Design of UAV target tracking controller based on visual servo

Hongfei Wang¹, Yongkang Shi¹*

¹School of Mechanical Engineering, Xinjiang University, Urumqi, Xinjiang Province, 8-30017, China
*Corresponding author’s e-mail: sunwenxj@sina.cn

Abstract. The task of tracking a moving target by the UAV has two major challenges: one is that the multi-angle and multi-scale changes caused by the target’s movement make the target detection more difficult, the other is that the stability and accurate of target tracking are difficult to guarantee. In order to solve these problems, using monocular vision to get the image information of the target and YOLOv4 is used to detect multi-angle and multi-scale moving targets. Then based on the image visual servo control method, the PID controller is designed, and the deviation between the target pixel and the ideal pixel is used as the input, the output of the controller is the desired pixel speed of the UAV in image coordinate system. The tracking speed of the UAV under the machine system is solved according to the image Jacobian matrix, and a tracking strategy is designed to improve the tracking accurate and stability. In the process of target tracking, the difference between the x-axis and y-axis coordinates of the target center point and the image center point is less than 30 pixels. The results show that the proposed algorithm can keep the target in the center of the image.

1. Introduction
With the development of computer vision technology, object tracking becomes more and more popular in the field of computer vision. At the same time, with the development of sensor technology and UAV technology, it is easy to realize target tracking based on UAV. Moreover, the application requirements in civil, industrial, military and other fields also make the realization of the UAV tracking system urgent. Therefore, a large number of scholars have emerged at home and abroad to study the field of UAV target tracking.

Liu Tianyu et al.[1]. applied the machine learning-based KCF target tracking algorithm and YOLOV4 target detector to the airborne target tracking scenario. The results show that the target tracking accuracy is high and the real-time performance is good. Ds-yolo (Dense-SPP Yolo)[2]. is proposed by Zhang Wei et al. The test results on Visdrone2019-DET dataset show that the mAP index of DS-YOLO algorithm is about 3% higher than that of myelov 3 algorithm, the detection speed is 89 FPS. Qiduo et al [3]. optimized YOLO network with k-means clustering algorithm to improve the detection ability of small targets, and designed a realizable PID controller for tracking moving targets. The actual flight test results show that the UAV can accomplish the mission with high dynamic target detection rate and accurate tracking rate. X. Qin et al. [4]. designed the position-attitude outer-loop controller, with the horizontal displacement between the target and the UAV as the control input, realized the effective tracking of the moving target. Xiao Zhengliang et al [5]. designed a tracking strategy (MSA) based on monocular vision, the algorithm effectively prevents the target from being lost in the process of tracking.

E. Çinta et al.[6]. proposed using YOLOv3-tiny combined with KCF to track the target, which improves the tracking speed. M. Rabah et al. [7]. put forward a target tracking algorithm of fuzzy PI
controller, which takes position and position change as input to adjust the parameters of PI controller. The experimental results of target tracking show that the controller can track the moving target very well. L. M. del Pozo López et al. [8] proposed a target tracking method that combines visual measurement information with navigation sensor information, real-time acquisition of the position and velocity of the moving target on the ground. The simulation results show the effectiveness of the proposed method and control algorithm. M. Lee et al. [9] divided the moving vehicle to generate the object’s center of mass, took the object’s center of mass as the tracking input, and used the Kalman filter to estimate the dynamic state of the target for tracking. The experimental results show that the algorithm has a good moving target detection and tracking accuracy.

In this paper, Yolo is used to detect the target, and the center value of target is calculated according to the return value of Yolo. The corresponding relationship between the projection point on the image plane and the plane coordinate system is obtained. A point on the image plane is selected as the target point, and image-based visual servo control is used to track the moving target. In order to prevent the moving object from losing, a tracking strategy based on the idea of image region partition is proposed.

2. Object detection based on Yolo
In this paper, YOLOV4 is selected as the target detector, considering the real-time and accuracy. The YOLO network turns the detection problem into a regression problem, in which the image is used only once as an input network, rather than being processed separately for each class. Using end-to-end detection method, the frame coordinates, target probability and confidence are obtained by detecting all pixels of the whole image.

Yolo uses convolutional neural network for feature extraction, which has a similar structure to Google Net, with two fully connected layers and 24 convolution layers. Convolution layer is used to extract image features, and fully connected layer is used to predict image position and class probability. First, using 3×3 convolution layer and then use 1×1 convolution layer to integrate multi-channel features, the 1×1 convolution layer can reduce the information collected from 3×3 convolution layers, using the way to reduce the image dimensions and simplify the operation.

Like other target detection algorithms based on neural network, YOLO algorithm’s core is data set, and its diversity is the key to optimize network model. A large number of samples with different lighting and different backgrounds were collected and labeled. It can give the position information of the target in the image for regression fitting.

3. Design of UAV target tracking controller based on visual servo
In this paper, a visual servo control algorithm based on image features is used to control the deviation between the actual position and the desired position in the image.

Image-based feedback system is divided into three parts: image acquisition, visual feedback and control strategy. Firstly, the image information is acquired by the camera on the UAV, and then the appropriate feedback signal is input into the control system by the feedback system to realize the motion control. For the image acquisition part, the imaging model of the airborne camera is established, and the corresponding relationship between the target pixel coordinates in the image plane and the 3D scene is obtained. At the same time, according to the imaging principle, the depth information relative to the target and the height of the target can be obtained.

3.1. Imaging principle of airborne camera
Figure 1 shows the imaging principle of a moving target on an onboard camera. \(O\) is the ground coordinate point corresponding to the center \(o\) of the camera image. Establish Descartes space coordinate system with \(O\) as origin, The back, right, and top of the target are in the positive direction of the \(y\), \(x\), and \(z\) axes, respectively. Suppose the target moves from \(O\) to \(P\), The projections of vector \(OP\) on the \(x\) and \(y\) axes are \(OP_x\) and \(OP_y\). The projection point of the pixel plane corresponding to the \(P\) point is \(P\). The projection of the pixel planes corresponding to \(OP_x\) and \(OP_y\) is \(OP_x'\) and \(OP_y'\). In addition,
take the lower left corner of the image as the origin of a Cartesian coordinate system, the right side is positive x axis, the upper side is positive y axis.

For the control of a moving target, a space Cartesian coordinate system is established, taking the position of the drone (point F) as the origin, and the rear, right and upper sides of the drone are in the positive direction of the y, x, and z axes. The focal length of the onboard camera is \( f \), \( d \) is the distance from the drone to the target, \( H \) is the height of the drone, \( 2h \) is the height of the target, \( 2L \) and \( 2W \) are the length and width of the images taken by the onboard camera, \( \theta \) is the elevation of the camera, respectively. Let \( \angle OFP = \theta \), The coordinates of Point P in the coordinate system \((x, y, z)\) are \((x_1, y_1, 0)\), \( P \) in the coordinate system \(x, y\) are \((x_2, y_2)\), thus, \( d \) can be expressed as:

\[
d = \frac{(H - h)}{\sin \theta}
\]

(1)

In fact, the height information of the target can be calculated by using the known information. Figure 2 shows the imaging principle of a target at \( 2h \) on a camera, with \( K \) and \( L \) as the starting and ending points, \( K \) and \( L \) are their projection points on the image plane. To sum up, \( \angle KQO = \pi / 2 - \theta \), \( OF = d \), \( OF = f \), \( KO = LO = h \). Make a point \( Q \) on line \( OO' \), make \( KQ \parallel OO' \), Then, we can get:

\[
KQ = h \cos \theta
\]

(2)

\[
OQ = h \sin \theta
\]

(3)

The equation (1) gives the height of the target \( h \)

\[
h = m(H - \sin \theta)/(m + f \sin \theta \cos \theta)
\]

(5)

3.2. Image-based visual servo control

In this section, the PID controller is designed based on the visual servo control method of image. According to the imaging principle of the airborne camera introduced in the previous section, the coordinates of the target in the image plane are \((x, y)\), The desired coordinates of the target in the image plane are \((x_d, y_d)\), The system generates an error signal \((e_x, e_y)\) based on the deviation between \((x_1, y_1)\) and \((x_2, y_2)\), The controller calculates the control speed \(v_d\) based on the image coordinate system according to the error signal. Then, the speed of UAV is calculated by Image Jacobian matrix, and the control of UAV is realized. \((e_x, e_y)\) may be expressed as:

\[
e_x = x - x_d, \quad e_y = y - y_d
\]

(6)

Mathematical model of PID control:
In this paper, the following control inputs for the UAV target are:

\[
\begin{bmatrix}
    u_x(t) \\
    u_y(t)
\end{bmatrix} = k_p \begin{bmatrix}
    e_{xp}(t) \\
    e_{yp}(t)
\end{bmatrix} + k_i \sum_{j=0}^{n} e_{xp}(j) + k_j \left[ e_{xp}(t) - e_{xp}(t-1) \right]
\]

(8)

where \( e_{xp}, e_{yp} \) is a scalar, and the desired velocity of the UAV in image coordinate can be calculated by the controller:

\[
\begin{bmatrix}
    \ddot{u} \\
    \ddot{v}
\end{bmatrix} = \frac{|x_d - x_i|}{2L} \begin{bmatrix}
    u(t) \\
    v(t)
\end{bmatrix}
\]

(9)

Among them, \(|x_d - x_i|/2L, |y_d - y_i|/2W\) are the ratio of the error to the pixels in the image.

3.3. Target tracking strategy

When the target’s projection point on the image is in a certain position, it is an important problem for UAV to adjust its position correctly. As shown in figure 3, for a drone to track a moving target, we always want the target to be in the center of the image after the drone has adjusted its position for a limited amount of time. Taking into account the optical error caused by the camera’s internal distortion, it is decided to localize the target image captured by the camera and define it as Figure 3.

The whole image region is divided into three parts, S1, S2, S3, L and W are the width and height of the image, and S1 is the circular region with the center of the image as the center and w/6 as the radius. S2 is the area where the center of the image is the center of the circle and W/3 is the radius minus S1. S3 removes the S1 and S2 regions for the entire image area. S1 is called the reliable region, S2 is
called the unreliable region, and S3 is called the extremely unreliable region.

The detailed implementation steps of the algorithm are shown in Table 1.

| Step   | Description                                                                                           |
|--------|-------------------------------------------------------------------------------------------------------|
| Step1  | The destination center is calculated based on the value returned by Yolo                             |
| Step2  | Calculates the Euclidean distance \( \text{dist} \) of the Target Center to the center of the image    |
| Step3  | \( \text{Dist} < \frac{\text{W}}{6} \), hover and wait; \( \frac{\text{W}}{6} < \text{dist} < \frac{\text{W}}{3} \), use the center of the image as the target point; \( \text{dist} > \frac{\text{W}}{3} \), select a point in the S2 area as the target point |
| Step4  | Calculates the desired pixel speed                                                                 |
| Step5  | Using 2-12 Jacobian matrix, the desired tracking velocity of UAV is obtained                         |
| Step6  | Repeat the steps 1~3                                                                                  |

### 4. Experimental Verification and result analysis

Simulation verification platform based on open source UAV simulation platform XTdrone[10]. Among them, a total of 600 images were collected in the simulation environment, but 600 images could not meet the requirements of training Yolo weight file, meet Yolo training requirements. The original dataset is expanded by 5 enhancement algorithms: translation, rotation angle, random color, contrast enhancement and brightness enhancement. Using the enhancement algorithm on the data set can not only solve the problem of insufficient data, but also get the network with stronger generalization ability and better adaptation to the application scenario.

The resolution of the image is set to 640 × 480. After many experiments, target tracking of UAV using PD controller. The PD parameters on the x and y axes are 0.8 and 0.3, respectively.

During target tracking, the coordinates of the center of the target are distributed as shown in figure 4 and figure 5. Throughout the experiment, the drone just hovered and waited for the target to appear in the camera’s field of view, then began to track. During the tracking process, the drone followed the target to move forward, and then to move to the right. The trace is shown in figure 6. Figure 4 and 5 show that the target is above the middle of the camera’s field of view during forward tracking. This is
because the drone doesn’t add pan control, but the target stays in the center of the image in the X axis during the tracking process, on the Y axis, the target is moving towards the center of the image. This shows that the effect of the designed controller is satisfactory.

The three axis velocity information of the UAV is shown in figures 7,8 and 9. The Z axis velocity of the UAV is always near 0, because the simulation experiment does not control the Z axis direction of
the UAV, the UAV always tracks at the set tracking altitude during the tracking process.

In the simulation, the tracking target starts to move toward the road at a steady speed, then turns right at the intersection to change the direction of movement, the tracking target speed is always the same. During the forward tracking, the linear velocity in the X axis of the UAV is always about 1 m/s, and the linear velocity in the y axis is always about 0. This is because the target is always moving in one direction at a constant speed, the drone only needs to change the speed of one direction to track. At frame 1,400, the target changes direction, after which, the drone moves right to track the target, and the linear velocity of the drone’s x axis rapidly drops to zero and remains near zero, the linear velocity on the y axis fluctuates around -0.5 m/s to continue tracking the target.

5. Conclusion
This paper mainly studies how to track multi-angle and multi-scale moving target for single UAV based on vision sensor. Therefore, the precision of target detection is enhanced by Yolov4 network. According to the imaging principle, the relationship between the pixel velocity of the projection point on the image of the moving target and the tracking velocity of the unmanned aerial vehicle (UAV) is studied, a method of measuring the height of a moving object based on its image is proposed. In order to improve the stability of UAV tracking target, a tracking strategy based on the idea of image region partition is proposed. According to the position of the target in the image, the corresponding control strategy is adopted. After the image is acquired on the UAV, the image-based visual servo control method is used to process the image to get the feedback signal and input it to the improved PID controller, which outputs the control instruction (speed), moving target tracking.

Reference
[1] Liu, C. og Wang, C. (2021) Air target tracking simulation based on kcf algorithm. LASER & INFRARED, 51(10): 1396-1400.
[2] Zhang, D. og Wang, C. (2021) Ds-yolo: A real-time small object detection algorithm on uavs. Journal of Nanjing University of Posts and Telecommunications(Natural Science Edition), 41(01): 86-98.
[3] Qi, D., Li, Z., Ren, B., Lei, P. og Yang, X. (2021) Detection and tracking of a moving target for uav based on machine vision. Í 2021 7th International Conference on Control, Automation and Robotics (ICCAR). pp. 173-178.
[4] Qin, X. og Wang, T. (2019) Visual-based tracking and control algorithm design for quadcopter uav. Í 2019 Chinese Control And Decision Conference (CCDC).pp. 3564-3569.
[5] Xiao, Z., Liu, D., Fei, B., Men, T., Zhou, Z., Zhang, X. og Zhou, Y. (2020) Moving target tracking for single uav in open outdoor environment. Í 2020 6th International Conference on Big Data and Information Analytics (BigDIA). pp. 317-324.
[6] Çintaş, E., Özyer, B. og Şimşek, E. (2020) Vision-based moving uav tracking by another uav on low-cost hardware and a new ground control station. IEEE Access, 8: 194601-194611.
[7] Rabah, M., Rohan, A., Mohamed, S. A. S. og Kim, S. H. (2019) Autonomous moving target-tracking for a uav quadcopter based on fuzzy-pi. IEEE Access, 7: 38407-38419.
[8] López, L. M. d. P., Boullosa, M. S. og Corcuera, J. J. N. (2021) Ground target motion estimation based on visual measurements from a uav platform. Í 2021 21st International Radar Symposium (IRS). pp. 1-10.
[9] Lee, M. og Yeom, S. (2018) Tracking of moving vehicles with a uav. Í 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS). pp. 928-931.
[10] Xiao, K., Tan, S., Wang, G., An, X., Wang, X. og Wang, X. (2020) Xtdrone: A customizable multi-rotor uavs simulation platform. Í 2020 4th International Conference on Robotics and Automation Sciences (ICRAS). pp. 55-61.