Visual saliency mechanism-based object recognition with high-resolution remote-sensing images

Lin He¹, Chen Li²

¹Beijing Institute of Space Mechanics and Electricity, Beijing, People’s Republic of China
²Beijing Engineering Technology Research Center of Aerial intelligent Remote Sensing Equipments, Beijing, People’s Republic of China

Abstract: Object recognition with remote-sensing image is widely used in many areas. Some objects are smaller and denser in the high-resolution images, such as the oil tank, ship, and aircraft. The recognition of this kind of objects is more difficult than the objects with low-resolution images, for example, bridges and airports. The recognition performance is more dependent on the shallower features. The contour of these objects is obvious, and the characteristics are quite different from the background, which satisfies the human visual saliency mechanism. Here, the authors propose a novel theme of object recognition method based on visual saliency mechanism for remote-sensing images with sub-meter resolution. The experimental results show that the proposed method performs best compared with other algorithms.

1 Introduction

Object recognition is applied to many areas. The existing traditional algorithms of remote-sensing object recognition and detection adopt the strategy of rough detection combined with precision detection. First, the object region is extracted from the input image, that is, the region of interest (ROI) is extracted from the original image to quickly extract the region of interest. As the general framework, object recognition and detection are carried out more accurately, and the false detection regions are removed, and the final object to be detected is found and confirmed [1]. For the extraction of candidate regions in the original image, scholars have proposed many schemes [2, 3]. They proposed a threshold segmentation method which combines the grey information and the edge information of the image. Others get the sparse description feature of the image by a multi-layer sparse coding method [4, 5], and calculate its saliency value according to the sparse feature of the image and segment the image according to the saliency value to get the candidate object area of the image. For the process of object recognition and validation after each candidate region is obtained, most researchers use a variety of feature descriptors to describe the detected object, and then use support vector machine or Adaboost classifier to further confirm. For example, in the literature [6], the author extracts the texture feature of the image and combines the shape feature of the object, and obtains a high-dimensional feature descriptor to describe the ship object. Similarly, in the previous work [7, 8], researchers proposed a local binary image model as the image object feature, and a deformable component model is also proposed in reference [9] to describe the image object feature. References [10, 11, 12] propose object recognition and detection techniques using feature analysis and meanshift algorithm. These algorithms are all for some specific close-range objects and are not involved in object recognition of aircraft, oil tanks and ships for remote sensing objects. Therefore, it is necessary to make a customised analysis of the characteristics of these objects.

On the basis of analysing the particularity of remote-sensing object recognition under high resolution, this paper designs an object extraction method based on ROI region of visual saliency mechanism for remote-sensing image objects with sub-meter resolution <1 m. Meanshift carries out image segmentation combined with saliency detection and feature extraction, and makes use of machine. The object recognition scheme of the learning algorithm is studied, and the aircraft, oil tank, and ship objects are tested and analysed comprehensively.

2 Framework

In this paper, for the object recognition task with high resolution, we also adopt the mode of coarse detection and precision detection. For this mode, the detection process can be described in two parts. The first part is the rapid acquisition of candidate regions and the second part is the description and recognition of the object. For high-resolution objects such as ships in this paper, taking the ship monitoring process as an example, the process of object recognition in high-resolution remote-sensing image using traditional machine learning algorithm is illustrated in Fig. 1.

The main step is to compute the saliency map of the input image using the improved Fourier transform saliency algorithm. Meanshift algorithm is used to segment the original image and merge the fragmented small areas in the original image. The basic shape features of the object are used to screen the following ROI regions, and the specific region extraction method will be described in detail in the following section. For the collected object sample files, the sample data set is made up in the ratio of positive and negative samples 1:3 and the sample description file is generated. Then, for the sample description file, the appropriate features are designed. To extract and describe the object features, the feature descriptors are sent to the support vector machine (SVM) classifier for learning, and the support vectors for subsequent recognition are output.

Finally, the extracted ROI regions are extracted and described, which are fed into the SVM detector and output the final recognition results.

3 Procedure

3.1 Step 1: ROI region extraction based on visual saliency mechanism

For the object with high resolution, the size of the object is smaller than the whole input image, and the background area of the object is more in the image. Therefore, it is necessary to locate the possible region of the object quickly, eliminate redundant information and extract the region with high similarity to the object. One of the most common research ideas is to extract regions by visual saliency. The ROI extraction algorithm based on visual saliency mechanism draws on the human visual selective attention mechanism and aggregates the pixels with significant local optical features into the region of interest. As shown in Fig. 2, this method mainly uses the colour and brightness information of
the image in Lab space. In this method, for an input image, first, the input image is filtered by using the Gauss filtering kernel to remove part of the noise as follows.

\[ S(x,y) = \| I_\mu - I_{\omega hc}(x,y) \| \]  

(1)

where \( S(x,y) \) is the computation value of the image with position \((x,y)\), and \( I_\mu \) is the mean value of channels through image filter as

\[ I_\mu = \begin{bmatrix} L_\mu \\ a_\mu \\ b_\mu \end{bmatrix} \]  

(2)

\[ I_{\omega hc}(x,y) = \begin{bmatrix} L_{\omega hc} \\ a_{\omega hc} \\ b_{\omega hc} \end{bmatrix} \]  

(3)

In this paper, saliency detection algorithm based on image colour features is used to calculate input image saliency map. Based on the visual saliency map, the ROI region in high-resolution remote-sensing object recognition is extracted in this paper. The contrast between the saliency map obtained from this saliency detection method and the original image is shown in Fig. 3.
The meanshift in the flow chart in Fig. 3 is essentially a clustering algorithm, and the meanshift algorithm is a non-parametric clustering algorithm for feature space, whose calculation method is essentially dependent on probability density estimation. For the high-resolution remote-sensing image in this paper, the result of meanshift segmentation is shown in Fig. 4. The same colour in the image represents the same region. It can be seen that meanshift can distinguish the object from the background for the data set used in this paper. Although there are many small fragmentary regions in the detection results, it can well remove large areas of background such as water.

At this point, most of the background in the reserved ROI region has been removed, because there are many small areas around the object, so in order to merge the object region into a region, the adjacent region is merged according to its bounding rectangular box. After merging, the final ROI region is extracted. After the above process, this paper completes the extraction of ROI region in the process of high-resolution remote-sensing image object recognition. Finally, the object ROI region extraction results of high-resolution remote-sensing images are shown in Fig. 5.

3.2 Step 2: description and recognition of object

After obtaining ROI candidate regions, the main task is to select the appropriate feature descriptors of ROI regions and objects, and use the classifier to get a segmentation plane between the object and negative samples. In this paper, the high-resolution remote-sensing objects are mainly small and many clumped objects, such as ships, oil tanks, aircraft etc. The difficulty of recognition is that the aircraft, ships, and other objects have different models and different scales. The difference between the data set and the angle of the same object is also larger. However, compared with the surrounding background, the research object with high resolution has more closed contour and prominent shape features. Therefore, in designing object feature description vectors, it is necessary to highlight the contour and shape of the object, and the extracted features have certain scale and rotation invariance.

In this paper, the shape features (aspect ratio, area size, proportion in the image), Hu invariant moments, and hog gradient histograms are used to identify large and medium-sized objects in high-resolution remote-sensing images. Support vector machine is a widely used supervised binary classification method based on statistical learning in image classification. The training process is to find the best segmentation of hyperplanes in two positive and negative samples.

4 Experimental results

In order to test the performance of the proposed algorithm, we collect the datasets for the application, including the oil tank, aircraft, ships, and other objects. For the training set and the test set, we randomly use the 20% of the source dataset as the test set, and the test results on the test set show the object recognition accuracy. The lowest recognition accuracy satisfies the need 85% of application, and the highest recognition accuracy is 92%. The part of results is shown in Figs. 6–9. The statistical results are shown in Table 1. (Figs. 10 and 11)
As shown in Table 1, we have more comparisons with other algorithms in the same conditions. As the databases are collected for the practical applications, influenced by background, objects, and illuminations, the recognition is more difficult than other popular open databases. So the recognition rate is only 85%, whereas the performances will be increased with the increasing of training samples number.

5 Conclusion

For the problems of remote-sensing object recognition under high resolution, this paper proposes the object recognition method based on ROI region extraction of visual saliency mechanism. The object recognition scheme using machine learning algorithm is proposed, and the recognition accuracy rate of the aircraft, oil tank, and ship objects is evaluated for the comparison. The lowest recognition accuracy satisfies the need 85% of application, and the highest recognition accuracy is 92%. The part of results is shown in Fig. 6–9. The statistical results are shown in Table 1. This method can be applied to other similar object recognition problems. The proposed methods are used to other image and video recognition systems.

| Table 1 | Statistical results of recognition accuracy |
|---------|-----------------------------------------|
| Algorithms | Oil tank, % | Aircraft, % | Ships, % | Other, % |
| Gabor + NN | 87 | 81 | 84 | 80 |
| wavelet + SVM | 86 | 80 | 78 | 79 |
| LOG + SVM | 86 | 79 | 81 | 81 |
| our method | 92 | 85 | 89 | 85 |

6 References

[1] Zhu, C., Zhou, H., Wang, R.: ‘A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features’, IEEE Trans. Geosci. Remote Sens., 2010, 48, pp. 3446–3456
[2] Li, Z., Yang, D., Chen, Z.: ‘Multi-layer sparse coding based ship detection for remote sensing images’, Proc. Int. Conf. of Information Reuse & Integration, California, USA, 2015, pp. 122–125
[3] Song, Z., Sui, H., Wang, Y.: ‘Automatic ship detection for optical satellite images based on visual attention model and LBP’, Proc. Int. Conf. of Electronics, Computer and Applications, Ottawa, Canada, 2014, pp. 1–5
[4] Cheng, M.M., Zhang, G.X., Mitra, N.J., et al.: ‘Global contrast based salient region detection’. Proc. Int. Conf. of Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, June 2011, pp. 409–416
[5] Li, Y., Xu, J.W., Zhao, J.F., et al.: ‘An improved mean shift segmentation method of high-resolution remote sensing image based on LBP and canny features’, Appl. Mech. Mater., 2015, 713–715, pp. 1589–1592
[6] Li, H., Gu, H., Han, Y., et al.: ‘Object-oriented classification of high-resolution remote sensing imagery based on an improved colour structure code and a support vector machine’, Int. J. Remote Sens., 2010, 31, (6), pp. 18–25
[7] Barzohar, M., Cooper, D.B.: ‘Automatic finding of main roads in aerial images by using geometric-stochastic models and estimation’, IEEE Trans. Pattern Anal. Mach. Intell., 1996, 18, pp. 707–721
[8] Zheng, P., Wang, R.: ‘A multiple conditional random fields ensemble model for urban area detection in remote sensing optical images’, IEEE Trans. Geosci. Remote Sens., 2007, 45, pp. 3978–3988
[9] Zamir, A.R., Shah, M.: ‘Image geo-localization based on multiple nearest neighbor feature matching using generalized graphs’, IEEE Trans. Pattern Anal. Mach. Intell., 2014, 36, pp. 1546–1558
[10] Elmqvist, B., Ard, J., Olsson, L.: ‘Land use studies in drylands: an evaluation of object – oriented classification of very high resolution panchromatic imagery’, Int. J. Remote Sens., 2008, 29, (24), pp. 7129–7140
[11] Su, W., Li, J., Chen, Y., et al.: ‘Textural and local spatial statistics for the object – oriented classification of urban areas using high resolution imagery’, Int. J. Remote Sens., 2008, 29, (11), pp. 3105–3117
[12] Yang, F., Xu, Q., Guo, F., et al.: ‘Ship detection from optical satellite images based on visual search mechanism’, Proc. Int. Conf. of Geoscience & Remote Sensing Symp., Milan, Italy, May 2015, pp. 3679–3682