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 reduce noise, and a new hybrid morphological approach is proposed for ROI extraction and feature enhancement. A detection operator based on the boundary shape information (BSI) is developed to detect preprocessed objects. The images from Google Earth are tested in the paper. Comparisons with popular object detection approaches are also discussed. The experimental results show that proposed approach is effective and feasible.

Key-Words: - BSI, Detection template, Object detection, High resolution satellite imagery, Vehicles, Aircrafts

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 small object detection emerged. However, satellite image data is very complex, and it costs too much time to recognize objects manually. So the automation and intelligence of target recognition are required.

Up to now, some studies about automatic target recognition or detection have been done [1~4]. For satellite images, most of study is about large objects detection. Some study for the small object detection is limited to aerial images. These approaches include Bayesian Network, the derivative of Gaussian model, object based on 3D model, local operator and image fusion approach [5~8] etc. Few studies are for small object detection based on high resolution satellite image. The existing methods mainly focus on Principal Component Analysis (PCA), Morphological Shared-Weight Neural Network (MSNN) and Danger Perceptive Network (DPN) [9~11]. 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, sometimes it is hard to obtain satisfactory samples in a practical work. 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 proposed 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 detection 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

Since many man-made objects usually have typical shapes, the shape of object boundary is very important information for detection. Most traditional shape detection approaches usually detect boundary via edge detection. However, the edge detection result is a kind of binary image where the gradual information of intensity is lost. Therefore, it is necessary to design the operators which can keep the shape and gradual information of object contours.

2.1 Mathematical Basis of the Detection Operator Based on BSI

The mathematical model of the object detection operator is based on 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 with the noise N(t) be $\overline{X}(t) = X(t) + N(t)$. 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 reduce noise.

It can be prove that Wiener filter is the optimal smoothing filter [12].

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

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

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

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

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

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.

The 1-D smoothing filter operator is extended to 2-D operator for smoothing filter. It is defined as

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

where x, y are the coordinates in image.

The optimal edge detection operator is merely the derivative of the smoothing operator [13]. The optimal edge detection operator is the gradient image for 2-D image.

2.2 The Generating Approach for the Detection Template Based on BSI

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

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

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

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

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

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

where $L'(\mathbf{x})$ is the gradient image of $L(\mathbf{x})$ which means a 2D 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 describing the vehicle shape, which is generated by simulated boundary and equation (8). Fig. 1 (c) is the 3D model of this template.

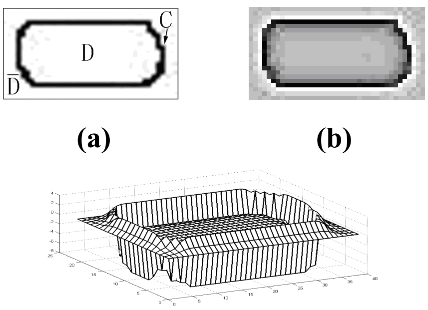


Fig. 1 Vehicle detection template and its 3D model. (a) An integral vehicle boundary. (b) The optimal edge detection and shape description template for vehicle detection. (c) The 3D model of the vehicle detection template.

As mentioned above, 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 generated detection template will be discussed in section 3.3.

3 The Object Detection Based on BSI

The proposed object detection approach consists of 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.

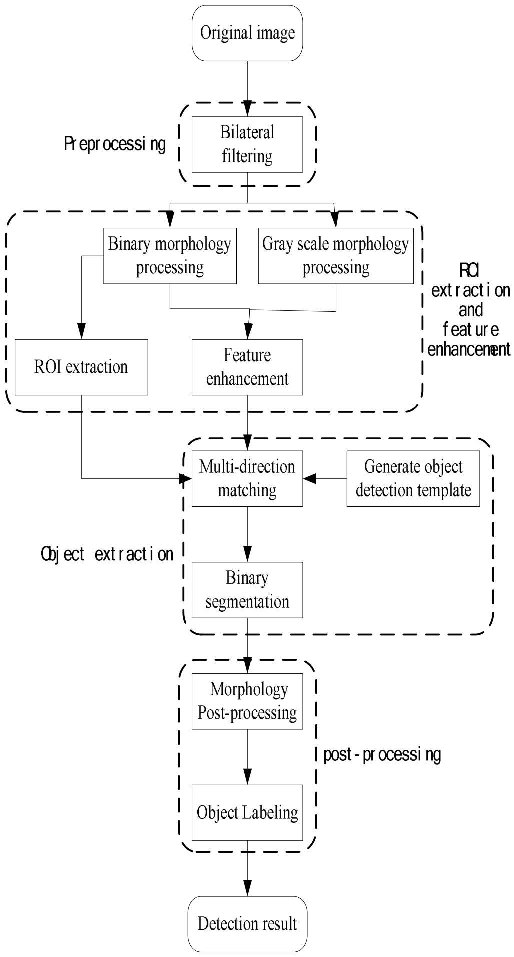


Fig. 2 The flow chart of the object detection approach based on BSI

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 [14,15]. As a result, not only noise is reduced but also useful edge information is kept. Fig. 4 shows an original image and its bilateral filtering results. It can be clearly seen that most road marks are removed and vehicle edges are not blurred obviously. The preprocessed image is good 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 consists of 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 algorithm is implemented by the morphological reconstruction. The result after binary morphological processing is the combination of regions containing most of the object pixels. Therefore, the shape and boundary information are stood out by fusing the gray morphological processing and the binary morphological processing. In Fig. 3, the image fusion is performed as following pseudo code.

```
if A(i,j) == 255
    B(i,j) = B(i,j) * 3;
else
    B(i,j) = B(i,j);
end
```

where $A(i,j)$ means the value of the binary morphological processing result image at the point

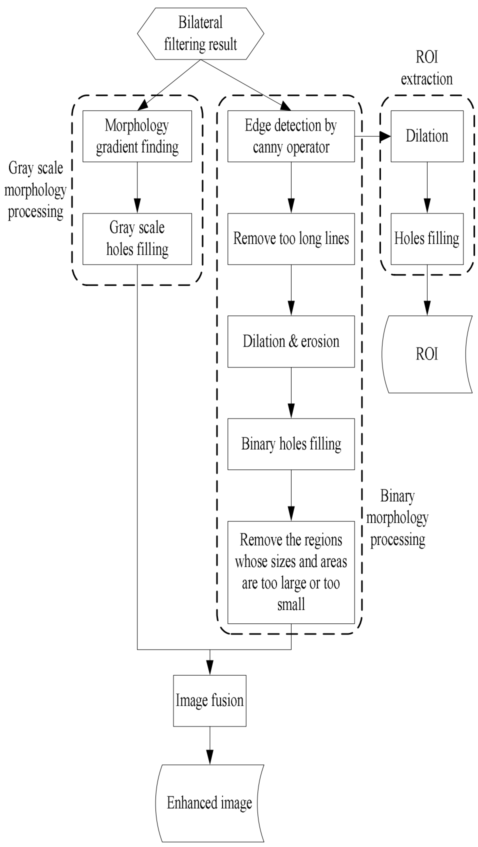
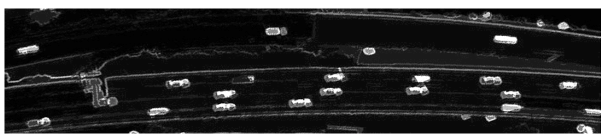


Fig. 3 The flow chart of ROI extraction and feature enhancement

(a) Original image

(b) Bilateral filtering results



(c) Feature enhancement results

Fig. 4 The example images of the bilateral filtering results and feature enhancement results.

(i,j), and *B*(i,j) means the intensity of the gray morphological processing result at the point (i,j).

After fusing the results of binary and gray scale morphology processing, the shape feature of object is enhanced. Moreover, this algorithm reduces the noise derived from the details of object. An example for feature enhancement result is shown in Fig. 4 (c). In the example, all vehicles become white 'blocks' whose edges are much more obvious than before, and the edges of noise are weaken effectively.

Since the objects locate in the places where there must be edges, edge detection is implemented, and the morphological dilation operator is used to fill holes. As a result, ROI is obtained.

3.3 Object Extraction

From the Fig. 4(c), it is obvious that enhanced objects are similar to the shape of the detection template generated in section 2. Therefore, the object detection runs on the enhanced image. In our study, every generated has only one direction for the same kind of object, but the objects 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 study, we consider four angles, i.e. 0, π/4, π/2 and 3π/4 [16].

The time complexity of usual matching algorithms is too great to apply in real detection work, especially for many or very large templates. So an approximate matching algorithm which is called integration weighted based on gradient distance 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 $p0$, which is the closest point to p on the boundary.

$$\mathbf{p_0} = \{\mathbf{p} \mid \max\nolimits_{\mathbf{p}\in C} L'(\mathbf{p})\} \qquad (9)$$

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

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

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

Every generated detection template should be matched with the detected ROI image. The last matched result at every pixel is the maximum of the 4 directional matched results. After obtaining the maximal matched result, 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. In order to remove noise from extracted object pixels, the morphological operation combining the area information and aspect ratio information of object is used for the post-processing. After the post-processing, the object regions are validated, and these 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 study. The detected objects contain vehicles and aircrafts. Table 1 shows the statistical detection results. The detection rate is the ratio of the number of detected objects correctly to the number of real objects. The greater the detection rates the better the results. 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 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. Among 10 areas, 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 the second resolution images (Area 11 to Area 16 in Table 1), 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. 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. Among listed areas, 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 collected high resolution images, the detection results are not good enough since object details are 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 trees or covered by the shadow of buildings.

4.2 Aircraft Detection Results

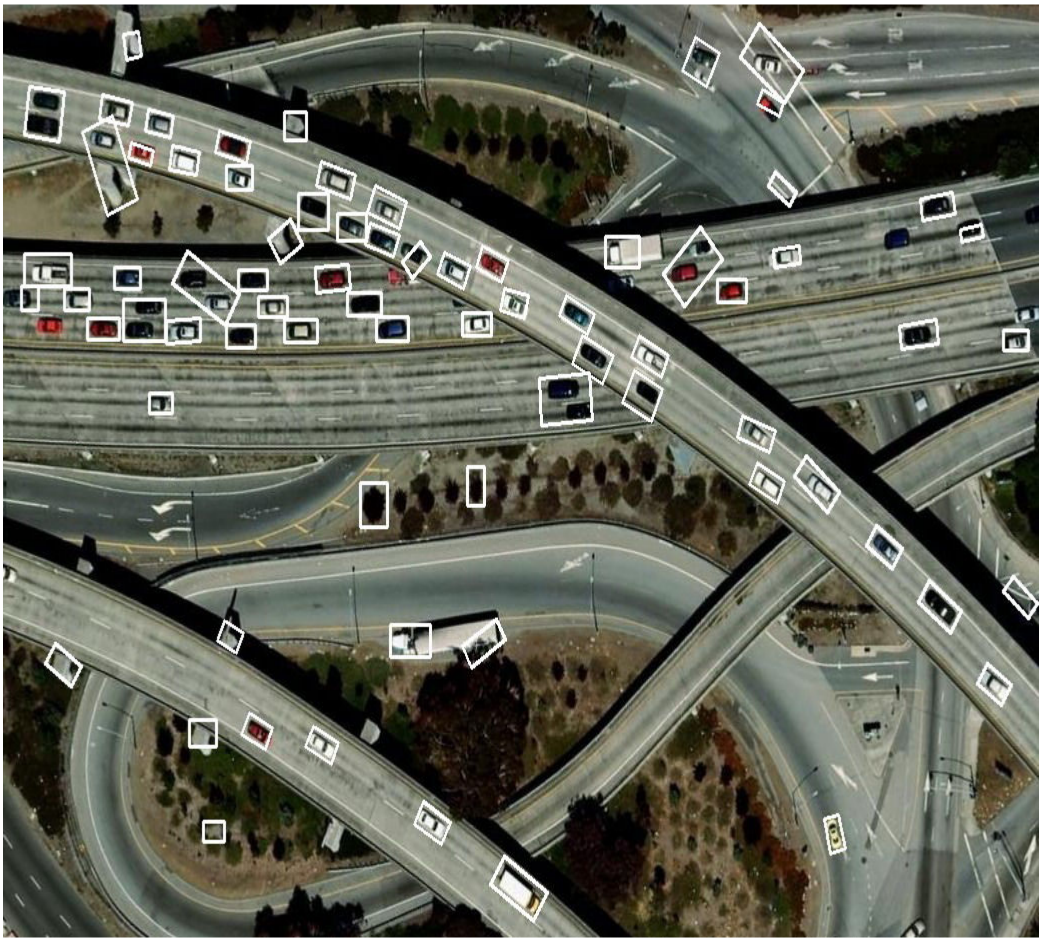
A common feature of aircrafts on the military airport is that the shape and size of the similar aircrafts are the same. In order to detect aircrafts with different shapes and sizes, several detection templates with different shapes and sizes should be generated. Fig. 5(c)(d) show two aircraft detection 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 generated templates depends on the kinds of aircrafts in the detected image. From detection results, it can be seen that proposed approach can be employed to detect different aircrafts efficiently. Since the noise in airports is less than in other area, a few of missing detections happen when the contrast of the aircrafts to the background is too weak. And the false detection never happen in our experiments.

4.3 Comparison with Other Approaches

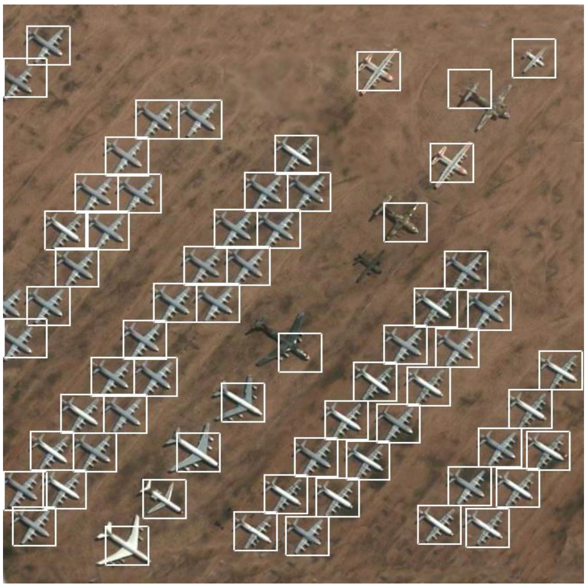
Table 1 shows the comparison results of three object detection approaches for high resolution satellite imagery. As is shown in Table 1, the PCA approach is comparable with the MSNN approach in detection rate. The false detection numbers by MSNN approach are a bit more than that by the PCA approach. However, the proposed approach is much better than above two approaches in the missing detection numbers, false detection numbers and detection rates. 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% when there is not so much noise in the images. But the two approaches are not so effective for Area 10 and Area 12 which contain much noise. But the BSI approach can reach a detection rate above 85% for every kind of images.

Compared with MSNN approach and PCA approach, the approach proposed in this paper doesn't need any sample for training, which solves the problem of object detection with few samples in real detection work. In contrast, MSNN and PCA need lots of training samples. In addition, since the

(a) Vehicles detection result 1



(b) Vehicles detection result 1



(c) **Aircrafts detection result 1**

(d) Aircrafts detection result 2

Fig. 5 Object detection results based on BSI

proposed approach detect objects via template matching, it is easier to be implemented by hardware.

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 also 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 easy to be implemented by hardware. The experimental results show that proposed 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 from more complicated background and reducing the detection time.

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

Acknowledges

This work was supported by the National Natural Science Foundation of China (40571102) and supported by the Fundamental Research Funds for the Central Universities (6081001).

References:

[1] V. P. de Araujo, R. D. Maia, M. F. S. V. D'Angelo and G. N. R. D'Angelo, Automatic Plate Detection Using Genetic Algorithm, *Proceedings of the 6th WSEAS International Conference on Signal, Speech and Image Processing*, Lisbon, Portugal, 2006, pp.43-48.

[2] S. Makdee, C. Kimpan and S. Pansang, Human face recognition using volume feature, fuzzyc-means and membership matching score, *Proceedings of the 5th WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics*, Venice, Italy, 2006, pp.302-307.

[3] J. Janta, P. Kumsawat, K. Attakitmongkol and A. Srikaew, A pedestrian detection system using applied log-Gabor, *Proceedings of the 7th WSEAS International Conference on Signal, Speech and Image Processing*, Beijing, China, 2007, pp.55-60.

[4] I. Parra, D. Femandez, M. A. Sotelo, P. Revenga, L. M. Bergasa, M. Ocana, J Nuevo and R. Flores, Pedestrian recognition in road sequences, *Proceedings of the 5th WSEAS International Conference on Signal Processing, Robotics and Automation*, Madrid, Spain, 2006, pp.273-278.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.

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

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

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

[15] 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.

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