

# Fast Aerial UAV Detection Based on Image Segmentation and HOG-FLD Feature

## Fusion

Li Xiaoping

School of Computer Science and Engineering  
Xi'an Technological University  
Xi'an, 710021, Shaanxi  
e-mail: 919083845@qq.com

Wang Yanhong

School of Science  
Xi'an Technological University  
Xi'an, 710021, Shaanxi  
e-mail: 29314998@qq.com

Lei Songze

School of Computer Science and Engineering  
Xi'an Technological University  
Xi'an, 710021, Shaanxi  
e-mail: lei\_sz@163.com

Xiao Feng<sup>a</sup>, Tian Penghui<sup>\*b</sup>

School of Computer Science and Engineering  
Xi'an Technological University  
Xi'an, 710021, Shaanxi  
e-mail: <sup>a</sup>544070146@qq.com  
<sup>b</sup>809576423@qq.com

**Abstract**—In order to detect non-cooperative target UAV quickly and accurately, a novel method of UAV detection method based on graph theory and HOG-FLD feature fusion is presented in this paper. In order to avoid the time-consuming full search, the candidate areas of the UAV are obtained through the selective search of the image segmentation and the similarity, and the features are extracted through the method of gradient orientation histogram fusion FLD linear to train the SVM classifier with generalization ability to identify the UAV. The method can detect the UAV quickly and accurately under complicated background and circumstances of various position and angle. Compared with the sliding window method based on image segmentation and HOG+SVM, the experimental results show that the speed of this method has been obviously improved with the same recognition accuracy.

**Keywords**—Image Segmentation; Graph Theory; HOG; FLD; SVM

## I. INTRODUCTION

In recent years, UAV have been widely used in the military field. Due to its advantages of small size, low cost, low noise, high maneuverability, high concealment and superior stability, UAV can be applied to various of fields in the world for the purpose of reconnoitering enemy troops, detecting danger zone, tracking target, electronic interference, communication relay, and even completing the task of attack by carrying small-scale offensive weapons. Therefore, to detect non-cooperative UAV is necessary. To completed the task of independently identifying aerial non-cooperative UAV and monitoring in real time to border areas of different country could realize economic and extensive border surveillance and play a security role for the country and society.

Recently, on the UAV detection technology of domestic and foreign scholars is rare. However, the difficulty of the UAV detection technology is mainly that the UAV in the picture is easily affected by the position, the angle, the distance from the camera and its own structure, which makes it difficult to detect the UAV. Above problem is extremely similar to difficulty of target detection technology and needs to research and practice. At present, the common methods used in target

detection are mainly following methods: the method based on the optical flow field [1], the inter-frame difference method and the background detection based on the target detection method. Optical flow target detection is a method of moving target, which the basic principle is: that a velocity vector is given to the moving target pixel, and the image will form an optical flow field, and its background will be obviously distinguished from the moving target vector. In the meanwhile, you can get the location of the target by analyzing dynamic image. This method is suitable for all kinds of backgrounds, and it has no requirement for background. Its drawbacks are that the number of pixels is too large, the distribution of optical motion fields is too wide and inaccurate, and the computation is too large. The method of inter-frame difference [2] is used to obtain the foreground of the target by using the threshold segmentation method to obtain the difference between the pixels of the adjacent two frames, whose calculation is small, and the ability of target recognition is poor when the illumination changes. In order to ensure the accuracy and integrity of UAV target detection, the traditional method of sliding window to detect UAV requires that the sample picture and the picture to be tested need to be scaled until the target of the image to be detected is less than or equal to the template size and the detection is stopped. So that the UAV which is larger than or equal to the template size can also be detected accurately after adjusting the pixel size of the whole image to be tested. However, the HOG feature [3-5] is extracted after the image is scaled. Extracting HOG feature need to traverse the whole image by sliding window [3] until the edge contour feature of high dimension can be obtained by matching, so that the traversal need to take a long time, which leads to the slow generation of feature descriptors, and the scale of each level requires a large amount of computation.

An image segmentation method based on graph theory and HOG-FLD feature fusion is proposed for UAV detection in this paper. This method is divided into training stage and test stage. In the training stage, it is necessary to unify and gray-scale processed the pixel size

of positive and negative samples set be collected in advance, and then it could obtain the feature vector with small dimension which can stably describe the UAV contour feature after the feature extraction is carried out by using HOG-FLD feature fusion method. Finally, the feature vector is input into the support vector machine model, the support vector machine classifier is trained after a series of statistical learning. In the test stage, the candidate regions are screened out by using the segmentation method based on graph theory and are extracted features according to the feature extraction method of the training stage, and then the extracted features are input into the trained classifier to carry out classification and detecting whether the candidate area includes the target UAV. Image segmentation uses the connected method of graph [6] or the minimum spanning tree to merge the region, which mainly determines whether to merge the two regions according to the similarity of each region of the image. This method could not only obtain the global features of the image and the candidate area of UAV, but also the calculation speed is very fast and the processing efficiency is also high. The features obtained by combining the HOG features of gradient histogram and the Fisher linear discriminant analysis (FLD) are selected as the features of the statistical learning training classifier in feature extraction stage. Because of the advantages that the features of small dimension and good robustness and good classification ability extracted by HOG-FLD feature fusion are rarely affected by the changes of local illumination, background change, target location and angle change, it is easier to train a classifier that is easy to classify and has strong generalization ability.

## II. SEGMENTATION Method Based on Graph Theory to Obtain Region of Interest

### A. Segmentation Method Based on Graph Theory

Image segmentation is a way to separate the image into regions with different characteristics and separate the interested objects from the background in the segmented blocks. Let interest objects and backgrounds in segmented regions show a clear sense of contrast based on human visual effects. Image segmentation method plays an important role in further

analysis of image such as image compression and image recognition.

The image segmentation method based on graph theory (Graph-Based Image Segmentation) and the region merging method are selected to segment the image in this paper. The image segmentation method based on graph theory takes the image as the object transforming the image into the right undirected graph, and uses the connected branch method of graph [6] to segment the image, and then it could obtain the region block. The region merging method is to merge the segmented region blocks according to the specific merging rules. Because of great different characteristics of the UAV and the background in texture, color, size and degree of agreement, regional block of great different characteristics need to be decomposed and regional block of small different features need to be merged [11], which follows the principle of [10] the optimal segmentation.

#### B. Acquisition Principle of UAV Region of Interest

The image contains information such as shape, size, color, texture and so on. Image will be segmented after the whole image is traversed by the search algorithm of graph theory. In the graph theory, the definition of the image - to - graph [12] is described as follows:

$G=(V,E)$  is an undirected graph with  $V$  as its vertex and  $E$  as its edge set. Any edge  $(v_i, v_j) \in E$  is defined as a weight function  $W(v_i, v_j)$  greater than zero. Any vertex  $v_i, v_j$  represents the pixel point of the graph to be processed in graph theory, and the weight of any edge represents the gray difference, the distance or some other features between the adjacent pixels of  $v_i, v_j$  and  $(v_i, v_j)$  in an image. Using Graph Cuts algorithm [7], the graph  $G$  is divided into several disjoint independent regions, and the region is defined as  $G'=(V, E')$ ,  $E'$  is a subset of  $E$ . The regions of the picture is segmented by the algorithm of Minimum Spanning Tree(MST) [14] that the difference of the pixels in the same region is measured through the maximum weight

edge, of which the largest weight edge of segmentation region values are defined as follows:

$$Int(C) = \max_{e \in MST(C,E)} \omega(e) \quad (1)$$

$$C \in V \quad (2)$$

Where  $C$  is the generated partition region, and  $MST(C,E)$  represents the set of partitioned regions or the Minimum Spanning Tree [17] that  $C$  generates in  $E$ .

The resulting dissimilarity between pixels within the segmented region could also be understood as the weight of the minimum edge that connects the vertices of two segmented regions, as defined below:

$$Dif(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2, (v_i, v_j) \in E} w((v_i, v_j)) \quad (3)$$

When there is no edge connection between the two partitioned regions, there is  $Dif(C_1, C_2) = \infty$ . The condition for the appearance of its partition boundary is defined as follows:

$$D(C, C) = \begin{cases} true & \text{if } Dif(C_1, C_2) > MInt(C_1, C_2) \\ false & \text{otherwise} \end{cases} \quad (4)$$

If true in the expression is satisfied, then the boundary is represented, then the region is segmented. otherwise, it is merged.

Where the smallest internal difference of division is defined as follows:

$$MInt(C_1, C_2) = \min(Int(C_1) + \tau(C_1), Int(C_2) + \tau(C_2)) \quad (5)$$

$\tau()$  is a threshold function, which is used to control the merging degree of the image segmentation region, which is defined as  $\tau(C) = k/|C|$ , where  $|C|$  is the number of the segmentation region or the vertex, and  $k$  is a constant parameter that is used to control the coarse granularity of its image segmentation area. The segmentation effect is shown in Figure 1 (b). That small regions will be merged means detecting whether the segmented region meets the regional boundary conditions, if not, then they will be merged. The regional boundary conditions could be measured by the following four types of similarity in the image:

1) Color [13] similarity  $S_{colour}(r_i, r_j)$  means that 25 bins one-dimensional color histogram of the three color channels

are acquired after image is normalized. Or that means that the three color components are combined into one dimensional feature vector, so that each region of the image could be represented as a 75 dimensional feature vector  $C_i = \{c_i^1, \dots, c_i^n\}$ , and the color calculation formula of the region is:

$$S_{colour}(r_i, r_j) = \sum_{k=1}^n \min(c_i^k, c_j^k)$$

$$C_i = \frac{size(r_i) \times C_i + size(r_j) \times C_j}{size(r_i) + size(r_j)} \quad (6)$$

Where  $size(r_i)$  is the number of pixels contained in the region ( $r_i$ ), and the number of pixels contained in the new region is  $size(r_i) + size(r_j)$ .

2) Texture [13] similarity  $S_{texture}(r_i, r_j)$  means that calculate Gaussian differential for 8 different directions of three color channels separately (its  $\sigma = 1$ ), 10 bins one-dimensional histogram of each direction of three color channels is respectively obtained after image is normalized, and then the region could be expressed as a 240-dimensional vector  $T_i = \{t_i^1, \dots, t_i^n\}$ , and its texture similarity calculation formula is as follows:

$$S_{texture}(r_i, r_j) = \sum_{k=1}^n \min(t_i^k, t_j^k) \quad (7)$$

3) Dimension similarity  $S_{size}(r_i, r_j)$ , which is used to merge smaller regions as early as possible. Its formulas are as follows:

$$S_{size}(r_i, r_j) = 1 - \frac{size(r_i) + size(r_j)}{size(im)} \quad (8)$$

Where  $im$  refers to the whole image.

4) The matching similarity  $fill(r_i, r_j)$  is used to merge the intersected regions as soon as possible. Its account form is as follows:

$$fill(r_i, r_j) = 1 - \frac{size(BB_{ij}) - size(r_i) - size(r_j)}{size(im)} \quad (9)$$

the similarity set  $S$  could be gotten by combining these four kinds of similarity, and the formula of the combination is as follows:

$$S(r_i, r_j) = a_1 S_{colour}(r_i, r_j) + a_2 S_{texture}(r_i, r_j) + a_3 S_{size}(r_i, r_j) + a_4 S_{fill}(r_i, r_j) \quad (10)$$

In order to judge whether the boundary condition of  $a_i \in \{0,1\}$  satisfies the similarity. The two segmented region blocks  $r_i$  and  $r_j$  which have the largest similarity are merged and a region  $r_i$  is obtained after the similarity degree of the segmented regions in the similarity set  $S$  is compared and sorted. The similarity between two adjacent regions of  $r_i$  and  $r_j$  in the set  $S$  is removed in the meanwhile. Then according the step of this method, we begin to calculate the similarity between the new merged region  $r_i$  and its current adjacent regions, and we add this similarity to the similarity set  $S$  again. At the same time, we also add  $r_i$  to the region set  $R$ , and we finally label the region with a rectangular box. The segmentation effect is shown in Figure 1 (c) and (d).

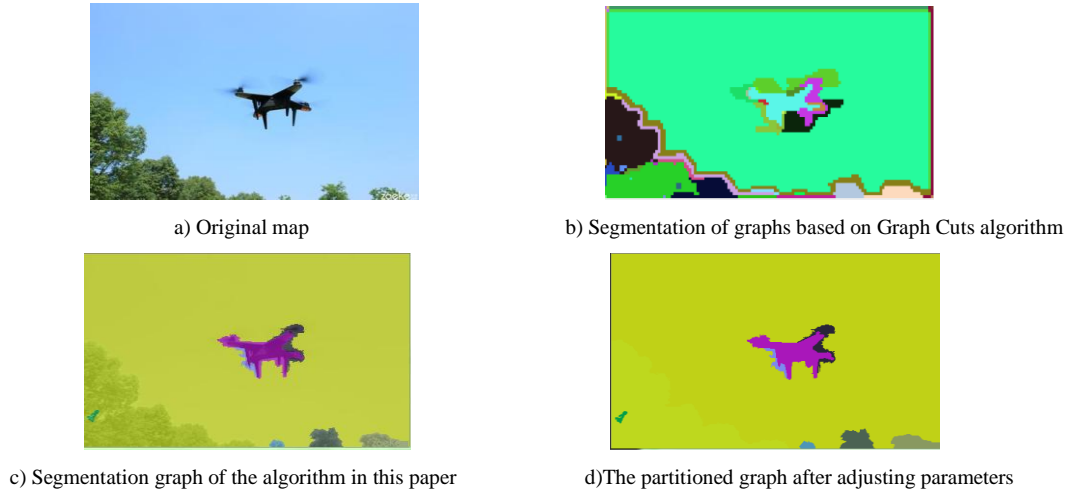


Figure 1. Image Segmentation

### III. TARGET UAV DETECTION BASED ON HOG-FLD FEATURE FUSION AND SVM DETECTION

The method could be divided into two stages from the perspective of practice: training and testing. In the training stage, the window with constant shape and size is used to traverse the whole UAV or non-UAV sample images, and the feature is extracted by HOG-FLD feature fusion method. A computable feature vector is obtained to describe the contour of the image, which is robust and easy to classify. The SVM classifier [8-9] for UAV detection could be obtained by statistical analysis and learning of these Eigenvectors. In the early stage of the test, the same method as the training stage is used to extract the feature of the regions to be detected, and the candidate targets of these areas are classified by the trained SVM classifier in the later stage to determine if the target in the candidate area is a UAV.

#### A. Feature Extraction From HOG-FLD Feature Fusion

Because the edge contour feature of UAV has strong stability and extensibility, the HOG feature with strong ability to describe contour is extracted from data set. However, because of the large dimension of HOG feature vector, it is unfavorable to the training of classifier. In order to train the classifier with good classification ability, it is necessary to reduce the dimension of extracted HOG feature by FLD [18] feature fusion and finally obtain the

feature vector with small dimension. To sum up, the feature extraction based on HOG-FLD feature fusion is adopted in this article.

The basis of feature extraction in this article is to extract feature of gradient histogram of oriented gradient (HOG). The thought of algorithm is as follows: calculating the gradient of selected image; The whole image is divided into rectangular cells of fixed size and equal size, each cell has pixel including  $m \times m$ ; the cell is divided into 9 undirected channels) or 18 directed channels, and the gradient histogram of each direction is voted on. The voting weight is the gradient value that has been calculated earlier. The divided cell units above are composed of fixed blocks of the same size. Each block contains  $n \times n$  cell elements, which the local eigenvector corresponding to each block is normalized in order to reduce the effect of illumination on the results. The feature vectors of these blocks are combined to form the HOG feature vectors of the image.

The steps of extracting HOG feature are as follows: graying the positive sample image of UAV, filtering the input image with Gamma correction method so that the image could achieve the standard contrast of color space and reduce the effects of local shadows and light changes on the image; dividing the image into a certain number of cells, and forming a fixed number of cells into a certain number of blocks of equal size as in the preceding paragraph, classifying the gradient range according to the above mentioned rules, we could

calculate the cell features in these blocks and finally connect all the blocks together to obtain the feature vector containing the whole target UAV image. Formula (11) is a formula for normalizing the image, formula (12) and formula (13) could calculate the gradient component of each pixel, formula (14) and formula (15) could calculate the size and direction of the gradient.

$$v_g \leftarrow \sqrt{v_g / (\|v_g\|_1 + \varepsilon)} \quad (11)$$

$$Gx_{(x,y)} = pi_{(x+1,y)} - pi_{(x-1,y)} \quad (12)$$

$$Gy_{(x,y)} = pi_{(x,y+1)} - pi_{(x,y-1)} \quad (13)$$

$$S_{(x,y)} = \sqrt{Gx^2_{(x,y)} + Gy^2_{(x,y)}} \quad (14)$$

$$\theta_{(x,y)} = \arctan(Gy_{(x,y)} / Gx_{(x,y)}) \quad (15)$$

Where  $v_g^*$  is the result of histogram normalization, and  $v_g$  is the extracted vector histogram;  $pi_{(x+1,y)}$ ,  $pi_{(x-1,y)}$ ,  $pi_{(x,y+1)}$ , and  $pi_{(x,y-1)}$  respectively denote the location of 4 pixel points;  $Gx_{(x,y)}$  and  $Gy_{(x,y)}$  respectively denote the coordinate position in the horizontal and vertical directions of the two pixels. And  $S_{(x,y)}$ ,  $\theta_{(x,y)}$  respectively denote the length of the gradient direction vector and the angle of the gradient direction vector.

In this article, a window including  $64 * 128$  pixel is used to scan the sample image and the image to be detected using a window including  $64 * 128$  pixel. The scanning step size is 8 pixels (scanning is in horizontal and vertical direction). The window is divided into cells including  $8 * 8$  pixel and forms  $8 * 16 = 128$  units. Then setting up four adjacent units up and down to the left and right as a block of pixels, a window contains 105 blocks

of pixels. A 3780 dimensional feature vector named HOG feature description value is generated in a window containing 105 pixel blocks according to the calculation steps of HOG. Its specific HOG algorithm is shown in Figure 2:

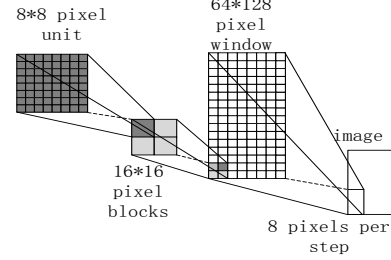


Figure 2. Principle Diagram of HOG Algorithm

On the basis of HOG feature, the linear subspace is constructed by using Fisher Linear Discriminant Analysis (FLD) [18]. By calculating the optimal projection matrix, the projection matrix for feature extraction of the training set is obtained, and the similarity of its projection vector is taken as the similarity degree  $S_{\cos}$  of cosine similarity, which purpose is to reduce the intra-class dispersion  $S_W$  as much as possible and to increase the inter-class dispersion  $S_B$  as much as possible (or that is in the training concentration, to make the sample data of the same UAV as close as possible, and the sample data of different UAV to be far away). In this way, the features with classification ability are extracted. When dealing with a problem that type is c, the mathematical formulas of inter-class dispersion  $S_B$  and intra-class dispersion  $S_W$  should be defined as follows:

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu) (\mu_i - \mu)^T \quad (16)$$

$$S_W = \sum_{i=1}^c \sum_{x_k \in \Omega_i} (x_k - \mu_i) (x_k - \mu_i)^T \quad (17)$$

Where  $\mu_i$  denotes the mean value of class  $\Omega_i$ ;  $\mu$  means the mean value of total sample;  $N_i$  means the number of samples of class  $\Omega_i$ . The optimal projection matrix  $W_{opt}$  could be obtained by solving the optimization problem such as formula (18), where  $S_W$  must be a nonsingular matrix (or that is

the total number of training samples  $N$  is greater than the characteristic dimension of UAV image):

$$W_{opt} = \arg \max_w \frac{|W^T S_B W|}{|W^T S_W W|} \quad (18)$$

$W_{opt}$  could also be obtained by solving the generalized eigenvalue problem such as the formula (19):

$$S_B W = S_W W \lambda \quad (19)$$

In order to solve the problem that the intra-class dispersion matrix  $S_w$  is singular, PCA principal component analysis (PCA) is used to reduce the dimension of the feature space (dimensionality reduction to  $N$ -cu), and then Fisher linear discriminant analysis (FLD) is used to deal with it. The projection vector  $y$  of test sample  $x$  is obtained according to formula (20):

$$y = W_{opt}^T x \quad (20)$$

Cosine similarity  $S_{cos}$  is used as the similarity measure of projection vector  $y$ . Where the cosine similarity  $S_{cos}$  of vector  $A = \{a_1, a_2, \dots, a_n\}$  and  $B = \{b_1, b_2, \dots, b_n\}$  is defined as follows:

$$S_{cos} = \frac{\langle A, B \rangle}{\|A\|_2 \|B\|_2} = \frac{\sum_{i=1}^n a_i b_i}{\|A\|_2 \|B\|_2} \quad (21)$$

### B. Support Vector Machine

Because the support vector machine (SVM) proposed by Vapnik has the advantages of simple system structure, global optimization, good generalization, and short training and prediction time [9], this paper uses SVM as a machine learning tool to calculate the rule of samples in order to achieve fast and efficient learning sample features and accurate classification purposes. The main idea of SVM is to deal with the linear inseparability of the original space by selecting the kernel function of Polynomial Kernel to correspond the data to the high-dimensional space. When the algorithm is used to realize the two-classification, the sample features such as

HOG must be extracted from the original space first, and then the sample features in the original space are represented as a vector in the high-dimensional space. In order to minimize the error rate of the two class classification problems, we need to find a hyperplane that is used to divide the two classes in the high dimensional space.

Let the sample set be  $x_i, y_i$ , where  $i=1,2,\dots,N$ ,  $x_i \in R^e$ ,  $y_i \in \{0,1\}$  is the class identifier. Then in  $e$  dimensional space, its linear discriminant function is as follows:

$$g(x) = w \cdot x + b \quad (22)$$

The formula of the classification surface equation is as follows:

$$w \cdot x + b = 0 \quad (23)$$

After normalization of the discriminant function, the following conditions must be satisfied for the two types of samples:

$$g(x) \geq 1 \quad (24)$$

The classification interval could be  $\frac{2}{\|w\|}$ , where the maximum requirement  $\|w\|$  for the classification interval is kept to a minimum, all samples must be correctly classified, and the following conditions must be met:

$$y_i [(w \cdot x_i) + b] - 1 \geq 0 \quad (25)$$

An SVM whose inner product function is  $k(x_i, x_j)$  is constructed with the following formulas (which can be understood as the formula for calculating the extreme value of a quadratic function with conditional constraints):

$$Q(a) = \sum_{i=1}^N a_i a_j y_i y_j k(x_i, x_j) \quad (26)$$

$$0 \leq a_i \leq C, \sum_{i=1}^N a_i y_i = 0$$

Its constraints are expressed as:

The formulas of the support vector machine that could be calculated are as follows:

$$f(x) = \text{sgn}\left(\sum_{i=1}^N a_i^* y_i k(x, x_i) + b^*\right) \quad (27)$$

Among them,  $b^*$  is a constant parameter, indicating the size of the threshold that needs to be classified.

### C. Sample Preparation and Classifier Training

In this article, 900 positive sample images and 900 negative sample images are selected, and 500 original images of positive samples are selected to calculate the aspect ratio of UAV. An image that aspect ratio is 1:2 is obtained, and the pixels are normalized to 256 \* 512 to avoid the effect of image size on the recognition effect. As described in the HOG feature section above, each image could produce 105 pixel blocks, and each pixel block could obtain a 36-dimensional feature vector, so a image will finally produce a 3780 dimensional HOG feature vector. The extracted HOG vector is used as the input vector of Fisher linear discriminant analysis algorithm to reduce the dimension of the whole vector. The dimension of the vector is adjustable and its parameter is determined according to the recognition efficiency of the experiment. Sending the vector reduced dimension to the SVM model, the SVM classifier is obtained to detect whether the UAV is included in the image under test.

## IV. PARAMETER ANALYSIS AND EXPERIMENTAL RESULTS

The environment used in this experiment is that the simulation software of Matlab R2015b based Windows 7 on 64-bit operating system. Where CPU of computer is i5-6500 and its installed memory is 4.0 GB, and the number of the test images are 200. The best detection effect is obtained by adjusting the important parameters that affect the experimental results.

### A. Parameter Selection of FLD

In the process of extracting feature vector, the dimension that needs to be reduced by Fisher linear discriminant analysis algorithm is the parameter k of FLD algorithm, which changes with the change of this parameter. The contrast graph of recognition time is

shown in Figure 3. It can be seen from the graph that when k is 50, the whole algorithm has the least recognition time.

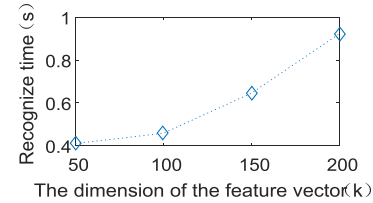


Figure 3. The time Chart with Parameter k Changed

### B. Experimental Effect of Aerial UAV Detection Algorithm Based on Region of Interest

Experiments show that in the whole algorithm, when the number of pixels in the divided blocks is 8\*8, SVM kernel function type is 2, when the threshold of segmentation is k=90 and sigma=10, The algorithm has the best overall recognition effect on accuracy rate and time, and the recognition result is shown in Figure 4:

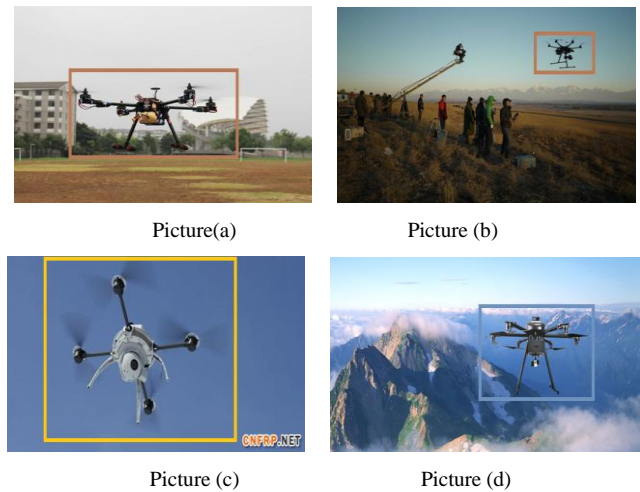


Figure 4. Identification Results

### C. Comparison of Detection Effect of Traditional HOG-SVM UAV Detection Algorithm

In order to verify the efficiency of the proposed method, the experimental results of the proposed method are compared with the experimental results of the HOG and SVM method based on image segmentation, and the experimental results are obtained by statistics. As shown in TABLE I, and the average recognition time of this method is faster than that of the latter method.



TABLE I. COMPARISON RESULTS OF THE TEST METHODS

Test Method	Accuracy Rate	Time
HOG-FLD+SVM	93.35%	0.090s
HOG+SVM	95.75%	0.160s

## V. TAG

The mechanism based on the region of interest to obtain the region of interest is used in this article. In the testing phase, the acquired regions of interest are input into the trained SVM classifier, which reduces the recognition time. The dimension reduction of Fisher linear discriminant analysis (FLD) makes SVM easy to train and in the phase of feature vector extraction. Compared the HOG-SVM detection method based on image segmentation with the region of interest based aerial UAV detection algorithm, this paper uses Matlab to carry out simulation experiments. The experimental results show that the detection algorithm based on region of interest is better than the sliding window method based on image segmentation to detect UAV in terms of accuracy and time.

## ACKNOWLEDGMENT

Fund projects: Key projects in the Industrial Field of Shaanxi (2016KTZDGY4-09), Scientific Research Program Funded by Shaanxi Provincial Education Department (17JK0364), National Natural Science Foundation of China (61572392), State and Provincial Joint Engineering Lab. of Advanced Network and Monitoring Control (GSYSJ2016008), Scientific Research Program Funded by Shaanxi Provincial Education Department (16JK1379).

## REFERENCES

- [1] J. Barron, D. Fleet, S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, vol.12, pp. 42-77, 1994.
- [2] A. Lipton, H. Fujiyoshi, R. Patil, "Moving target classification and tracking from real-time video," *Proc IEEE Workshop on Applications of Computer Vision*, pp. 8-14, 1998.
- [3] DALALN, TRIGGS B, "Histograms of oriented gradients for human detection," *Proc IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2005.
- [4] Qiang Zhu, Shai Avidan, Mei Chen Yeh, and Kwang Ting Cheng, "Fast human detection using a cascade of histograms of oriented gradients," *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, 2006.
- [5] Suard F, Akotomamonjy A R, Bensrhair A, et al, "Pedestrian Detection Using Infrared Images and Histograms of Oriented Gradients," *Proceedings Intelligent Vehicle Symposium*, pp. 206-212, 2006.
- [6] Felzenszwalb P F, Huttenlocher D P, "Efficient Graph-Based Image Segmentation," *International Journal of Computer Vision*, vol.59, pp. 167-181, 2004.
- [7] Humayun A, Li F, Rehg J M. RIGOR, "Reusing Inference in Graph Cuts for Generating Object Regions," *Proc. IEEE Computer Vision and Pattern Recognition*, pp. 336-343, 2014.
- [8] Vapnik V N, "The nature of statistical learning theory," Springer-Verlag, New York, pp. 37-69, 1995.
- [9] Guo Mingwei, ZhaoYuzhou, Xiang Junping, et al, "A Survey of Target Detection Algorithms Based on Support Vector Machine," *Control and Decision*, vol. 29, pp. 192-200, 2014.
- [10] Zhang Han, He Dongjian, "An Image Segmentation Method Based on Texture Information and Graph Theory," *Computer Science and Engineering*, vol. 50, pp. 180-184, 2014.
- [11] Zhai Jiyou, Zhuang Yan, "Significant Detection of Boundary Prior and Adaptive Region Merging," *Computer Engineering and Application*, 2017.
- [12] Chen Shanchao, Fu Hongguang, Wang Ying, "Application of An Improved Graph Segmentation method in Tongue Image Segmentation," *Computer Engineering and Application*, vol. 48, pp. 201-203, 2012.
- [13] Yan Yu, Song Wei, "Color and Texture Mixed Descriptor Image Retrieval Method," *Computer Science and Exploration*, pp. 1-8, 2016.
- [14] Ye Qing, Hu Changbiao, "An improved Image Segmentation Method Based on Graph Theory," *Computer and Modernization*, vol. 253, pp. 64-67, 2016.
- [15] Sande K E A V D, Uijlings J R R, Gevers T, et al, "Segmentation as selective search for object recognition," *International Conference on Computer Vision. IEEE Computer Society*, pp. 1879-1886, 2011.
- [16] Girshick R, Donahue J, Darrell T, et al, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," *Computer Science*, pp. 580-587, 2014.
- [17] Wang Ping, Wei Zheng, Cui Weihong, "A Minimum Spanning Tree Image Segmentation Criterion Based on Statistical Learning Theory," *Information Science Edition, Journal of Wuhan University*, vol. 42, pp. 878-883, 2017.
- [18] Belhumeur P, Kriegman D, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.19, pp.711-720, 1997.