

# Research on target location of unmanned aerial vehicles in parallel path

*Huadong Luo\**, Jun Yan, Haifeng Ling, Guangna Zhang, Qing Wang, Hongchuan Luo, and Tao Zhu

Graduate School of PLA Army Engineering University, Nanjing Jiangsu, China

**Keywords:** Parallel path search, SSD image detection algorithm, Machine vision, Direction cosine matrix.

**Abstract.** This paper mainly studies how to use the stereo vision system that combines the monocular vision with parallel path search to locate the target. When the unmanned aerial vehicle (UAV) searches in the mission area according to the parallel path, the SSD image detection algorithm based on deep learning is adopted to detect and identify the target in the area. The image coordinate information is inversely calculated by using the pixel coordinate information fed back by machine vision. The auxiliary coordinate system is established according to the relationship of angle position between the track line and the basic coordinate system in the parallel path. Combining the position relation and the attitude direction information of UAV, the target position conversion relation between the imaging coordinate system and the auxiliary coordinate system is solved by using the direction cosine matrix. Combined with the coordinate information of UAV, the coordinate position of the target point in the basic coordinate system is finally solved through three coordinate conversion operations. In order to avoid the single calculating error of the target coordinates, the weighted average operation is carried out. On the basis of not changing the search trip of the parallel path, the target location function is preliminarily realized through the reverse solution and the weighted average operation of the target coordinates.

## 1 Introduction

At present, the application scopes and fields of target location are extensive, and machine vision is widely used to implement target location in all walks of life [1][2][4]. Compared with multi-vision, monocular vision has less research on static target location. There are many limitations in target location using monocular vision, but monocular vision also has its outstanding advantages, such as small computation, fast speed and economical benefit. Monocular vision can locate static targets by stereo vision algorithm. The evaluation of the machine vision target location algorithm is mainly reflected in three aspects: stability, location accuracy and location time-consuming. In the process of using stereo vision

---

\* Corresponding author: [lhdzgn@163.com](mailto:lhdzgn@163.com)

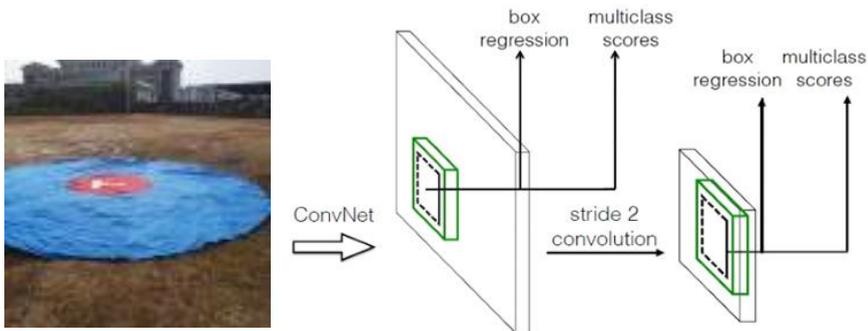
algorithm to implement location, it is easy to be affected by noise interference, illumination change and other factors, which will have a negative impact on system stability. At the same time, a large number of operations and time consumption cannot meet the real-time location requirements. Therefore, this paper employs a new way to carry out the target location by using monocular vision to construct the stereo vision system, and uses the stereo vision system combines the monocular vision with parallel path search to locate the target.

## 2 SSD image detection algorithm based on deep learning

### 2.1 Platform selection

SSD image detection algorithm based on deep learning is a one-stage method [4][5]. The design idea of this method is to conduct intensive sampling at different scales and different locations on the target image according to certain characteristics, and then use CNN algorithm for classification and regression [5][6][7]. For the reason that this method is a one-stage method, it has the advantage of faster speed. However, it is relatively difficult to conduct uniform and dense sampling on the images because of the uneven distribution of the samples and the backgrounds.

SSD algorithm is used to detect and identify the target, and a multi-scale feature graph of CNN network is used for detection [6][8]. Its basic architecture is shown in Fig.1.



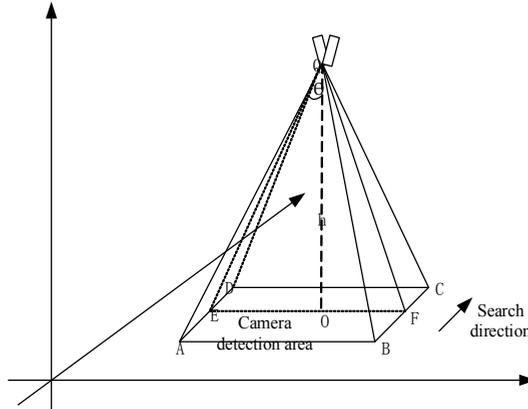
**Fig. 1.** The basic architecture of SSD algorithm.

To make the data set based on the target style, and add labels to the training images employing labelImg tool, then the four files under Main folder were generated by code, which are named as test.txt, train.txt, val.txt and trainval.txt, respectively. The implementation of SSD on TensorFlow is based on the TensorFlow version of balancap (<https://github.com/balancap/SSD-Tensorflow>), which is used to implement the inference process of SSD method.

### 2.2 Pixel coordinate conversion

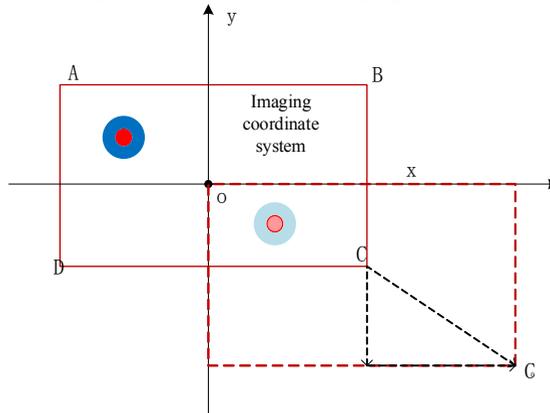
This case implements SSD300. There are actual target and phantom target as output values. In this paper, we ignore the latter phantom target and mainly discuss the positioning of the former actual target in the parallel path in detail. For the detected actual target, we can obtain the center point pixel coordinates of the target, as well as a height value and a width value, both the two values are represented by pixel coordinates.

In this paper, we assume that the flight attitude has no impact on the imaging, and the camera is connected to the unmanned aerial vehicle (UAV) through the cradle, which is vertically downward to search in a parallel path.



**Fig. 2.** A diagram of imaging during the search.

The target pixel coordinates are extracted and converted to imaging coordinates. The camera imaging coordinates are established with the center point of the image as the origin of coordinates, but the origin of coordinates constructed during the target detection is the upper-left endpoint of the target image. There is a displacement difference between the origin of pixel coordinates and the origin of UAV camera imaging coordinates.



**Fig. 3.** Imaging coordinate and pixel coordinate transformation diagram.

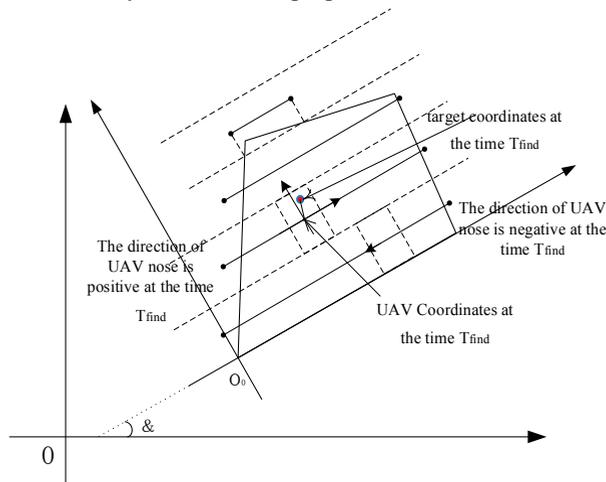
That is,  $x_{\text{imaging coordinates}} = x_{\text{pixel coordinates}} + k_x$ , and  $y_{\text{imaging coordinates}} = y_{\text{pixel coordinates}} + k_y$ . Where in, pixel coordinates can be simply understood as imaging coordinates. During target detection, the images will be compressed and restored, and then represented according to pixel coordinates.  $k_x$  is half of the image length of the camera, and  $k_y$  is half of the image width of the camera.

### 2.3 UAV direction determination

In the process of parallel path search, the camera collects data information, and the onboard computer detects the target by running the SSD target detection program. When the target is found, the flight parameters of the UAV are first taken to obtain the real-time coordinate information of the UAV. The UAV carries out the search according to the parallel path.

During the search, the flight directions of the UAV are two relatively fixed directions, that is, the direction along the parallel path tends to the positive  $X$ -axis and the direction along the parallel path tends to the negative  $X$ -axis. Here, we specify that the direction along the parallel path to the positive  $X$ -axis is positive, and the direction along the parallel path to the negative  $X$ -axis is negative. When the parallel path is perpendicular to the  $X$ -axis, there will be no incremental change along the  $X$ -axis, then the direction tends to the positive  $Y$ -axis will be positive, and the direction tends to the negative  $Y$ -axis will be negative.

Acquired the last two coordinates  $T_1(x_{T_1}, y_{T_1})$  and  $T_2(x_{T_2}, y_{T_2})$  in the UAV flight parameters, where  $T_2$  is the later moment. Let  $k_0 = y_{T_2} - y_{T_1}$ ,  $k_1 = x_{T_2} - x_{T_1}$ . When  $k_0$  or  $k_1$  is positive, the direction of UAV nose is positive; when  $k_0$  or  $k_1$  is negative, the direction of UAV nose is negative. Let the longitude and latitude coordinate values of the UAV at the time  $T_{find}$  to be  $T_{find}(x_{T_{find}}, y_{T_{find}})$ , set the longitude and latitude coordinates as the basic coordinate system. Besides, the coordinate system constructed with the bottom edge of the parallel path as the cartesian  $x$ -axis as the search coordinate system, and the coordinate system constructed on the principle of machine recognition imaging is the imaging coordinate system. The origin of the imaging coordinate system is the longitude and latitude coordinate point of the UAV in the basic coordinate system. In the implementation of parallel path search, the path line of UAV is parallel and the included angle with the positive direction of  $X$ -axis in the basic coordinate system is the fixed value  $\&$ . However, there are special conditions that the direction of UAV nose has binary values and is perpendicular to the  $Y$ -axis.



**Fig. 4.** UAV directional position diagram at the time  $T_{find}$ .

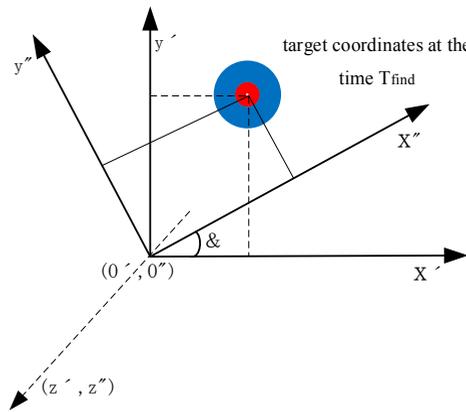
## 2.4 Coordinate conversion

The target location solution at  $T_{find}$  mainly includes the conversion from pixel coordinates to imaging coordinates, from imaging coordinates to auxiliary coordinates, and from auxiliary coordinates to basic coordinates. First, we set up the auxiliary coordinate system, taking the coordinate point of the UAV at  $T_{find}$  as the origin of coordinates, and the axes parallel to the

$X$ -axis and  $Y$ -axis as  $x'$ -axis and  $y'$ -axis. There is a translation relationship between the auxiliary coordinate system  $O'x'y'$  and the basic coordinate system, that is

$$\begin{cases} x = x' + a \\ y = y' + b \end{cases} \quad (1)$$

where  $a$  and  $b$  are the two coordinate values  $x_{T_{find}}$  and  $y_{T_{find}}$  of the points  $T_{find}(x_{T_{find}}, y_{T_{find}})$ . When the direction of the UAV nose is positive, the auxiliary coordinate system is rotated  $\&$  degrees clockwise around the axis  $(z', z'')$  to obtain the imaging coordinate system  $O''x''y''$ .



**Fig. 5.** Imaging coordinates and auxiliary coordinates at the time  $T_{find}$ .

Construct auxiliary coordinate system  $i$  and the imaging coordinate system  $j$ , so that the auxiliary coordinate system and the imaging coordinate system are the right coordinate systems  $O'x'y'$  and  $O''x''y''$  with the same origin. Assume that there is a target point P, whose coordinates in the auxiliary coordinate system and the imaging coordinate system are respectively

$$[x', y', z']^T \quad (2)$$

$$[x'', y'', z'']^T \quad (3)$$

The coordinate transformation relation of the target point P from the imaging coordinate system to the auxiliary coordinate system is essentially the vector sum that projects the components of the imaging coordinate system to the auxiliary coordinate system

$$\begin{cases} x' = x'' \cos(x', x'') + y'' \cos(x', y'') + z'' \cos(x', z'') \\ y' = x'' \cos(y', x'') + y'' \cos(y', y'') + z'' \cos(y', z'') \\ z' = x'' \cos(z', x'') + y'' \cos(z', y'') + z'' \cos(z', z'') \end{cases} \quad (4)$$

The formula can be simplified as

$$[r]_i = [{}^iR_j] \cdot [r]_j \quad (5)$$

Among them

$$[{}^iR_j] = \begin{cases} \cos(x', x'') & \cos(x', y'') & \cos(x', z'') \\ \cos(y', x'') & \cos(y', y'') & \cos(y', z'') \\ \cos(z', x'') & \cos(z', y'') & \cos(z', z'') \end{cases} \quad (6)$$

The above equation is the rotation matrix in the general form of the transformation from the imaging coordinate system  $j$  to the auxiliary coordinate system  $i$ . For the auxiliary coordinate system  $i$ ,  $[{}^iR_j]$  describes the attitude of the imaging coordinate system, which is mathematically called the attitude matrix. All the elements in the attitude matrix are represented by the cosines between the coordinate axes, so the attitude matrix is also called the direction cosine matrix.

When the direction of the UAV nose is positive, that is, when  $k_0 > 0$  or  $k_0=0$  and  $k_1 > 0$ , the UAV imaging coordinate system is obtained through the auxiliary coordinate system rotating the fixed angle  $\&$  around the  $Z$ -axis, that is

$$[{}^iR_j(Z, \&)] = \begin{bmatrix} \cos\& & -\sin\& & 0 \\ \sin\& & \cos\& & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

which can be simplified as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\& & -\sin\& \\ \sin\& & \cos\& \end{bmatrix} \begin{bmatrix} x'' \\ y'' \end{bmatrix} \quad (8)$$

That is, there is a coordinate transformation relationship between the imaging coordinate system and the basic coordinate system

$$\begin{cases} x = x'' \cos\& - y'' \sin\& + a \\ y = x'' \sin\& + y'' \cos\& + b \end{cases} \quad (9)$$

When the direction of UAV nose is negative, that is, when  $k_0 < 0$  or  $k_0=0$  and  $k_1 < 0$ , the UAV imaging coordinate system should be regarded as the auxiliary coordinate system rotating fixed angle  $(\pi + \&)$  around the  $Z$ -axis, which can be obtained as follows

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} -\cos\& & \sin\& \\ -\sin\& & -\cos\& \end{bmatrix} \begin{bmatrix} x'' \\ y'' \end{bmatrix} \quad (10)$$

When  $k_0=0$ , there is no rotation between the imaging coordinate system and the auxiliary coordinate system. When  $k_0=0$ , the angle between the imaging coordinate system and the auxiliary coordinate system is 180 degrees.

## 2.5 Weighted average of the target position

The UAV carries out target search according to the parallel path planning. When the target is found, the coordinate value of the target point can be preliminarily obtained through the calculation. During the flight, the weather situation will have a certain influence on the flight attitude. In order to solve the influence of UAV flight attitude on target positioning accuracy, we employ the weighted average of target position in this paper. In the process of detecting the target, the UAV will continuously detect the target status of each frame image and calculate the target location. Then the calculated results are compared with the existing target information in the stack and weighted average is made after the same decision condition is satisfied, that is  $Q \leq \frac{r}{2} + m$ , where  $r$  is the target radius,  $m$  is the average of the target positioning error, and determine coefficient can be modified according to weather situation.

After it is determined as the same target, the weighted average of the target position coordinates is carried out. The specific weighted average method is to assign a weight value to each position information, and the weight value of single target position information is  $x_0$ . When it is determined to be the same target, whenever the position information is weighted, the position information of the target in the stack is assumed to be  $t$ , the weight value is  $x$ , the newly received position information is  $t_0$ , and the weight value is  $x_0$ . Then the new position information  $t_{new}$  is represented as

$$t_{new} = \frac{t \cdot x + t_0 \cdot x_0}{x + x_0} \quad (11)$$

After the stack is updated, the new position information is  $t_{new}$ , and the weight value of the new position information is  $x_{new} = x + x_0$ . The weight value of the position information in the stack increments to the weight value of the newly received position information. As the target detection program runs, the position information in the stack will become more accurate by processing the images of more frames.

## 3 Summary and outlook

This paper completes the process of UAV target detection and recognition as well as target location and processing. By constructing the training data set, SSD in-depth learning algorithm is used to process the image information captured by the camera, obtain the target pixel coordinates and convert them into imaging coordinate information. Combine the location information of UAV transmitted by flight control with the location relationship between the path location of parallel path search and the basic axis, and use the direction cosine matrix to implement coordinate transformation, the target location is preliminary achieved. Besides, single calculating error is eliminated through the weighted average of the target location, and the coordinate position of the target is more accurate.

## References

1. Qun Z, Dandan W, Shikai S, Boyuan Z, Yu H. Research Status and Development of Multi UAV Coordinated Formation Flight Control. Journal of Harbin Institute of Technology, vol. 49, no. 3, pp. 1-4, 2017.

2. Jian N. The Research and Development of Machine Vision Localization Technology Based on Pattern Shape Feature. Jiangnan University, 2013.
3. Chunping Yin, Liaoni W, Hao W. Research on Key Technology of Wheel Location Based on Machine Vision. Chinese Journal of Automotive Engineering, vol. 3, pp. 157-161, 2016.
4. Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot Multi-Box Detector. Computer Science, vol. 18, no. 2, pp. 55-70, 2015.
5. Lim T. R, Guntoro A. T. Car Recognition Using Gabor Filter Feature Extraction. Proceedings of Asia-Pacific Conference on Circuits and Systems, vol. 20, no. 3, pp. 13-30, 2002.
6. Micha Drozd, Tomasz Kryjak. FPGA Implementation of Multiscale Face Detection Using HOG Features and SVM Classified. Image Processing & Communications, vol. 21, no. 6, pp. 27-44, 2016.
7. Jaderberg M, Simonyan K, Vedaldi A, et al. Reading Text in The Wild with Convolutional Neural Networks. International Journal of Computer Vision, vol. 1, pp. 1-20, 2016.
8. Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-time Object Detection with Region Proposal Networks. 29th Annual Conference on Neural Information Processing Systems, NIPS 2015, Montreal, Canada, pp. 18-30, 2015.