Multiple target recognition of UAV based on image processing

Weidong Deng¹, Daquan Tang, Keke Lu and Guanzheng Tang
Naval Aviation University, Yantai 264001, China

¹E-mail: 276541889@qq.com

Abstract. Unmanned aerial vehicle (UAV) has been widely used in military and civil fields due to its unique advantages. Aiming at the principle of UAV in reconnaissance field and target recognition field, a method of UAV multi-target recognition based on image processing is proposed, and the validity of the method is verified by building an UAV platform and field experiments. At the same time, the display interface of image processing results is developed, which can get the UAV recognition results in real time. Finally, the experimental results are analyzed, and the factors affecting the accuracy of UAV target recognition are elaborated.

1. Introduction
In recent years, with the continuous advancement of science and technology, unmanned technology has gradually become a research hotspot. Due to its small size, strong mobility and low radar radiation surface, the drone has been widely used in the battlefield. At the same time, with the continuous development of the Internet, drones have developed rapidly in express delivery and traffic monitoring.

On the battlefield, when the drone performs tasks such as reconnaissance and strike, because the battlefield targets are many and the environment is complex, the drone must accurately detect the target position and distance. In life, the traffic inspection drone is shooting traffic conditions. When recording a violating vehicle, it must identify and detect the passing vehicle.

In order to better solve the above military and civilian problems, this paper takes the UAV system to reconnoitre and identify the ground fixed target as the application background, and focuses on verifying the real-time recognition ability of the UAV target, and based on the accurate identification of the ground target, the distance of each target from the center-point.

2. Research status
According to the similar characteristics of enemy drones on the battlefield, the algorithm based on sparse auto-encoder is used to fuse the underlying visual features to attack enemy goals, effectively improve the accuracy and robustness of recognition [1]. The literature [2] proposes a target recognition algorithm for laser and infrared fusion. The simulation results show that the result is very good. Aiming at the problem that the recognition accuracy is low due to the inconspicuous object characteristics and the different recognition angles in the complex background, the improved target recognition algorithm based on template matching is proposed in [3]. The algorithm uses the dominant gradient direction with large amplitude as the feature quantity. At the same time, the DOT algorithm is used to remove the secondary gradient features, and the affine projection transformation algorithm is integrated to improve the recognition accuracy of multi-target and multi-view.
3. Experimental procedure

3.1. Image acquisition and enhancement
The process of acquiring images by the drone can be simply divided into three sections: from take-off to reaching the target area, during the reconnaissance of the target area, and from the target area to the take-off point. In these three processes, due to the unstable flight state of the drone, factors such as the interference of the ground debris may affect the processing of the image, resulting in a decrease in the recognition accuracy. Therefore, the image is processed correspondingly before the image processing to make it more suitable for processing than the original image. This operation mode is image enhancement.

Image enhancement is performed on the basis of gradation transformation, and gradation transformation is performed in the spatial domain of the image [4]. The spatial domain of an image refers to the image plane itself, and the processing in the spatial domain is to directly manipulate the pixels in the image [5].

Image enhancement is a related operation performed before image processing, which enhances the features required in the image, suppresses other unwanted image features, and achieves the purpose of improving image quality and enhancing recognition effect [6]. The image enhancement method used in this paper is: image enhancement based on Laplacian.

Image enhancement based on Laplacian. In the acquired image, due to weather, illumination and other factors, there will be blurred regions in the image, and even if there is a grayscale change, the grayscale feedback in the image is not obvious. In order to highlight grayscale jumps, people enter the field using sharpening spatial filters. The main purpose of sharpening is to highlight the transition portion of grayscale [7]. The Laplacian is an isotropic differential operator whose filter is an isotropic filter whose response is independent of the direction of the mutation of the filter image [8].

The Laplacian operator of a two-dimensional image function \( f(x, y) \) is defined as:

\[
\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}
\]

(1)

To be more suitable for digital image processing, change it to a discrete form:

In the \( x \) direction:

\[
\frac{\partial^2 f}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y)
\]

(2)

In the \( y \) direction:

\[
\frac{\partial^2 f}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y)
\]

(3)

The discrete Laplacian operators that satisfy the two variables of these three formulas are:

\[
\nabla^2 f(x, y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)
\]

(4)

The Laplacian is characterized by restoring the background characteristics of the image and maintaining the result of the sharpening process [9], but it is important to note that the Laplacian not only grays out during image sharpening. Sharpening of the mutated region also sharpens the noise, so it is generally necessary to smooth the image when sharpening using the Laplacian operator [10]. The transformation effect is shown in Figure 1 and Figure 2.
3.2. **Optimal global domain value processing based on Otsu method**

After performing grayscale transformation and image enhancement on the image, the image is optimally binarized. The fundamental principle of binarization is to divide the image pixel into two pixel values 0 and 255, and set a field value in the pixel space. When the actual pixel is higher than the set field value, it is divided into 255, when it is lower than the set field value, it is marked as 0 [11]. The purpose of thresholding is to minimize the average error introduced in the process of assigning pixels to two or more classifications [12].

Otsu is a global threshold processing method, also known as the maximum inter-class difference method [13]. It is widely used because it is simple to calculate and is not affected by image brightness and contrast. The principle of Otsu algorithm is to divide the content of the image into two parts according to the gradation characteristics of the image. When the variance between the two is large, it means that the two parts of the image are different, and the recognition process will be the probability of both identifying errors is small [14].

The calculation of the Otsu algorithm is performed directly on the histogram of the image.

Let L denote the different gray levels in a digital image of size pixels, the number of pixels representing the gray level, and the total number of pixels in the image is:

$$MN = n_0 + n_1 + n_2 + \cdots + n_{L-1}$$  \hspace{1cm} (5)

Normalized histogram has components

$$p_i = \frac{n_i}{MN}.$$  \hspace{1cm} (6)

Select a domain value $T(k) = k, 0 < k < L-1$, after selecting the threshold, the gray value of the image is divided into $C_1[0,k]$ and $C_2[k+1,L-1]$. Then the probability that the pixel is assigned to $[0,k]$ is:

$$P_1(k) = \sum_0^k p_i \hspace{1cm} (6)$$

The average gray value is:

$$m_1(k) = \sum_0^k iP(i/C1) = \sum_0^k iP(C1/i)P(i)/P(C1) = \frac{1}{P_1(k)} \sum_0^k ip_i \hspace{1cm} (7)$$

The probability that a pixel is assigned to $[k+1,L-1]$ is:

$$P_2(k) = \sum_{k+1}^{L-1} p_i = 1 - P_1(k) \hspace{1cm} (8)$$

The average gray value is:
\[ m_2(k) = \frac{1}{P_z(k)} \sum_{i=k+1}^{l-1} i P(C2 / i) P(i) / P(C2) \]  \tag{9}

The global gray mean is:
\[ m_G = \sum_{i=0}^{l-1} i p_i \]  \tag{10}

The gray scale variance of all pixels in the image is:
\[ \sigma_G^2 = \sum_{i=0}^{l-1} (i - m_G)^2 p_i \]  \tag{11}

The variance between classes is:
\[ \sigma_B^2(k) = \frac{(m_G P_i(k) - m(k))^2}{P_i(k)(1 - P_i(k))^2} \]  \tag{12}

Maximize \( \sigma_B^2(k) \) as:
\[ \sigma_B^2(k^*) = \max_{0 \leq k \leq L-1} \sigma_B^2(k) \]  \tag{13}

Get the best threshold \( k^* \).

3.3. The Edge detection

Image processing is mainly to extract the required information from the image. Generally speaking, the image contains many contents in the perspective of the lens, which is useless information. In the experiment, the main information needs to be extracted from the whole image. Edge detection technology is needed.

The so-called edge refers to the dividing line of different regions. In the field of image processing, it mainly refers to the method of image segmentation using gray-scale mutation [15], as shown in Figure 3. Common edge detection operators include first-order differential operators and second-order differential operators. The Roberts operator belongs to the first-order differential operator and is the earliest edge detection operator using two-dimensional templates for edge detection [16]. The Sobel operator is performed on the 3*3 region and belongs to the first-order differential operator, as shown in Figure 4. Its advantage is that it can suppress noise better [17].

| -1 | 0 |
|----|---|
| 0  | -1|

Figure 3. Roberts 2*2 template.

| -1 | -2 | -1 |
|----|----|----|
| 0  | 0  | 0  |
| 1  | 2  | 1  |

Figure 4. Sobel 3*3 template.

This article uses the LOG filter. The LOG filter is a combination of Gaussian (GAUSS) filtering and Laplacian filtering, that is, smooth filtering is first performed with a Gaussian filter to over-modulate the noise and extract the edges, so the effect is better [18]. The principle is to perform Gaussian filtering first to perform Laplacian detection, and then to determine the edge position by zero point [19].

Its two-dimensional Gaussian formula is:
\[ G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \]  \tag{14}

According to the Laplacian formula, the second-order partial derivative in the \( x,y \) direction is:
\[
LOG(x, y) = -\frac{1}{\pi \sigma^2} \left[ 1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2 + y^2}{2\sigma^2}}
\]

Here \(x, y\) is the distance from the other position of the template to the center position:

\[
LOG(x, y) = -\frac{1}{\pi \sigma^2} \left[ 1 - \frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2} \right] e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2}}
\]

### 3.4. Experimental simulation and analysis

In order to carry out better experimental verification, the selected task area is a square of 30 meters multiply 30 meters, which is divided into 9 small squares of the same area to represent different task areas, wherein the middle square is set as the center circle. Shape, in each small square, place squares (1m-2m), triangles (side length 1m-2m) and circles (diameter 1m-2m) with different sizes and sizes, and set the target color to red for detection and test identification, the experimental site is shown in Figure 5.

The drone adopts a carbon fiber body. In order to ensure the safety of the drone, the wing is located below the beam bar, and the support bar is internally tested. The 3DR digital transmission with the frequency band of 915MHZ is mounted above, and the FPV radio telemetry air-to-ground module is used to improve the unmanned accuracy of the positioning of the machine identification is equipped with HER+ with carrier phase difference technology see Figure 6 for details. It is equipped with GoPro motion camera to improve the accuracy of UAV recognition. At the same time, since the UAV autonomously takes off and land during the execution of the mission, the original flight control on the basis of the program, changes were made in the aspects of landing judgment, speed information, posture information, position information, and the like.

The principle of the algorithm used in this paper is to use the polygon to approximate a contour when the shape is analyzed, so that the number of vertices is reduced, the algorithm principle is relatively simple, and the core is to continuously find the farthest point of the polygon to form a new polygon until the shortest distance is less than specified accuracy.

By setting up flight routes for the drones, the drones can achieve autonomous take-off and landing, autonomous observation, and the data is transmitted back to the ground display interface.

Compared with the proposed method, the support vector machine (SVM) algorithm is widely used in the field of pattern recognition, but SVM is only excellent in the classification problem of the two categories.

Different performance, from the results, it can be found that its circular recognition accuracy is high, the triangle recognition accuracy is 25%, and the square recognition accuracy is 50%, the experimental results are shown in Figure 7.
The polygon approximation algorithm used this time has a triangle recognition accuracy of 100%, a square recognition accuracy of 100%, and a circular recognition accuracy of 50%, the experimental results are shown in Figure 8.

Figure 7. Support vector machine test result.  Figure 8. Polygonal approximation method.

The factors that influence the experimental results are:
1) GoPro motion camera resolution and graphics transmission and anti-interference ability. In the process of recognizing the target, the ground target is photographed by the camera. Since the drone is photographed in the air and the distance is set to 21 meters to 25 meters, the resolution of the camera has a great influence on the quality of the image;
2) High impact. As we all know, the distance between the distance and the clarity of the graphic photos has a direct impact. The closer the distance, the more accurate the measurement of the side length (diameter) of the graphic, and the recognition accuracy increases accordingly. Considering the flight performance of the drone itself and the flight conditions on the actual battlefield, the height of the drone in this experiment is set at 20 meters, which is more practical.
3) Other distracting items and irresistible factors. During the flight of the drone, the flight attitude is stable, and the accuracy of the flight altitude is very important for target recognition; in addition, the drone is equipped with stable operation of each system, signal interference at the scene, and light intensity, for no one. The recognition result of the machine has a great influence.

4. Conclusions
This paper mainly verifies the multi-target recognition method of UAV based on image processing. By constructing the UAV system, we can improve recognition parameters in the recognition algorithm program and detect the target recognition accuracy of the UAV. At the same time, develop the display window of the ground station to make the identification data processed in real time to obtain the nature of the target, and the distance parameter of each target from the center point is displayed.

This experiment takes the UAV in the battlefield reconnaissance and the application in the civil field as the research background, but the target is relatively simple in the experiment process, and the background of the site has no noise interference, so it is not universal, but for further realization of the drone. Multi-target recognition provides research directions and ideas, and then continues to explore according to the problems of this experiment, making it more adaptable.
References
[1] Bing X, Duan Z 2018 UAV target recognition algorithm based on fusion of SAE and bottom visual feature Infrared & Laser Engineering
[2] Tong X Y, Wu R and Yang X F 2018 Infrared and laser fusion target recognition method Infrared and Laser Engineering 283(5) 167-174
[3] Ding Y L, Zhao Q and Li Y B 2018 Improved target recognition algorithm based on template matching Journal of Shandong University (Engineering Science) 228(02) 5-11
[4] Zhang X H and Yang K T 2007 Image enhancement based on maximum entropy principle and grayscale transformation Opto-Electronic Engineering 34(2) 84-87
[5] Liu D Q 2012 Application of Histogram Processing in Image Spatial Domain Enhancement Processing Journal of Jianghan University(Social Science) 40(5) 45-48
[6] Wang M 2011 Research on Target Feature Recognition Algorithm Based on Image Processing Harbin Engineering University
[7] Jiao Y B 2011 Implementation of a Clearing Method for Grayscale Blurred Image Asia-Pacific Conference on Information Network and Digital Content Security(APCID)
[8] Tian S Y, Wang X M and Zhao X Q 2012 Image Denoising Algorithm Based on Laplacian Operator and Image Patching Journal of Computer Applications 32(10) 2793-2797
[9] Wu H B and Liu Z 2008 Color image sharpening based on Laplacian Computer Development & Applications 21(9) 27-28
[10] Rosenfeld A, Kak A C 1982 Digital picture processing - Volume 1, Volume 2 Computer Science & Applied Mathematics New York Academic Press Ed 6(2) 113-116
[11] Valizadeh M, Armanfard N, Komeili M, et al. 2009 A novel hybrid algorithm for binarization of badly illuminated document images Computer Conference CSICC
[12] Liu D J 2009 Research on Image Segmentation Algorithm Based on Threshold Beijing jiao tong university
[13] Otsu N 2007 A Threshold Selection Method from Gray-Level Histograms IEEE Transactions on Systems Man & Cybernetic 9(1) 62-66
[14] Cao L 2008 Research on Image Segmentation Algorithm Based on Otsu Theory Wuhan university of Technology
[15] Zhang A H 2004 Research on Image Segmentation Technology Based on Fuzzy Clustering Analysis Huazhong University of Science and Technology
[16] Wang F C, Zhang H and Gong L M 2016 Improved Roberts image edge detection algorithm Journal of Detection & Control 38(2) 88-92
[17] Qu Y D, Cui C S and Chen S B 2005 A fast subpixel edge detection method using Sobel – Zernike moments operator Image & Vision Computing 23(1) 11-17
[18] Zhang L , Zhang W and Liu X 2017 Existence of positive solutions for integral boundary value problems of fractional differential equations with p-Laplacian Advances in Difference Equations (1) 36
[19] Marr D Hildreth E 1980 Theory of edge detection Proceedings of the Royal Society of London 207(1167) 187-217