A Novel Object Detection Approach Based on the Boundary Shape Information from High Resolution Satellite Imagery

XIAOSHU SI, XUEMIN HU, HONG ZHENG
Joint Lab for Intelligent Computing and Intelligent Systems, School of Electronic Information
Wuhan University
129 Luoyun Road, Wuhan, Hubei 430079
CHINA
hong-zheng@126.com, huxuemin2003@163.com

Abstract: - This paper presents a novel approach of detecting special objects from high resolution satellite imagery. In this approach, a bilateral filtering is used to denoise firstly, and a new morphological approach which combines gray scale morphological processing and binary morphological processing is proposed for ROI extraction and feature enhancement. A detection operator based on the boundary shape information (BSI) is developed to detect enhanced objects. The experiments on images from Google Earth are discussed in the paper. The experimental results show that proposed approach is effective and feasible. Compared with other object detection approaches from high resolution satellite imagery such as PCA or MSNN, the proposed approach has better detection performance.

Key-Words: - BSI, Detection template, Object detection, Vehicle, Aircraft, High resolution satellite imagery

1 Introduction

With the development of science and technology, the technology of optical satellite develops quickly and some new applications of the satellite imagery like object detection emerged. However, satellite image data is too complex to operator manually. So the automation and intelligence of target recognition are required.

Up to now, lots of study about automatic target recognition or detection have been done. However, most of them are about large objects detection, such as bridge detection and airport detection [1,2]. Some study for the small object detection are limited to aerial images. These approaches include Bayesian Network, the derivative of Gaussian model, 3D model and background approach, local operator and image fusion approach [3~6] etc. But study for small object detection based on high resolution satellite image is few. Up to this point, typical approaches include Principal Component Analysis (PCA), Morphological Shared-Weight Neural Network (MSNN) and Danger Perceptive Network (DPN) [7~9]. These approaches work well in their object detection experiments. But these approaches need a lot of training samples, and the detection result depends on the training results of object samples seriously. However, we sometimes can't obtain satisfactory samples in a practical work, so these methods can't meet our practical demand. In order to solve the problem, an object detection approach based on the boundary shape information is proposed in this paper. This approach can detect the objects with special shape without training samples. The experimental results on satellite images from Google Earth show that this approach is effective and feasible.

The paper is organized as follows. In Section 2, the mathematical basis for proposed detection operator and how to generate this kind of operator are described. In Section 3, the detailed procedure of object detection is presented. The experimental results are shown in Section 4, and the conclusions are given in Section 5.

2 Mathematical Basis and Generating Approach for Detection Operators

Some man-made objects are very clear from the high resolution satellite imagery, and these objects usually have typical shapes. So the shape of boundary is very important information. For usual shape detection approaches, the boundary must be detected firstly. Operators such as Roberts, Sobel, Prewitt, Canny are usually used for edge detection. However, there are some problems for usual shape detection. Firstly, the edge detection result using these typical operators is a kind of binary image where the gradual information of intensity is lost compared with the gray scale image. Secondly, the boundary after edge detection using these typical operators can not guarantee a close curve and the open curve can't describe the shape of objects perfectly. Therefore, an operator that can detect the intensity change of edge and describe the shape of object is required.

2.1 Mathematical Basis of the Detection Operator Based on BSI

As introduced in above paragraphs, an operator which can detect the intensity change of edge and describe the shape of object should be designed. Since the shape is contact with boundary, an optimal boundary detection operator is required firstly. The mathematical model of the object detection operator is introduced via finding an optimal frequency domain edge detection operator.

Take finding the optimal edge detection operator of 1-D signal for example. Let the range of an ideal step signal X(t) be α.

$$X(t) = \begin{cases} 0 & t < 0 \\ \alpha & t \geq 0 \end{cases} \tag{1}$$

Since the real signal is not ideal, let the signal after adding the noise N(t) be $\widetilde{X}(t) = X(t) + N(t)$. Let the autocorrelation function of noise be $R_N(t) = \mu * \delta(t)$, where $\delta(t)$ is the pulse function and μ is the weighted coefficient. So an optimal smoothing filter operator h is required to denoise.

It can be prove that Wiener filter is the optimal smoothing filter [10].

$$H(w) = \frac{S_X(w)}{S_X(w) + S_N(w)} \tag{2}$$

where $S_X(w)$ and $S_N(w)$ are the spectral densities of X and N.

Since the noise is a white noise and signal is a step signal, $S_X(w) = (\alpha^2/4)\delta(w) + \alpha^2/(4\pi^2 w^2)$ and $S_N(w) = \mu^2$. Through the equation (2), the optimal smoothing filter can be obtained by

$$H(w) = \frac{\frac{\alpha^2}{4}\delta(w) + \frac{\alpha^2}{4\pi^2 w^2}}{\frac{\alpha^2}{4}\delta(w) + \frac{\alpha^2}{4\pi^2 w^2} + \mu^2} = \frac{d^2}{d^2 + 4\pi^2 w^2} \quad (d = \frac{\alpha}{\mu}) \tag{3}$$

After anti-Fourier transform, the filter in spacial domain can be shown as

$$h(t) = \frac{d}{2}\exp(-d\,|t|) \tag{4}$$

where d is defined in equation (3). t is independent variable. From the probability distribution, it can be seen that the equation (4) pertains the double exponential distribution. So the optimal smoothing filter operator is the double exponential function.

Extend the 1-D smoothing filter operator to 2-D image. The 2-D operator for smoothing filter is defined as

$$\bar{h}(x, y) = h(\sqrt{x^2 + y^2}) \tag{5}$$

where *x, y* are the coordinates in image.

The optimal edge detection operator is merely the derivative of the smoothing operator [11]. The optimal edge detection operator is the gradient image for 2-D image.

2.2 The Generating Approach of the Detection Template based on BSI

Since the optimal edge detection operator has been found, the shape description function should be added. We take the vehicle object for example in this paper. The intensity change of the vehicle object edge is described by a function, and the edge is assumed to be a smooth simple successive boundary. Let *D* be the connected region of the vehicle including the boundary *C*. $\overline{D}$ is the region outside of *D*, and *C* is the vehicle boundary which is shown as Fig. 3(a). A level function *L* which is shown as follows should be found.

$$L(\mathbf{x}) = \begin{cases} \min_{\mathbf{z} \in C} \|\mathbf{x} - \mathbf{z}\| & \mathbf{x} \in D \\ -\min_{\mathbf{z} \in C} \|\mathbf{x} - \mathbf{z}\| & \mathbf{x} \in \overline{D} \end{cases} \tag{6}$$

where **x** is the coordinates of any point in the image, and **z** is the point in boundary. $\|\mathbf{x} - \mathbf{z}\|$ shows the Euclidean distance between **x** and **z**. So $L(\mathbf{x})$ is the minimum distance between point **x** and boundary C when **x** is inside of object region D. When **x** is outside of object region, $L(\mathbf{x})$ means the negative of the minimum distance between point **x** and boundary C. So the object detection operator $f(\mathbf{x})$ can be generated using the level function $L(\mathbf{x})$.

$$f(\mathbf{x}) = h'(L(\mathbf{x})\sigma) \tag{7}$$

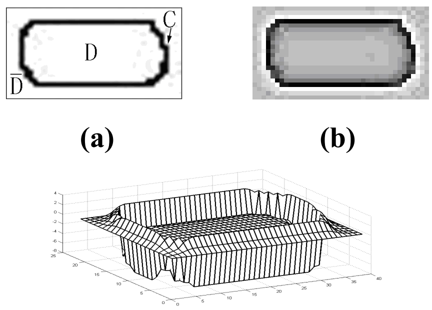
where *h* is the double exponential function, which is defined by $h(t) = (1/\sigma)\exp(-|t|/\sigma)$. σ is the coefficient of the double exponential function. $\sigma = 1/d$ and *d* is given in equation (3). Then the object detection operator $f(\mathbf{x})$ is obtained using equation (6) and equation (7).

$$f(\mathbf{x}) = \begin{cases} -\dfrac{1}{\sigma}\exp(-L(\mathbf{x}))L'(\mathbf{x}) & L(\mathbf{x}) \ge 0 \\ \dfrac{1}{\sigma}\exp(L(\mathbf{x}))L'(\mathbf{x}) & L(\mathbf{x}) < 0 \end{cases} \tag{8}$$

where $L'(\mathbf{x})$ is the gradient image when the function $L(\mathbf{x})$ is a 2-D image.

Fig. 1 shows a vehicle detection template and its 3D model. Fig. 1(a) is a simulated integral vehicle boundary. Fig. 1(b) is the vehicle detection template which can describe the vehicle shape. The detection template is generated by simulated boundary and equation (8). Fig. 1 (c) is the 3D model of this template.

As mentioned above, it is clear that a special object detection template can be generated by the equation (8) and the pre-set object boundary. How to detect the special object using the detection template will be discussed in section 3.3.



(a) (b)

(c)

Fig. 1 Vehicle detection template and its 3D model. (a) An integral vehicle boundary. (b) The optimal edge detection and shape description template for vehicle detection. (c) The 3D model of the vehicle detection template.

3 The Object Detection Based on BSI

The proposed object detection approach includes four steps: preprocessing, Region of Interest (ROI) extraction and feature enhancement, object extraction and post-processing. Fig. 2 shows the flow chart of the object detection approach.

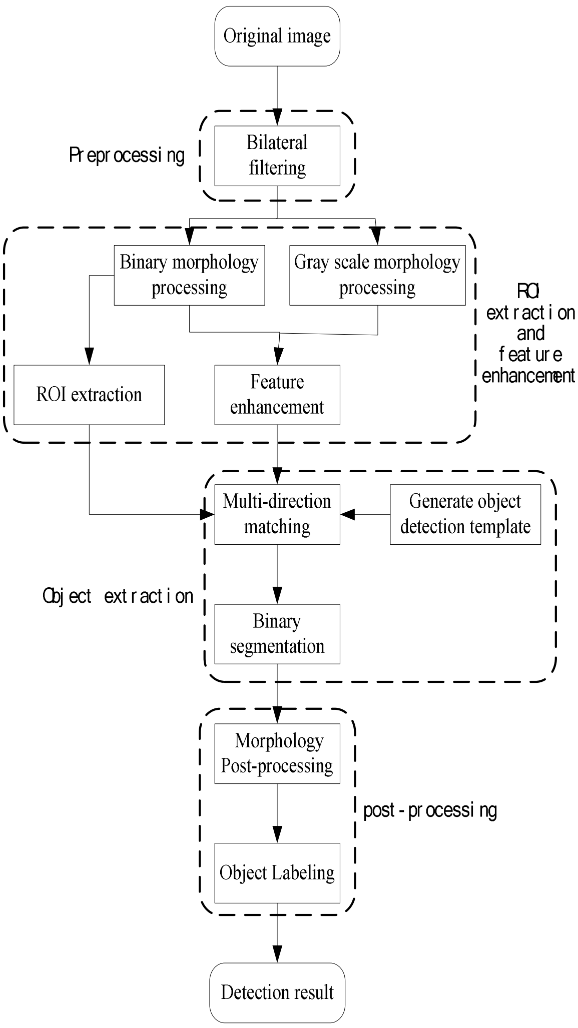


Fig. 2 The flow chart of the object detection approach based on BSI

3.1 Preprocessing

Image preprocessing aims to reduce background noise. In this paper, a bilateral filtering is used for preprocessing. Bilateral filtering converts the weights coefficients of Gauss filtering into the multiply results of Gauss function and image intensity, and the optimized weight coefficients are convolved with images [12,13]. As a result, not only noise is reduced but also useful edge information is kept. Fig. 4(a) shows an original image and Fig. 4(b) shows the result of bilateral filtering. It can be clearly seen that most road marks are removed and vehicle edges are not affected obviously. The preprocessing provides a good image for ROI extraction and feature enhancement.

3.2 ROI Extraction And Feature Enhancement

Although preprocessing removes some background noise, the low contrast between the objects and background or other details of objects will disturb the consistency between objects and the generated model. In order to remove this disturbance and reduce the searching area, a ROI extraction and feature enhancement approach which combines gray morphology and binary morphology is proposed. The detailed procedure is shown in Fig. 3 which includes three parts: gray morphological processing, binary morphological processing and ROI extraction. Gray morphological processing is to find the morphological gradient and binary morphological processing is to complement the details of object edge. The gray holes filling is implemented by the

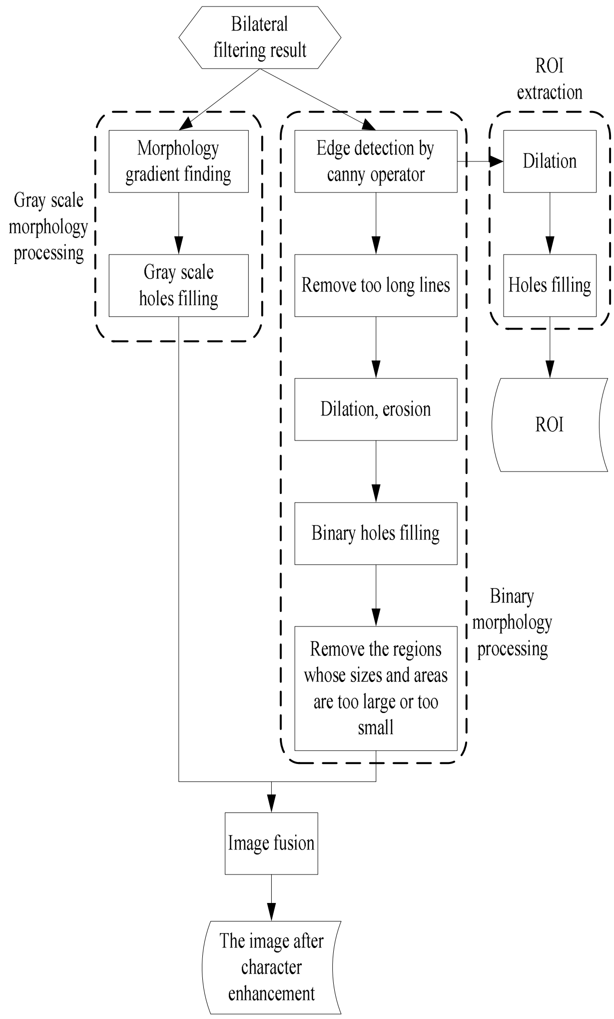
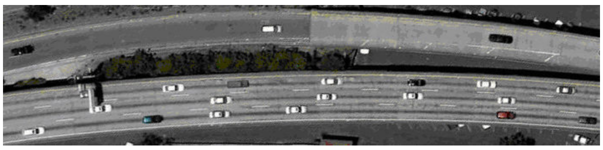


Fig. 3 The flow chart of ROI extraction and feature enhancement

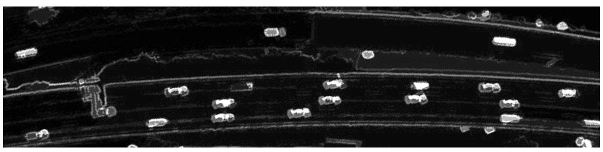
morphological reconstruction. The result after binary morphological processing is a combination of regions containing all most of the object pixels. Therefore, the shape information and boundary information are stood out by fusing the gray morphological processing and the binary morphological processing. In Fig. 3, the procedure 'Image fusion' is performed as following steps. If the value of the point in binary morphology processing result is true, the value of the same pixel in gray morphology processing result is multiplied by k which is the enhancement coefficient. In our experimental k is 3. After fusing the results of binary morphology processing and gray scale morphology processing, the shape feature of object is enhanced. Moreover, this algorithm reduces the noise derived from the details of object. A feature enhancement result is shown in Fig. 4 (c). In the image, vehicles all become white 'blocks' whose edges are much more obvious than before, and the edges of noise are weaken effectively.

Because the objects are in the places where there must be edges, edge detection is also implemented by Canny operator firstly. Then the morphological dilation operator is implemented and the holes are filled to acquire the ROI.



(a) Original image

(b) Bilateral filtering result



(c) Feature enhancement result

Fig. 4 The images of bilateral filtering result and feature enhancement result

3.3 Object Extraction

From the Fig. 4(c), it is obvious that objects after feature enhancement are very similar to the shape of the detection template generated in section 2. Therefore, the object detection is based on the image after feature enhancement. In this experiment, every template we generated has only one direction for the same kind of object. Because the objects in image may be toward any direction, several directions should be considered. An affine transformation for rotation with four directions is used in this paper. In our experiment, the rotated angles include 0, π/4, π/2 and 3π/4 [14].

The time complexity of usual matching algorithms is too great to apply in real detection work especially when the templates are too many or the size of the template is too large. So an approximate matching algorithm which is called integration weighted by distance based on gradient is presented in this paper.

A hypothetic shape boundary (Fig. 1.(a)) is required before generating the detection template. And then we integrate the gradient along the hypothetic shape boundary. Let p be a point on boundary C. Assume that C is close to the object, then p pick up the maximum gradient at $p0$, the closest point to p on the boundary.

$$\mathbf{p_0} = \{\mathbf{p} \mid \max_{\mathbf{p}\in C} L'(\mathbf{p})\} \tag{9}$$

where $L(\mathbf{x})$ is the level function which was introduced in section 2. So the matching response $R(\mathbf{x})$ of a point $\mathbf{x}$ in the detected image is computed as follows.

$$R(\mathbf{x}) = \int_{\mathbf{u}\in C} f(\mathbf{x}-\mathbf{u}) I(\mathbf{u}) \|\mathbf{u}-\mathbf{p_0}\| d\mathbf{u} \tag{10}$$

where f is the detection template, and I is the detected image. This algorithm only integrates the gradient along shape boundary, so the time complexity is much less than other matching algorithms.

Every detection template generated as the hypothetic shape should match with the detected image in all considered directions for all the pixels in ROI. The last matched result in every pixel is the best one of the 4 directional matched results. After obtaining the best matched result image, a threshold T is used to segment the matching result into a binary image. The threshold T is chosen manually according to the type of the object.

3.4 Post-processing and Object Labeling

The object pixels have been separated from the background after segmenting the matched result image. However, the noise is too much and the background is too complicated in some satellite images. In order to improve the detection result, the morphological operation combining the area information and aspect ratio information of object is used for the post- processing. After the pose-processing, the object regions are validated. These object regions are labeled using minimum bounding rectangle.

4 Experimental Results

Some images including roads, parking lots and airports from Google Earth are collected in our experiment. The detected objects contain vehicles and aircrafts. Table 1 shows the detection statistical results. We evaluate the result with the detection rate which is the ratio of the number of detected objects correctly to the number of real objects. The higher the detection rate the better the result. The detection results show that the approach presented in this paper can detect vehicles and aircrafts effectively.

4.1 Vehicle Detection Results

The vehicle detection is tested at two different resolutions. The test images are from all over the world which contains the area in city, countryside, desert and so on. The size of vehicles is 55*25 pixels in the first resolution images. The detection results are shown in Table 1 (Area1 to Area 10). These images are obtained at a low altitude, so the visual range of the images is narrow, which only contain one road or several roads. Fig. 5(a) shows a vehicle detection result of an area from USA where vehicles are very clear. Area 10 is an image from Wuhan in China. This image is not clear and it contains much noise. But most of vehicles in this image are detected correctly except 3 black vehicles whose contrast to the background is too weak.

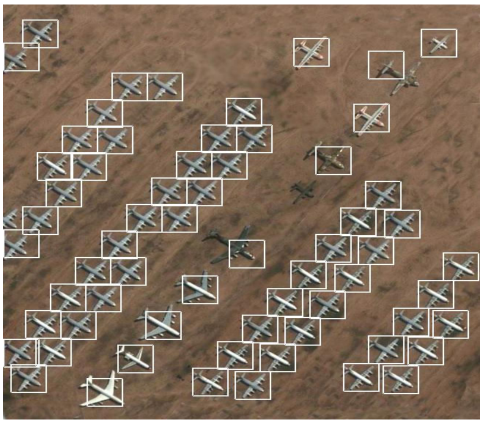
In second resolution images, the size of vehicle is 13*27 pixels. These images are obtained at a higher altitude, so the visual range of the images is larger than the first one (Area 11 to Area 16 in Table 1). The noise in these images is more serious and background is more complicated. The noise contains trees, houses, overpasses, traffic lines on road, shadow of some buildings and so on. Fig. 5(b) shows one of these detection results. Area 12 is an image from USA whose definition is high. Area 15 is a desert area from Iraq whose definition is much lower than Area 12. From the detection results, it is obvious that the approach presented in this paper is robust and can be used to detect the vehicles from different background images efficiently. Compared with these high resolution images, the detection results are not good enough since object details become fuzzy. The corner of house, non-green tree and traffic lines whose shapes are similar to rectangle lead to false detection, and the missing detections mostly happen when the vehicles are under the tree or covered by the shadow of buildings.

4.2 Aircraft Detection Results

A common feature of aircrafts on the military airport is that the shape and size of the similar aircrafts are the same and park together. Different type aircrafts park on different places. If there are aircrafts in different shapes or in sizes on an airport, several detection templates with different shapes and different sizes should be generated. Fig. 5(c)(d) show two aircraft detection results. Three different templates are generated to detect the aircrafts in Fig. 5(c), and two different templates are generated for detecting the aircrafts in Fig. 5(d). The number of

(a) Vehicle detection result 1

(b) Vehicle detection result 2



(c) Aircrafts detection result 1

(d) Aircrafts detection result 2

Fig. 5 Object detection results based on BSI

templates which should be generated depends on the kinds of aircrafts in the detected image. From these results, it is clear that this approach can be applied to detect different aircrafts on airport efficiently. Since the noise in airports is less than in other area, a very small number of missing detections happen when the contrast of the aircrafts to the background is too weak. And the false detection never happen in our experiment.

4.3 Comparison with Other Approaches

Table 1 shows the contrastive results of three object detection approaches for high resolution satellite imagery. As is shown in Table 1, the PCA approach is comparable with the MSNN approach in detection rate. The false detection numbers by MSNN approach are a bit more than that by the PCA approach. However, the proposed approach is much better than the two approaches mentioned above in detection results which include the missing detection numbers, false detection numbers and detection rates. The land area is very complicated and the noise in these images is very diverse, so the robustness advantage of BSI approach in detection result is obvious for vehicle detection. The aircraft detection results of the three approaches are comparable because the noise in airport images is much less than the noise in land area images. The PCA approach and the MSNN approach can get a detection rate above 80% only if there is not so much noise in the images. So the two approaches are not so effective for Area 10 and Area 12 which contain much noise. Whereas the BSI approach can reach a detection rate above 85% for every kinds of images including the two 'bad' images.

Compared with MSNN approach and PCA approach, the approach proposed in this paper has some other advantages. Firstly, this approach doesn't need any sample for training, which solves the problem of object detection with few samples in real detection work. Compared with the approach like MSNN and PCA which need a lot of training samples, this approach can reduce the time of

Table 1 The statistical results of object detection

Site	Object	No. of objects	No. of detected objects			No. of missing objects			No. of false objects			Detection rate %		
			PCA	MSNN	BSI	PCA	MSNN	BSI	PCA	MSNN	BSI	PCA	MSNN	BSI
Area 1	vehicle	24	29	26	25	3	4	2	8	6	3	87.5	83.3	91.7
Area 2	vehicle	16	15	15	16	2	3	0	1	2	0	87.5	81.3	100
Area 3	vehicle	18	19	19	18	5	4	1	6	5	1	72.2	77.8	94.4
Area 4	vehicle	22	19	18	22	5	5	1	2	1	1	77.3	77.3	95.5
Area 5	vehicle	33	32	33	33	3	4	1	2	4	1	90.9	87.9	97
Area 6	vehicle	20	20	20	20	0	0	0	0	0	0	100	100	100
Area 7	vehicle	21	22	23	20	2	3	1	3	5	0	90.5	85.7	95.2
Area 8	vehicle	36	37	39	34	5	6	3	6	9	1	86.1	83.3	91.7
Area 9	vehicle	24	26	24	25	0	2	1	2	2	2	100	91.7	95.8
Area 10	vehicle	22	14	19	20	8	6	3	0	3	1	63.6	72.7	90.9
Area 11	vehicle	25	30	31	27	4	5	3	9	11	5	84	80	88
Area 12	vehicle	81	90	94	87	17	19	11	26	32	17	79	76.5	86.4
Area 13	vehicle	54	60	62	54	6	7	4	12	15	4	88.9	87	92.6
Area 14	vehicle	70	83	81	78	8	10	3	21	21	11	88.6	85.7	95.7
Area 15	vehicle	27	31	34	28	4	4	2	8	11	3	85.2	85.2	92.6
Area 16	vehicle	55	57	59	54	8	10	6	10	14	5	85.5	81.8	89.1
Airport 1	aircraft	53	48	49	52	5	4	1	0	0	0	90.6	92.5	98.1
Airport 2	aircraft	65	60	59	63	6	7	2	1	1	0	90.8	89.2	96.9
Airport 3	aircraft	12	10	10	12	2	2	0	0	0	0	83.3	83.3	100
Airport 4	aircraft	17	16	18	17	2	1	0	1	2	0	88.2	94.1	100
Airport 5	aircraft	6	6	6	6	0	0	0	0	0	0	100	100	100

detection because it only needs to generate several templates according to the shape of objects. Secondly, many approaches for object detection are based on a condition that the region of interest must be extracted manually or by GIS. This approach can extract ROI automatically. Thirdly, this approach can be applied to other object detection besides vehicles and aircrafts since the approach is able to generate the detection templates according to detected object shapes.

5 Conclusions

An object detection approach is proposed in this paper. This approach is based on the boundary shape information of detected objects. According to the boundary of objects, the detection template of any shape is generated to detect this kind of objects from high resolution satellite imagery. In this paper, a bilateral filtering is used to reduce the noise in an image. Moreover, a ROI extraction approach and a feature enhancement algorithm are proposed to improve the detection results. In addition, proposed matching algorithm is so easy to be implemented by hardware. The experimental results show that this approach can be well applied to not only vehicle objects and aircraft objects, but also other objects which have a closed boundary. Further work would mainly focus on detecting the objects in more complicated background and reducing the time of detection.

Acknowledges

This work was supported by the National Natural Science Foundation of China (40571102) and Self-regulated Science and Research Foundation of Wuhan University (6081001).

References:

[1] W. Wu, D. Yin, R. Zhang, Y. Liu and J. Pan, Bridge recognition of median-resolution SAR images using pun histogram entropy, *Chinese Optics. Letters*. Vol.7, No.7, 2009, pp. 572-575.

[2] N. Di, M. Zhu and Y. Wang, Real-time detection of airport runway by extracting line feature, *Optics and Precision Engineering (in Chinese)*. Vol.17, No.9, 2009, pp. 2336-2341.

[3] T. Zhao and R. Nevatia, Car Detection in Low Resolution Aerial Image, *in Proceedings of International Conf. Computer Vision.* Vancouver, Canada, 2001, pp.710-717.

[4] H.Moon and R.Chellappa, A.Rosenfeld, Performance analysis of a simple vehicle detection algorithm, *Image and Vision Computing*. Vol.20, 2002, pp.1-13.

[5] S. Hinz and A. Baumgarter, Vehicle Detection in Aerial Images Using Genetic Feature, Grouping, and Context, *DAGM 2001, LNCS*, Vol.2191, 2001, pp.45-52.

[6] X. Perrotton, M. Sturzel and M. Roux, Automatic Object Detection on Aerial Images Using Local Descriptors and Image Synthesis, *ICVS 2008, LNCS 2008*, pp.302-311.

[7] L. Gong, Vehicle Detection Using Principal Component Analysis, *Computer and Robert Vision, Haralick, Shapiro*, 1993.

[8] H. Zheng and Y. Yu, Vehicle Detection from High Resolution Satellite Imagery Based on the Morphological Neural Network, *Journal of Harbin Engineering University (in Chinese)*, Vol.27, Suppl, 2006, pp.189-191.

[9] H. Zheng, X. Hu, X. Si and W. Yang, A Novel Object Detection Approach for Satellite Imagery Based on Danger Theory, *in Proceedings of ICINIS 2008*, pp.445-448.

[10] G. R. Cooper and C. D. McGillem, *Probabilistic Methods of Signal and System Analysis*. Oxford, U. K, Oxford Univ. Press 1999.

[11] H. Moon and A. Rosenfeld, Optimal Edge-Based Shape Detection, *IEEE Transactions on Image Processing.* Vol.11, No.11, 2002, pp.1209-1226.

[12] M. Jin and J. Song, A Self-adjust Approach for Bilateral Filtering, *Opto-electronic Engineering (in Chinese)*. Vol31, No.7, 2004, pp.65-67.

[13] G. Chen, K. Yang, R. Chen and Z. Xie, A gray-natural logarithm ratio bilateral filtering method for image processing, *Chinese Optics. Letters*. Vol.6, No.9, 2008, pp.648-650.

[14] R. C. Gonzalez, R. E. Woods and S. L. Eddins, *Digital Image Processing Using MATLAB* Publishing House of Electronics Industry, Beijing, 2005.