm performance proposed in this paper, the precision, success rate, and consuming time are analyzed and compared with the traditional KCF tracking (KCFT) method [28] and tracking learning detection tracking (TLDT) method [29].

Partial tracking pictures of various methods for scene 2 at different distances and light conditions are compared in Figure 12. As shown in Figure 12, the adaptabilities to the

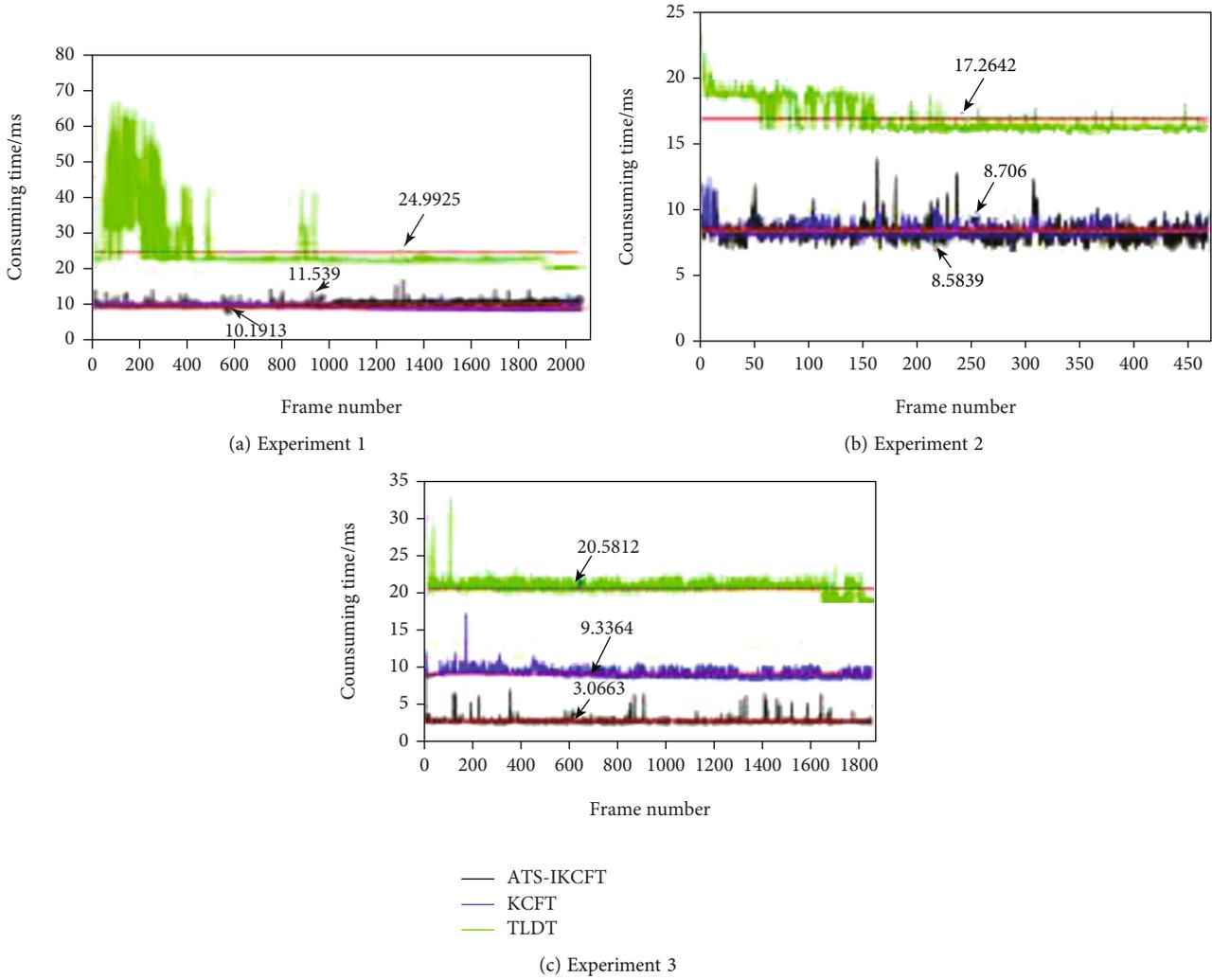


FIGURE 15: The consuming time of various tracking algorithms.

target scale change of the KCFT and TLDT method are poorer than ATS-IKCFT. When the target scale changes, it will bring into background interference information or loss some of the target feature information, which causes the accumulation of tracking errors and the drift of the tracking region. The tracking accuracy and success rate are shown in Figures 13 and 14. It shows that the tracking accuracy and success rate of ATS-IKCFT are higher than those of both KCFT and TLDT method. The ATS-IKCFT algorithm has high robustness and accuracy for different environments.

In terms of real-time tracking, as shown in Figure 15, the ATS-IKCFT tracking algorithm in this paper has reached a high level. The average processing time in various environments is within 11 ms and it reaches 3 ms under better environment, which meets the requirements of high-dynamic and high-frequency output. The consuming time of three experimental results is shown in Figure 15.

#### 4. Conclusions

Drogue detection and target tracking algorithm is significant for probe-drogue refueling system of AAR. In this paper, a

novel PRPL multistage image processing strategy based on vision navigation is proposed for the drogue docking tasks. The detection preprocessor, detection recognizer, target tracking predictor, and target tracking locker are designed and implemented for the detection and tracking of drogue. The preprocessing cascade classifier accelerates the target detection speed and improves the robustness of target detection. The histogram adaptive threshold classifier makes it possible to adjust the appropriate threshold for target segmentation under different lighting conditions. The tracking feedback phase enables the tracker to adapt to the change of the target scale, reduces the accumulation of tracking errors greatly, and ensures the effectiveness of the tracking model. Experiments are conducted with a full-scale drogue model to provide high-fidelity moving images by a micro six-rotor aircraft. Results show that the proposed PRPL multistage image processing strategy for drogue has better accuracy, real-time capability, and reliability than conventional methods, which can meet the requirements for AAR in docking stage.

Future work will focus on the development of relative navigation algorithm between drogue and UVA based on

vision system. It will provide a meaningful reference for the practical application of AAR technology.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

## Acknowledgments

The work described in this paper was supported by Funding of Jiangsu Innovation Program for Graduate Education (Grant No. KYLX15\_0276).

## References

- [1] F. O. MAO, "A survey of the dynamics and control of aircraft during aerial refueling," *Nonlinear Dynamics & Systems Theory*, vol. 4, no. 4, pp. 375–388, 2008.
- [2] L. Wu, S. Yongrong, and H. Bin, "Modeling method and control strategy for hose-drogue aerial refueling system," *Trans. Nanjing Univ. Aero. Astro.*, vol. 35, no. 2, pp. 371–382, 2018.
- [3] Y. Zhu, Y. Sun, W. Zhao, B. Huang, and L. Wu, "Relative navigation for autonomous aerial refueling rendezvous phase," *Optik-International Journal for Light and Electron Optics*, vol. 174, pp. 665–675, 2018.
- [4] M. L. Fravolini, M. Mammarella, G. Campa, M. R. Napolitano, and M. Perhinschi, "Machine vision algorithms for autonomous aerial refueling for UAVs using the USAF refueling boom method," *Studies in Computational Intelligence*, vol. 304, 2010.
- [5] D. B. Wilson, A. H. Gktoan, and S. Sukkarieh, "Drogue motion estimation using air-to-air observations," in *Proceedings of Australasian Conference on Robotics and Automation*, Melbourne, Australia, 2014.
- [6] M. L. Fravolini, A. Ficola, G. Campa, M. R. Napolitano, and B. Seanor, "Modeling and control issues for autonomous aerial refueling for UAVs using a probe-drogue refueling system," *J. of Aerospace Science Technology*, vol. 8, no. 7, pp. 611–618, 2004.
- [7] B. Huang, Y. Sun, and Q. Zeng, "Real-time drogue detection and template tracking strategy for autonomous aerial refueling," *Journal of Real-Time Image Processing*, vol. 17, no. 3, pp. 437–446, 2020.
- [8] K. D. Zhao, Y. R. Sun, Y. Zhang, and H. Li, "Monocular visual position and attitude estimation method of a drogue based on coaxial constraints," *Sensors*, vol. 21, no. 16, p. 5673, 2021.
- [9] S. Li, R. Lu, L. Zhang, and Y. Peng, "Image processing algorithms for deep-space autonomous optical navigation," *Journal of Navigation*, vol. 66, no. 4, pp. 605–623, 2013.
- [10] G. E. Lightsey and J. A. Christian, "An on-board image processing algorithm for a spacecraft optical navigation sensor system," *Journal of Spacecraft & Rockets*, vol. 49, no. 2, pp. 337–352, 2010.
- [11] D. Luo, J. Shao, Y. Xu, and J. Zhang, "Docking navigation method for UAV autonomous aerial refueling," *SCIENCE CHINA Information Sciences*, vol. 62, no. 1, 2019.
- [12] W. R. Williamson, G. J. Glenn, and V. T. Dang, "Sensor fusion applied to autonomous aerial refueling," *Journal of Guidance Control & Dynamics*, vol. 32, no. 1, pp. 262–275, 2009.
- [13] J. Kimmett, J. Valasek, and J. L. Junkins, "Autonomous aerial refueling utilizing a vision based navigation system," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Monterey, California, 2002.
- [14] J. L. Hansen, J. E. Murray, and N. V. Campos, *The NASA Dryden Flight Test Approach to an Aerial Refueling System*-NASA/TM-2005-212859.
- [15] L. Pollini, M. Innocenti, and R. Mati, "Vision algorithms for formation flight and aerial refueling with optimal marker labeling," in *AIAA Modeling and Simulation Technologies Conference and Exhibit*, San Francisco, California, 2005.
- [16] X. F. Wang, X. W. Kong, and J. H. Zhi, "Real-time drogue recognition and 3D locating for UAV autonomous aerial refueling based on monocular machine vision," *Chinese Journal of Aeronautics*, vol. 28, no. 6, pp. 1667–1675, 2015.
- [17] B. Huang, Y. R. Sun, Y. F. Zhu, Z. Xiong, and J. Liu, "Vision pose estimation from planar dual circles in a single image," *Optik*, vol. 127, no. 10, pp. 4275–4280, 2016.
- [18] C. Martínez, T. Richardson, P. Thomas, J. L. du Bois, and P. Campoy, "A vision-based strategy for autonomous aerial refueling tasks," *Robotics and Autonomous Systems*, vol. 61, no. 8, pp. 876–895, 2013.
- [19] Y. Yin, D. Xu, X. Wang, and M. Bai, "Detection and tracking strategies for autonomous aerial refuelling tasks based on monocular vision," *International Journal of Advanced Robotic Systems*, vol. 11, no. 7, p. 97, 2014.
- [20] S. Gao, Y. Cheng, and C. Song, "Drogue detection for vision-based autonomous aerial refueling via low rank and sparse decomposition with multiple features," *Infrared Physics & Technology*, vol. 60, pp. 266–274, 2013.
- [21] Y. Qin, H. Wang, and P. Yao, "Drogue detection and tracking method for monocular-vision-based autonomous aerial refueling," in *Proceedings of 2016 IEEE Chinese Guidance, Navigation and Control Conference*, Nanjing, 2016.
- [22] B. Huang, Y. R. Sun, and B. W. Yang, "Drogue image detecting and tracking based on iterative least squares ellipse fitting," *Journal of Image and Graphics*, vol. 19, no. 8, pp. 1202–1209, 2014.
- [23] X. Xu, H. Duan, Y. Guo, and Y. Deng, "A cascade adaboost and CNN algorithm for drogue detection in UAV autonomous aerial refueling," *Neurocomputing*, vol. 408, pp. 121–134, 2020.
- [24] C.-I. Chen, R. Koseluk, C. Buchanan, A. Duerner, B. Jeppesen, and H. Laux, "Autonomous aerial refueling ground test demonstration—a sensor-in-the-loop, non-tracking method," *Sensors*, vol. 15, no. 5, pp. 10948–10972, 2015.
- [25] H. Nasser, "Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 11, pp. 3592–3607, 2011.
- [26] F. Pedro, "Object detection with discriminatively trained part-based models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.
- [27] S. Hare, S. Golodetz, A. Saffari et al., "Struck: structured output tracking with kernels," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 10, pp. 2096–2109, 2016.

- [28] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 583–596, 2015.
- [29] Z. Kalal, K. Mikolajczyk, and J. Matas, "Tracking-learning-detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1409–1422, 2012.