Research on Image Static Target Recognition and Extraction Based on Space Intelligent System

Yufei Huang(✉), Jia Xu, Xiangyu Lin, Xiongwen He, Ran Zhang, and Wenjie Li

Beijing Institute of Spacecraft System Engineering, Beijing, China
billyyufei@sina.com

Abstract. With the continuous development of remote sensing technology, remote sensing images and other multimedia information in the civil field has become more and more widely used. Unmanned aerial vehicle (UAV)/satellite image recognition and extraction of ground targets has become one of the important means of information acquisition in the civil field. At the same time, intelligent systems (such as mobile phones) have gradually become the core of UAV/nano-satellites. In this paper, through JAVA, the recognition and extraction of static targets in remote sensing images are realized. This method provides a theoretical basis for future image processing in remote sensing applications.

Keywords: Space intelligent system · Remote sensing image · Static target recognition and extraction

1 Introduction

Space smart platforms, such as smart phones, are becoming the core of UAVs/nano satellites. Since April 2013, NASA has used Orbital Science's Antares rocket to launch mobile satellites into space and explore their applications. This paper discusses the application of target recognition and extraction in remote sensing image in intelligent platform, analyzes the method of static target recognition and extraction in remote sensing image based on intelligent system platform, and finally demonstrates the method by using JAVA to verify the vehicle in parking lot, which provides theoretical basis for remote sensing satellite image acquisition and processing in the future. The foundation.

2 Principles

The core principles of static target recognition and extraction based on intelligent system platform are: imaging a certain area of the ground by satellite or unmanned aerial vehicle (UAV) camera; focusing on the scale problem in static target extraction based on satellite image after pre-processing, the scale conversion of vehicle satellite image in parking lot is carried out. In the evaluation and analysis of scale effect, the

M. Jia et al. (Eds.): WiSATS 2019, LNICST 281, pp. 91–99, 2019.
https://doi.org/10.1007/978-3-030-19156-6_9

correlation between scale and object extraction is used to extract static objects, and all the static objects identified in the image are numbered in order. The specific task flow chart is as follows (Fig. 1):

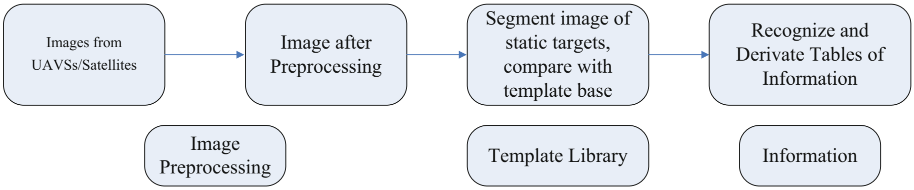

Fig. 1. Task process diagram.

2.1 Image Preprocessing

Image Enhancement

Image distortion often occurs in the process of acquiring images, which makes the obtained image different from the original image to some extent. For some imaging conditions, the weak signal is highlighted to make it easier to distinguish. The methods of image enhancement are divided into frequency domain method and spatial domain method. Spatial filtering is based on the relationship between pixels and adjacent pixels. Neighborhood processing method in spatial domain is used to highlight some features of the image. In each transform domain, the transformed coefficients, such as the coefficients of Fourier transform, DCT transform and so on, are modified to manipulate the image, and then inversely transformed to obtain the processed image.

(1) Enhancement of Grayscale Transformation

In case of $x = f(i,j)$ as the gray value of the original image, $y = F(i,j)$ as the gray value of the transformed image, the relationship of the function $y = h(x)$ as follows:

$$y = h(x) = \begin{cases} \frac{y_1}{x_1}x, 0 \leq x < x_1 \\ \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1, x_1 < x < x_2 \\ \frac{255 - y_1}{255 - x_1}(x - x_2) + y_2, x_2 < x \leq 255 \end{cases} \quad (1)$$

(x_1, y_1) and (x_2, y_2) are the two turning points. These two turning points determine the slope of three straight lines.

(2) Histogram Transformation Enhancement

The cumulative distribution function is used in histogram equalization because the cumulative distribution function is a monotone increasing function (control size relationship) and the range is 0 to 1 (control crossing problem), so the cumulative distribution function is used in histogram equalization.

In case of the area of the image $f(i,j)$ is A, and there are L gray levels. n_i is used to represent the pixel list of the grayscale r_i. Histogram equalization is to establish a transformation T, that is, to construct a monotone increasing function is as follows:

$$r' = T(r), r \in [f_{\min}, f_{\max}] \tag{2}$$

The gray value r of the original image $f(i,j)$ is transformed into the gray $r' \in [g_{\min}, g_{\max}]$ of the output image $g(i,j)$.

The formula for histogram equalization is as follows:

$$r' = \left\lfloor \frac{L-1}{A} \sum_{k=0}^{i} n_k + 0.5 \right\rfloor, (i = 0, 1, \ldots, L-1) \tag{3}$$

(3) Smoothing Filtering

There are two types of smoothing filtering: one is fuzzy; the other is noise cancellation. Smoothing filtering in spatial domain usually uses simple averaging method, which is to find the average brightness of adjacent pixels. The size of the neighborhood is directly related to the smoothing effect. The mean filtering formula for setting A is as follows:

$$F(i,j) = \frac{1}{M} \sum_{(i',j') \in A} f(i',j') \tag{4}$$

(4) Median Filtering

The basic principle of median filtering is to replace the value of a point in a digital image or sequence with the median value of each point in a neighborhood of the point, so that the value of the surrounding pixels close to the true value, thereby eliminating isolated noise points. It uses a neighborhood A containing odd points, which is called a sliding window, and replaces the gray value of the pixel in the center of the window with the median gray value of each point in the window.

(5) High Pass Template Filtering

The high pass filter can smooth the high-frequency components and weaken the low frequency. The edge and detail of the image are mainly located in the high frequency part, and the blur of the image is caused by the weak high frequency component. High pass filter can be used to sharpen the image, in order to eliminate blurring and highlight edges. Therefore, high-pass filter is used to let the high-frequency component pass through, so that the low-frequency component is weakened. The 5 convolution templates commonly used for high pass filters are sometimes referred to as convolution kernels.

$$H_1 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}, H_2 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}, H_3 = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix},$$
$$H_4 = \frac{1}{7}\begin{bmatrix} -1 & -2 & -1 \\ -2 & 19 & -2 \\ -1 & -2 & -1 \end{bmatrix}, H_5 = \frac{1}{2}\begin{bmatrix} -2 & 1 & -2 \\ 1 & 6 & 1 \\ -2 & 1 & -2 \end{bmatrix} \tag{5}$$

Simulation takes H_5 as an example.

Target Edge Detection

In this paper, Roberts operator is used, which uses the difference between two adjacent pixels in diagonal direction to approximate the gradient amplitude to detect the edge, so it is also called four-point difference method. Vertical edge detection is better than oblique edge detection. It has high positioning accuracy and is sensitive to noise and cannot suppress the influence of noise. The formula for the Roberts operator is as follows:

$$|gradf(x,y)| = \sqrt{(f(x+1,y+1) - f(x,y))^2 + (f(x+1,y) - f(x,y+1))^2} \tag{6}$$

$$|gradf(x,y)| = |f(x+1,y+1) - f(x,y)| + |f(x+1,y) - f(x,y+1)| \tag{7}$$

Image Morphology Design

This section is mainly reflected in noise filtering, area filling, expansion and corrosion, to a certain extent, can eliminate the resulting contour loss and other phenomena. The algorithm used in noise filtering is first open operation, then closed operation; region filling operation first assigns 1 to the point within the boundary as seed point, and then fills it by iterative method; expansion and corrosion operations are defined as follows:

$$\begin{cases} (f \oplus b)(s,t) = \max\{f(s-x,t-y) + b(x,y) \big| (s-x,t-y) \in D_f, (x,y) \in D_b\} \\ (f \odot b)(s,t) = \max\{f(s+x,t+y) + b(x,y) \big| (s+x,t+y) \in D_f, (x,y) \in D_b\} \end{cases} \tag{8}$$

D_f and D_b are the definitions of image f and structure element b, which are determined by the width and height of the image.

2.2 Image Segmentation

In this paper, a threshold based image segmentation method is used in image segmentation and initial recognition. The aim is to divide the pixel set according to the gray level, and each subset is formed into a region corresponding to the real scene. Each region has the same attributes, while the adjacent regions do not have the same attributes. The basic principle is to divide pixels into several classes by setting characteristic thresholds. Common features include gray or color features directly from the original image, and features transformed from the original gray or color values. In this paper, we use the best threshold segmentation iterative algorithm. The steps are as follows:

The steps are as follows:

1. The maximum and minimum gray values of the image are calculated by $P_{\max}$ and $P_{\min}$, respectively. The initial value of the threshold T is as follows:

$$T_0 = (P_{\max} + P_{\min})/2 \tag{9}$$

2. Then, according to the threshold $T_k (k = 0, 1, \ldots, k)$, the image is divided into two parts: the target and the background, and the average gray value P_O and P_B of the two parts are calculated, as follows:

$$\begin{cases} P_O = \dfrac{\sum_{p(i,j) < T_k} p(i,j)\omega(i,j)}{\sum_{p(i,j) < T_k} \omega(i,j)} \\ P_B = \dfrac{\sum_{p(i,j) > T_k} p(i,j)\omega(i,j)}{\sum_{p(i,j) > T_k} \omega(i,j)} \end{cases} \tag{10}$$

In the upper form, $p(i,j)$ is the gray value of the image at (i,j), and $\omega(i,j)$ is the weight coefficient of the gray value $p(i,j)$, usually taking $\omega(i,j) = 1.0$.

3. The new threshold is calculated as follows:

$$T_{k+1} = (P_O + P_B)/2 \tag{11}$$

4. If $T_k = T_{k+1}$, then the program ends, and the final T_k is the best threshold T. Otherwise, $k \leftarrow k + 1$ returns to step 2.

2.3 Image Recognition and Extraction

The method of image recognition is based on template matching algorithm. A commonly used measure for template matching is the sum of squares of errors in the corresponding regions of the template and the source image, which measures the difference between the blocks and templates in the original image, i.e. the square error measure. If $f(x,y)$ is the source image of $M \times N$ and $t(j,k)$ is the template image of $J \times K (J \leq M, K \leq N)$, the error square sum measure is defined as follows:

$$D(x,y) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} [f(x+j, y+k) - t(j,k)]^2 \tag{12}$$

For upper type expansion, the following formula can be obtained.

$$D(x,y) = \sum_{j=0}^{J-1}\sum_{k=0}^{K-1}[f(x+j,y+k)]^2 - 2\sum_{j}^{J-1}\sum_{k}^{K-1}t(j,k)\bullet f(x+j,y+k) + \sum_{j=0}^{J-1}\sum_{k=0}^{K-1}[t(j,k)]^2 \quad (13)$$

If:

$$\begin{cases} DS(x,y) = \sum_{j=0}^{J-1}\sum_{k=0}^{K-1}[f(x+j,y+k)]^2 \\ DST(x,y) = 2\sum_{j=0}^{J-1}\sum_{k=0}^{K-1}t(j,k)\bullet f(x+j,y+k) \\ DT(x,y) = \sum_{j=0}^{J-1}\sum_{k=0}^{K-1}[t(j,k)]^2 \end{cases} \quad (14)$$

In the above formula, $DS(x,y)$ denotes the energy of the corresponding area of the template in the source image. It is related to the pixel position (x,y), but the change of $DS(x,y)$ is slow with the change of the pixel position (x,y). $DST(x,y)$ template and the corresponding region of the source image are correlated, which changes with the change of the pixel position (x,y). When the template $t(j,k)$ matches the corresponding region of the source image, the maximum value is obtained. $DT(x,y)$ represents the energy of the template, which is independent of the pixel position (x,y) of the image, and can be calculated only once, which reduces the amount of calculation. After the comparison of all pixels, finding the smallest error is the result of matching (Fig. 2).

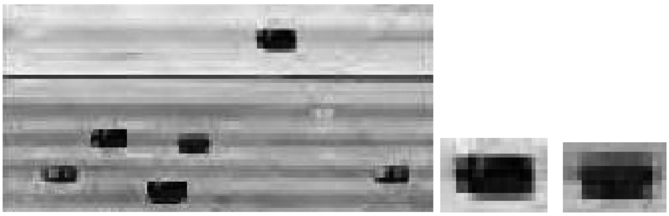

Fig. 2. Source images and partial template images

3 Simulation

This section mainly compiles and implements the above contents. The parking lot is used as the setting scene. The contrast expansion parameters are set as follows: X1 = 50, Y1 = 30, x2 = 200, y2 = 210. The template of high-pass template filtering is H5. The simulation results are as follows (Figs. 3, 4, 5, 6 and 7):

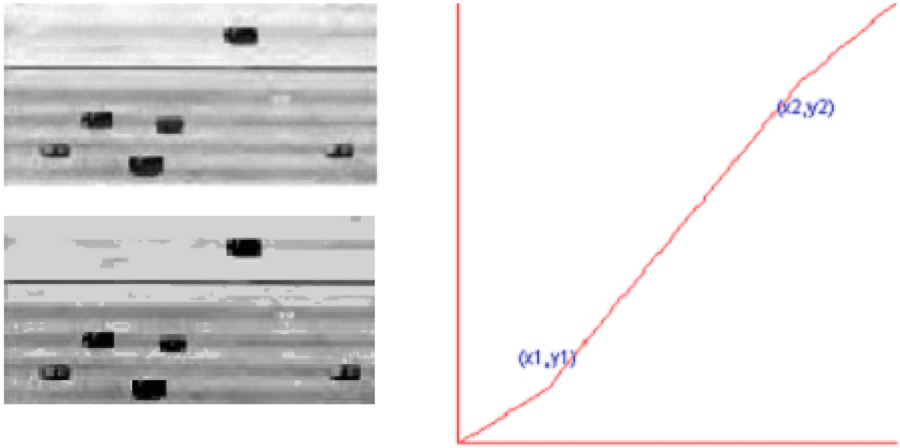

Fig. 3. Visual preview of the original image and contrast extension, as well as comparison of effects (top-left picture is the original one, bottom-left picture is the processed one, as contrast extension)

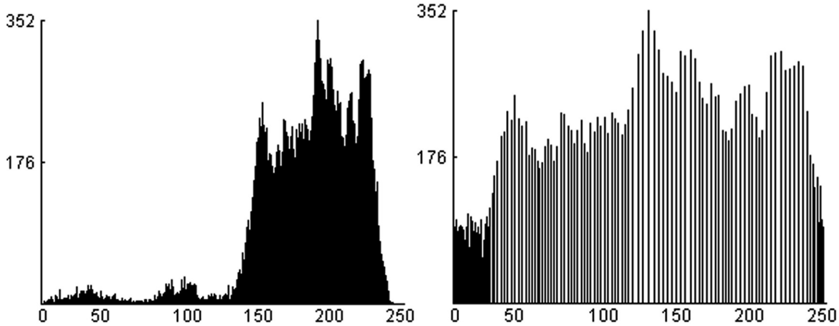

Fig. 4. Histogram comparison before and after homogenization

Fig. 5. Template selection and effect comparison of high pass template filtering

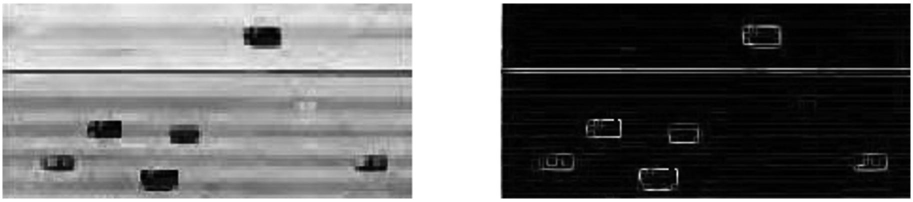

Fig. 6. Target edge detection based on Roberts operator

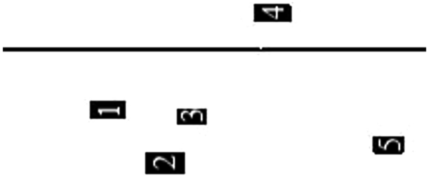

Fig. 7. The example of image morphology and image segmentation and recognition algorithm.

There are six target vehicles in the source image. After a series of processing, one of the targets in the lower left corner is regarded as an invalid target because the edge is not obvious. The other five targets can be recognized in the image, and the purpose is basically achieved.

4 Summary

In this paper, the key technologies of remote sensing image recognition and extraction system based on space intelligent system platform are analyzed and the strategy is studied. The image pretreatment, target edge detection, image segmentation and preliminary recognition are studied, which provides a theoretical basis for the future remote sensing satellite image acquisition and processing.

References

1. Ren, C., Huang, H., Tan, Y., et al.: Vehicle identification from remote sensing image based on image symmetry. Remote Sens. Land Resour. **28**(4), 135–140 (2016)
2. Mishra, R.K.: Automatic moving vehicle's information extraction from one-pass worldview-2 satellite imagery. Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci. **39**, 323–328 (2012)
3. Zhao, Q., Chang, B., Wang, Y., Huang, H., Tan, Y., et al.: Fuzzy segmentation and genetic algorithm based road vehicle extraction method from high resolution aerial image. Bull. Surv. Mapp. (8), 62–66 (2017)
4. Cao, T., Shen, L.: A method for vehicles detection method based on traffic remote sensing. In: The Eighth China Intelligent Transportation Annual Meeting, pp. 389–394 (2013)

5. Wu, D., Liu, D., Puskas, Z., et al.: A learning based deformable template matching method for automatic rib centerline extraction and labeling in CT images. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Providence, pp. 980–987. IEEE (2012)
6. Li, S., Xu, Y., Sun, W., et al.: Remote sensing image recognition for vehicles based on self-feedback template extraction. J. South China Univ. Technol. (Nat. Sci. Ed.) **42**(5), 97–102 (2014)
7. Chen, X., Xiang, S., Liu, C.L., et al.: Vehicle detection in satellite images by hybrid deep convolutional neural networks. IEEE Geosci. Remote Sens. Lett. **11**(10), 1797–1801 (2014)