

The Target Classification of Optical Remote Sensing Image Based on Hierarchical Features and AdaBoost Algorithm

Xiaofei Ji, Ningli Qin and Yangyang Wang

College of Automation, Shenyang Aerospace University, Shenyang, 110136, China
jixiaofei7804@126.com

Abstract

In order to accurately classify and recognize specific targets in remote sensing images, a novel recognition method based on hierarchical features and AdaBoost algorithm is proposed. Hierarchical features can represent the distribution characteristics of targets effectively by applying the image pyramid idea to the BoF-SIFT feature extraction. The AdaBoost algorithm with support vector machine(SVM) as the weak classifiers is used for targets classification. The proposed classification and recognition method is tested on our remote sensing image dataset. The results verified the method can achieve higher recognition accuracy.

Keywords: remote sensing image, BoF-SIFT feature, hierarchical features, SVM, AdaBoost

1. Introduction

With the continuous development of remote sensing technology and pattern recognition technology, the target classification and recognition of remote sensing image, especially optical remote sensing images, has become the most important direction of remote sensing image processing and analysis field gradually. The research plays an extreme important role in the military, topography and so on [1].

The target classification and recognition of optical remote sensing image includes two steps: feature extraction and description, classifier training. The feature will directly affect target recognition accuracy. Yang [2] proposed a texture classification algorithm based on the SOFM(self-organizing feature map) network model. The algorithm extracted the local texture spectrum feature of remote sensing image to characterize the overall image gray distribution. Chen [3] proposed a new feature extraction method of circular target. It preserves many advantages, such as lower consumption, higher detection rate and stronger anti-noise ability. Liu [4] proposed an aircraft recognition method based on coarse-to-fine. It extracted Sobel edge feature and weighted the influence on the pixel position. Many experiences show that the proposed method achieved better recognition accuracy. Because the methods mentioned above only use single feature in the recognition process, the recognition rate is limited. The multi-features fusion is an effective solution to recognize the multiple class targets accurately. Hsieh [5] combined outline, wavelet coefficients and original bitmap feature to recognize aircraft type. Zhu [6] proposed a ship recognition method based on the feature fusion of shape and texture in the optical remote sensing images. The proposed method achieved a better recognition rate. However the rotational robustness of these features is relatively weaker. Furthermore the distribution characteristics representation of target feature points is not adequate.

According to the problems mentioned above, a novel remote sensing image target classification and recognition method based on hierarchical features and AdaBoost algorithm is proposed. Hierarchical features combined the hierarchical thought of image pyramid and BoF-SIFT feature. AdaBoost algorithm is achieved by cascading a number

of SVM weak classifiers to build a strong classifier. The proposed algorithm diagram is shown in Figure 1.

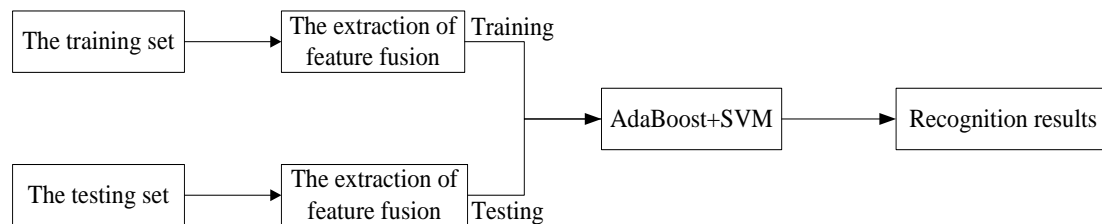


Figure 1. The Diagram of the Proposed Method

2. BoF-SIFT Algorithm Description

SIFT(Scale-invariant feature transform) feature not only overcomes the scaling and brightness invariance which many other features don't meet, but also has rotational robustness. BoF-SIFT feature can intuitively describe the similarity among the same targets and the difference among the different targets. So it plays a key role in multi-targets classification and recognition of optical remote sensing image.

2.1. SIFT Feature Extraction

SIFT feature, a local feature descriptor, was proposed by David Lowe in 1999 [7] and made further development and improvement in 2004 [8]. The descriptors of modulus value and direction are computed in order to match corresponding feature points between different images. SIFT feature matching algorithm can solve the matching problem between two images with translation, rotation and affine transformation. It has relatively strong matching ability in contrast to other nine kinds of local descriptors [9]. Overall, SIFT descriptor have the following advantages: local feature representation, rich information, multi-quantity, relatively fast speed, strong scalability.

Therefore SIFT feature is used as basic feature in this paper. Features extraction of an image is divided into two steps. One is feature points detection and the other one is feature points SIFT descriptor generation. SIFT feature extraction process is shown in Figure 2.

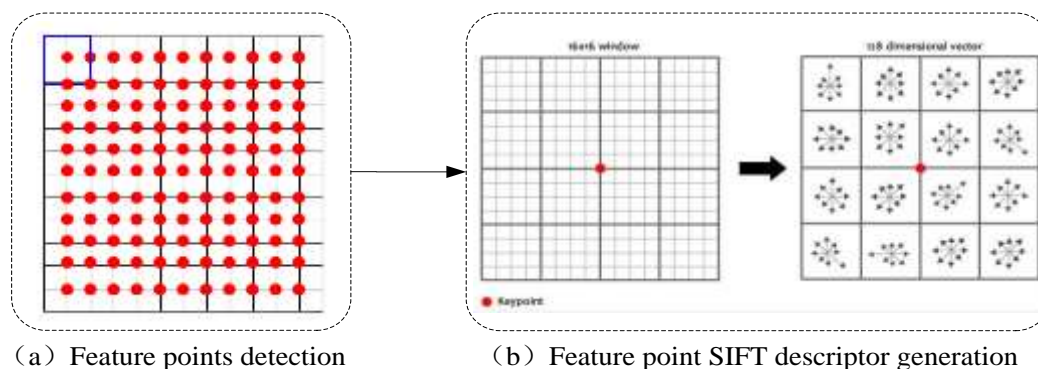


Figure 2. SIFT Feature Extraction Process

Feature points detection: As is shown in Figure 2 (a), the input image is divided into n blocks, corresponding to n feature points. SIFT features are extracted for every feature points respectively.

Feature point SIFT descriptor generation: As is shown in Figure 2(b), SIFT descriptor generation need to take sample for the 16×16 neighborhood window whose center is feature point. The window is divided into 4×4 sub regions as 16 seed

points. The gradient direction of 16 seed points are counted in eight directions by statistic cumulative value of gradient direction histogram. The main neighborhood gradient direction of the feature point is represented by the peak of histogram to ensure rotational robustness. The feature vector is $4*4*8=128$ dimensions. It is an abstract representation of the image information. Gradient value and direction of each pixel around each feature point is defined as:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (1)$$

$$L(x, y) = \arctan \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (2)$$

Here (x, y) is coordinate of feature points. $L(x, y)$ is Gauss convolution of the image. The scale of L is the scale of feature point. The position, scale and orientation information of each feature can determine a SIFT feature area.

2.2. Generate BoF-SIFT Descriptor

BoF (Bag of Features) [10] is one of the popular visual descriptors used for visual data classification. BoF is inspired by a concept called Bag of Words. A bag of words is a sparse vector of occurrence counts of words; that is, a sparse histogram over the vocabulary. In computer vision, a bag of visual words of features is a sparse vector of occurrence counts of a vocabulary of local image features.

The first step of BoF-SIFT descriptor generation is: build the codebook of SIFT features. K-means++ algorithm [11] is used to cluster the set of feature descriptors to form the codebook of visual words. It is an indirect clustering method which can keep the visual words in the codebook have higher intra-class similarity and relatively low inter-class similarity. The second step is: the BoF-SIFT descriptor generation for given image. The feature points can be replaced approximately by vocabulary in the codebook. By counting the number of times which is each word in the codebook appearing, the input image can be represented by a K dimensions value vector. The K-Means++ algorithm is more effective to select the initial seed point of clustering than K-Means algorithm. It is also known as the Lloyd algorithm, the algorithm steps are as follows:

The BoF-SIFT descriptors of different targets are shown in Figure 3. The BoF-SIFT descriptors of the same target with different rotation angle are shown in Figure 4. From which we can find the BoF-SIFT preserve bigger inter-class difference and better rotational robustness.

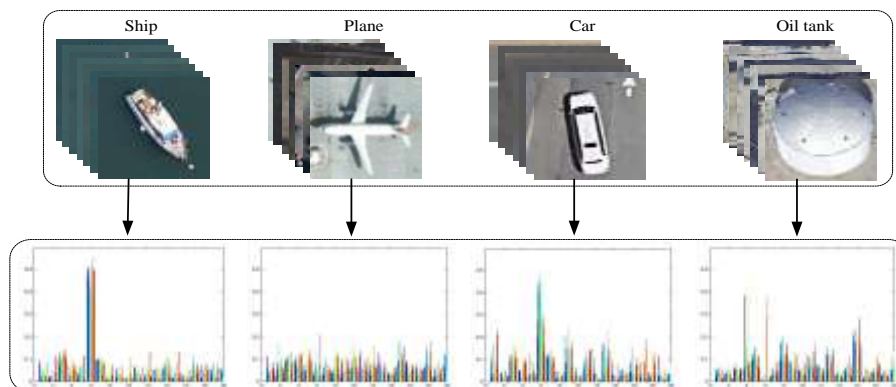


Figure 3. The BoF-SIFT Descriptors of Different Targets

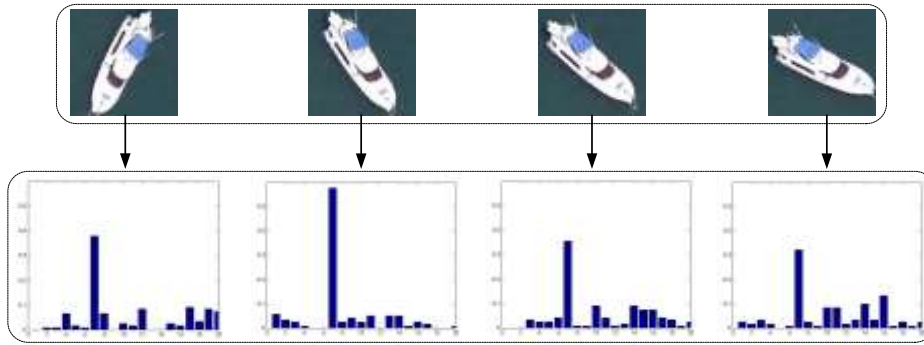


Figure 4. The BoF-SIFT of the Same Target with Different Rotation Angle

3. Hierarchical Features Extraction

Because of the distribution characteristics representation capability of the traditional BoF-SIFT descriptor is not adequate, hierarchical features was applied to the field of human action recognition [12] which proposed by combining pyramid idea [13] with the BoF-SIFT descriptor. It can not only represent the local features and global features effectively, but also represent the distribution characteristics of image feature points. The whole hierarchical features extraction process is shown in Figure 5. The detail is following:

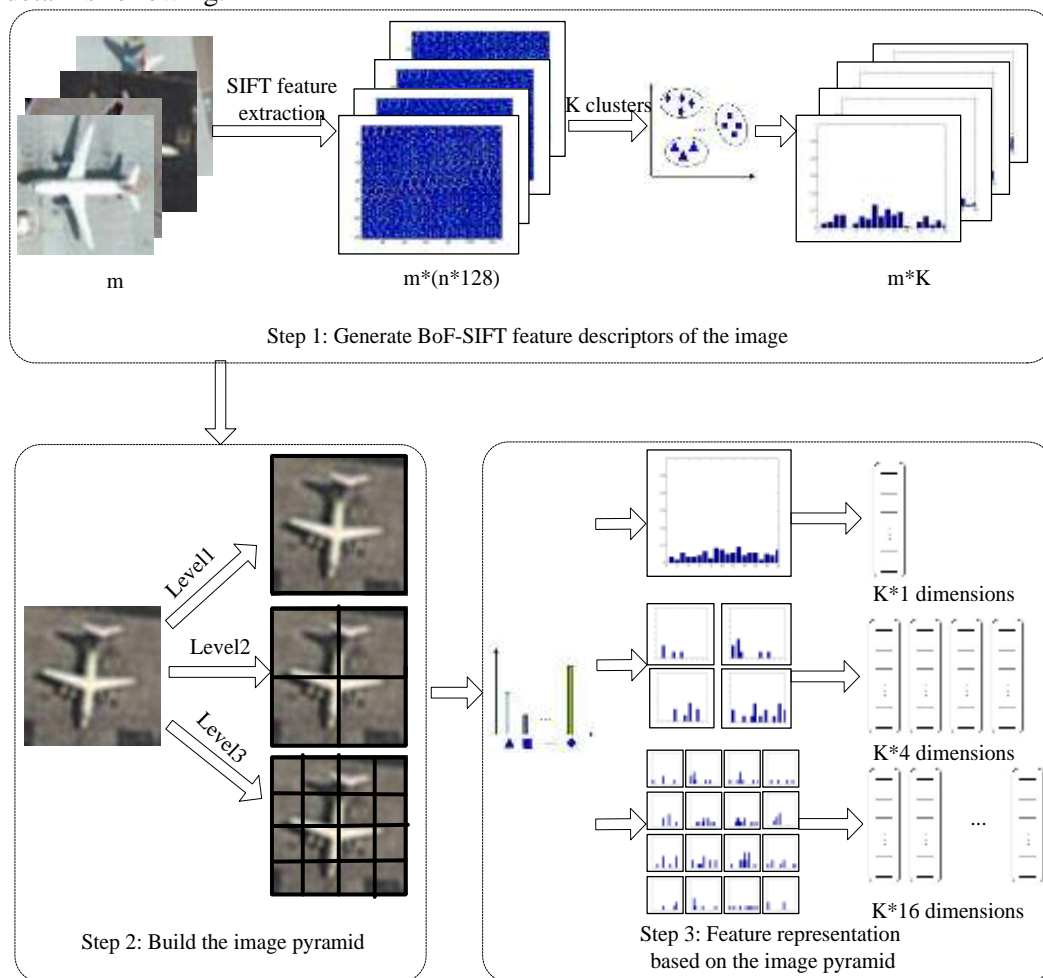


Figure 5. Feature Extraction Algorithm Diagram Based on Pyramid Structure

Step 1: Generate BoF-SIFT feature descriptors of the image: the number of samples is chosen as m . Each image is evenly divided into n blocks. The number of cluster centers is defined as K . The final dimensions of BoF-SIFT features is $m \cdot K$.

Step 2: Build the image pyramid: the original image is divided into different blocks to construct three levels image pyramid. The first level is the whole image. The second level is divided into 4 sub blocks. The third level is divided into 16 sub blocks.

Step 3: Feature representation based on the image pyramid: The visual vocabulary histogram of BoF-SIFT descriptors are generated by 1:1:1 weighing in each level feature of image pyramid. The feature descriptors of each image is represented by a $K \cdot 16$ dimensions vector.

4. Recognition Algorithm

AdaBoost algorithm with SVM as weak classifier is set as the recognition algorithm.

4.1. Brief Introduction of the SVM Principle

SVM [14] is proposed by Corinna Cortes and Vapnik in 1995. It shows many unique advantages in the small sample, nonlinear and high dimensional pattern recognition. It also overcomes the dependence on the sample numbers of traditional statistical pattern recognition method.

SVM utilizes the mapping of nonlinear to deal with nonlinear classification problem. The samples are transformed to high dimension by selecting appropriate kernel function $k(x_i, x)$, so that the transformed sample can be linear separable. Common kernels include Polynomial kernel, Radial Basis kernel and sigmoid kernel. The optimal discriminate function of its nuclear form is:

$$f(x) = \text{sgn} \left(\sum_{i=1}^n y_i a_i^* k(x_i, x) + b^* \right) \quad (3)$$

Here, sgn is sign function. y_i is Lagrange function. $a_i \geq 0$ is Lagrange coefficient. x_i is support vector. $b^* = -\frac{\max_{y_i=1}(w^* \cdot x_i) + \min_{y_i=-1}(w^* \cdot x_i)}{2}$.

Radial Basis function(RBF) is a radially symmetrical scalar function. It is generally defined as the monotone function of the Euclidean distance from any point x in space to a center x_i . It can be written as $K(\|x - x_i\|)$. Then the role is often partial, that is when x is far away from x_i , the function value is very small. The form of radial basis

kernel function is $K(x, x_i) = \exp\{-\frac{\|x - x_i\|^2}{\sigma^2}\}$. The SVM with RBF is corresponding to a radial basis function classifier.

4.2. AdaBoost Algorithm Based on SVM

AdaBoost algorithm is one of the most famous Boosting algorithms. AdaBoost as the most common algorithm has been used in various fields of machine learning. AdaBoost [15] is a kind of iterative algorithm. Its core idea is to construct a strong classifier by training multiple weak classifiers according to the different training sample subset with the different sample distribution.

AdaBoost algorithm achieves new training subset by calculating and adjusting the weight of each sample in the iteration process. The corresponding initial weight of each sample is the same. Given sample distribution is D_i , the initial training subset S_i is obtained from the total training samples S under the sample distribution. The initial

weak classifier $h_1(x)$ is achieved by learning the corresponding training subset S_1 . In boosting learning, each sample is associated with a weight, and the weights are updated dynamically by using a multiplicative rule according to the errors in previous learning so that more emphasis is placed on those examples which are erroneously classified. The iterative process is to train the new sample subset S_i in the new sample distribution D_i every time and get new weak classifiers $h_i(x)$. The process of training sample and acquiring new weak classifier is called an iteration every time. After t iterations, we can obtain t weak classifiers $H(X) = \{h_1(x), h_2(x), \dots, h_t(x)\}$. Finally according to certain rules these weak classifiers are combined to a high classification accuracy classifier $h_f(x)$ (The stronger classifier). SVM is proposed as weak classifiers in paper. The AdaBoost algorithm is shown in Figure 6:

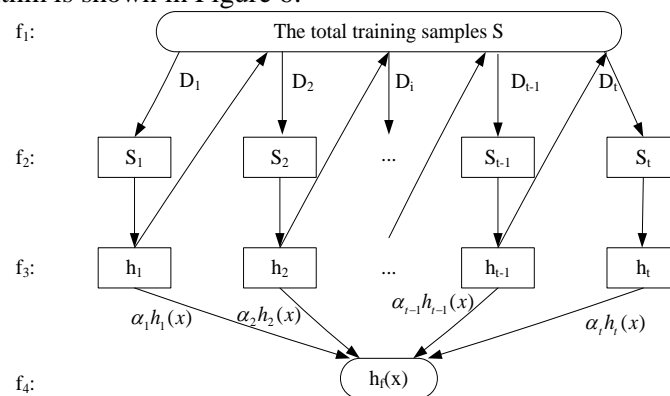


Figure 6. Schematic Diagram of AdaBoost Algorithm

5. Experiment and Results Analysis

The experiments are performed on a new optical remote sensing image dataset by MATLAB software. The dataset contains four kinds of targets, *i.e.*, ship, plane, car and oil tank. Each kind of target contains 74 samples, so that are 296 images in our dataset. The classification and recognition of target is performed on the complicated background. The part of target samples in the dataset is shown in Figure 7.



Figure 7. The Target Samples in the Dataset

5.1. Parameter Test

The classification experiments are performed on our dataset by using SVM with RBF kernel function. The results show that the extracted $4*4*8=128$ gradient information of feature points is the most appropriate. At the same time, considering the influence of different parameters setting on the experimental results, such as the number of training set samples Tr , the number of cluster centers K and the number of pyramid level L , the experiments are carried out on the different settings of these parameters in order to achieve optimal parameters.

1. The effects of the number of training set samples and the number of cluster centers on recognition rate

Given the training set samples Tr are 68, 108, 148, 188 images and the test samples Te are 108 image. Given the number of cluster centers K are 10, 20, 30, 40 and level of the Pyramid level L is 3. The recognition rate is obtained by using SVM with RBF kernel. The obtained recognition rate curves are shown in Figure 8.

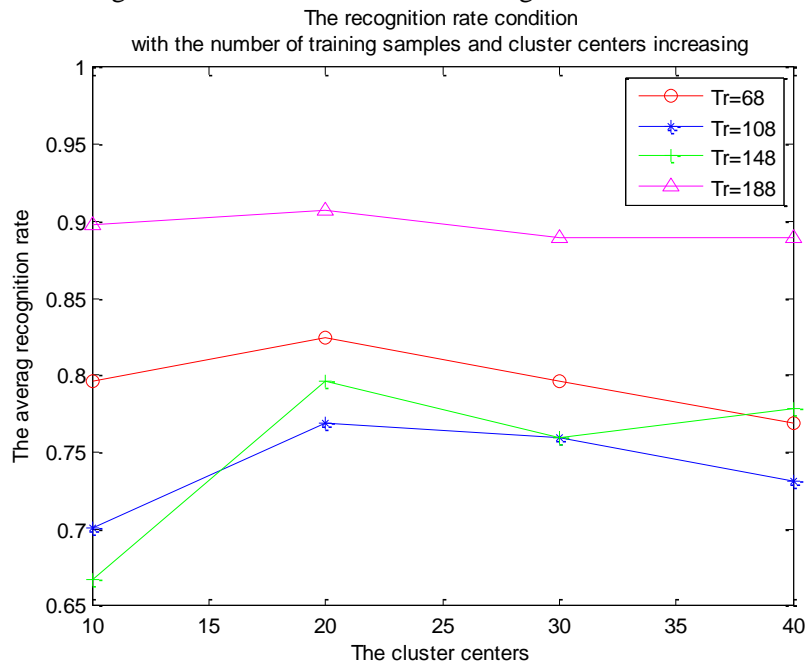


Figure 8. The Average Recognition Rate on the Different Number of Training Set Samples and Cluster Centers

As them can be seen from the chart, when the number of training samples is 188 and the number of cluster centers is 20, the recognition rate of the proposed method reaches the maximum value.

2. The effects of Pyramid level number on recognition rate

Under the optimal parameters above, that is $Tr=188, K=20$, multi-layer features of image Pyramid(not the highest layer) is trained and recognized by RBF kernel based SVM. The average recognition rate is shown in Table 1. The obtained recognition confusion matrix of multi-layer levels features is shown in Figure 9. The results show that the recognition rate is the highest when $L=3$.

Table 1. The Recognition Rate of Each Layer Images Pyramid

Pyramid layers	The feature vector dimension	The average recognition rate
1	20	82.41%
2	20+80	90.73%
3	20+80+320	90.74%

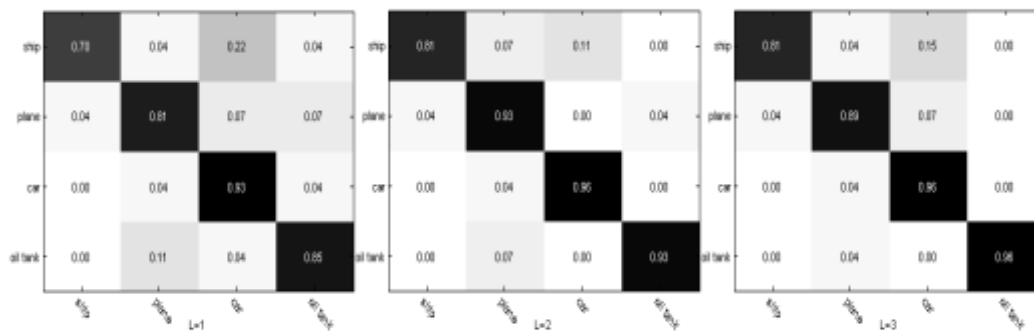


Figure 9. The Recognition Confusion Matrix of Images Pyramid from 1-3 Layer

5.2. The Recognition Results Comparison on the Optimal Parameters

Using the optimal parameters obtained on Section 5.1, $Tr=188$, $K=20$, $L=3$, the average recognition rate and running speed comparison of the following three methods is shown in Table 2.

Table 2. The Recognition Rate Comparison of Different Methods

Experiment number	The recognition methods	The average recognition rate	running speed/seconds
Experiment 1	BoF-SIFT feature+ RBF kernel SVM	82.41%	0.018427
Experiment 2	Hierarchical BoF-SIFT feature +RBF kernel SVM	90.74%	0.095406
Experiment 3	Hierarchical BoF-SIFT feature +(RBF kernel SVM+AdaBoost)	93.52%	0.028114

Although the running speed of experiment 2 is slower than experiment 1, corresponding recognition rate is more higher. It shows that the fusion of hierarchical features and BoF-SIFT feature represents the global and local characteristics of the targets effectively. Experiment 3 shows that the combination of AdaBoost+SVM has greatly improved the recognition rate and running speed on basis of experiment 2.

6. Conclusions

The proposed recognition method based on hierarchical features and AdaBoost algorithm has advantage on representing the distribution characteristics of feature points in the relative to the traditional single feature. The combination of the AdaBoost and SVM improves the performance of the classifier. Experiments show that the method has certain feasibility in the field of multiple targets classification and recognition of optical remote sensing image. The follow-up will consider the further fusion of feature

level and decision level in order to improve the recognition rate.

Acknowledgments

The project supported by the Program for Liaoning Excellent Talents in University (No. LJQ2014018) and Natural Science Foundation of Liaoning Province(201602577, 2015020101)

References

- [1] X. Ji and N. Qin, "A review of target detection classification and recognition method based on optical remote sensing image", *J. Journal of Shenyang Aerospace University*, vol. 1, no. 32, (2015), pp. 23-31.
- [2] B. Yang, H. Zhao, Z. Zhao and L. Zhang, "An improved target texture classification and recognition algorithm of remote sensing image", *J. Microelectronics & Computer*, vol. 9, no. 21, (2004), pp. 111-113.
- [3] Z. Chen, J. Liu and G. Wang, "A new circle targets extraction method from high resolution remote sensing imagery", *Proceedings of the 4th Advanced Computational Intelligence (IWACI), Wuhan, China, (2011) October 19-21*, pp. 529-533.
- [4] G. Liu, X. Sun, K. Fu and H. Wang, "Aircraft recognition in high-resolution satellite images using coarse-to-fine shape prior", *J. Proceedings of IEEE Transactions on Geoscience and Remote Sensing Letters*, vol. 3, no. 10, (2013), pp. 573-577.
- [5] J. Hsieh, J. Chen, C. Chuang and K. Fanl, "Aircraft type recognition in satellite images", *J. Proceedings of Vision, Image and Signal Processing*, vol. 3, no. 152, (2005), pp. 307-315.
- [6] C. Zhu, H. Zhou, R. Wang and J. Guo, "A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features", *J. Proceedings of IEEE Transactions on Geoscience and Remote Sensing*, vol. 9, no. 48, (2010), pp. 3446-3456.
- [7] D. Lowe, "Object recognition from local scale-invariant features", *Computer vision, 1999. The proceedings of the 7th International Conference On Computer Vision, Kerkyra, Greece, (1999) September 20-27*, pp. 1150-1157.
- [8] D. Lowe, "Distinctive image features from scale-invariant keypoints", *J. International Journal of Computer Vision*, vol. 2, no. 60, (2004), pp. 91-110.
- [9] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: a survey", *J. Foundations and Trends in Computer Graphics and Vision*, vol. 3, no. 3, (2008), pp. 177-280.
- [10] G. Sharma, S. Chaudhury and J. Srivastava, "Bag-of-features kernel eigen spaces for classification", *Proceedings of the 19th International Conference on Pattern Recognition, Tampa, FL, (2008) December 8-11*, pp. 1-4.
- [11] D. Arthur, S. Vassilvitskii, "k-means++: The advantages of careful seeding. Proceedings of the 18th annual ACM-SIAM symposium on Discrete algorithms", *Society for Industrial and Applied Mathematics, New Orleans, Louisiana, (2007) January 7-9*, pp. 1027-1035.
- [12] J. Wang, P. Liu, M. She, A. Kouzani and S. Nahavandi, "Human action recognition based on pyramid histogram of oriented gradients", *International Conference on Systems, Man, and Cybernetics (SMC), Anchorage, AK, United states, (2011) October 9-12*, pp. 2449-2454.
- [13] S. Lazebnik, C. Schmid and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories", *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, American, (2006) June 17-22*, pp. 2169-2178.
- [14] Q. Huang, G. Wu, J. Chen and H. Chu, "Automated remote sensing image classification method based on FCM and SVM", *Proceedings of the 2nd International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE), Nanjing, China, (2012) June 1-3*, pp. 1-4.
- [15] S. He, L. Wang, Y. Xia and Y. Tang, "Insulator recognition based on moments invariant features and Cascade AdaBoost classifier", *Proceedings of the 2nd International Conference on Mechatronics and Control Engineering(ICMCE), Dalian, China, (2013) August 28-29*, pp. 362-367.

Authors



Xiaofei Ji, received her M.S. and Ph.D. degrees from the Liaoning Shihua University and University of Portsmouth, in 2003 and 2010, respectively. From 2003 to 2012, she was the Lecturer at School of Automation of Shenyang Aerospace University. From 2013, she holds the position of Associate Professor at Shenyang Aerospace University. She is the IEEE member, has published over 40 technical research papers and 1 book. More than 20 research

papers have been indexed by SCI/EI. Her research interests include vision analysis and pattern recognition. She is the leader of The Program for Liaoning Excellent Talents in University (No. LJQ2014018). E-mail: jixiaofei7804@126.com (Corresponding author)



Ningli Qin, receives her B.Eng. Degree in Measuring and Control Technology and Instrumentations from Qilu University of Technology, Jinan, China, in 2013. She is currently a graduate student studying for Master degree in the school of Automation, Shenyang Aerospace University. Her research is focus on the remote sensing image processing and analysis. She has published 1 research papers in this research direction.



Yangyang Wang, received her M.S. degrees from the Shenyang Aerospace University, in 2006. She is currently a graduate student studying for Doctor degree in the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics. Her research is focus on the human action modeling and recognition. She has published over ten research papers in this research direction.