

The information extraction of Gannan citrus orchard based on the GF-1 remote sensing image

S Wang¹, Y L Chen

¹ School of Architectural and Surveying Engineering, Jiangxi University of Science and Technology, 341000 Ganzhou, China
1125551714@qq.com

Abstract. The production of Gannan oranges is the largest in China, which occupied an important part in the world. The extraction of citrus orchard quickly and effectively has important significance for fruit pathogen defense, fruit production and industrial planning. The traditional spectra extraction method of citrus orchard based on pixel has a lower classification accuracy, difficult to avoid the "pepper phenomenon". In the influence of noise, the phenomenon that different spectrums of objects have the same spectrum is graveness. Taking Xunwu County citrus fruit planting area of Ganzhou as the research object, aiming at the disadvantage of the lower accuracy of the traditional method based on image element classification method, a decision tree classification method based on object-oriented rule set is proposed. Firstly, multi-scale segmentation is performed on the GF-1 remote sensing image data of the study area. Subsequently the sample objects are selected for statistical analysis of spectral features and geometric features. Finally, combined with the concept of decision tree classification, a variety of empirical values of single band threshold, NDVI, band combination and object geometry characteristics are used hierarchically to execute the information extraction of the research area, and multi-scale segmentation and hierarchical decision tree classification is implemented. The classification results are verified with the confusion matrix, and the overall Kappa index is 87.91%.

1. Introduction

The orange region of Gannan in Jiangxi province has become the largest planting area in the world, which production is the third in the world, the first in China. To 2011, Ganzhou navel orange planting area of the city reached 119000 Hm², which annual output reached 1363900 T. The brand of Gannan navel orange has been famous at home and abroad (Wang J H, Ganzhou Citrus Research Institute, 2015). How to quickly and efficiently access the Gannan orange planting area and other related information has important significance for guiding the local fruit production and industrial planning. The development of remote sensing technology provides advanced technical means for obtaining the information of the Gannan orange planting area.

With the development of remote sensing technology, remote sensing image classification of vegetation experienced the development stage of visual interpretation, supervised classification and unsupervised classification, artificial neural network, fuzzy mathematics, expert system classification, decision tree and object-oriented interpretation, which provides some new techniques for classification of vegetation. In the classification of high resolution remote sensing images, due to the irregularity of vegetable coverage areas relative to the artificial objects, there is a certain degree of difficulty in classification, and the lower classification accuracy is always a difficult problem to solve. Yuan J G [1] has done a research on the forest vegetation classification based on NDVI (Normalized Difference Vegetation Index), which showed higher classification accuracy. Brian D. wardlow et al (2008) [2] has demonstrated that the classification of crops in the United States using hierarchical classification method based on the time series data of MODIS-NDVI achieved a high accuracy. Sulong et al [3], under the support of GIS (Geographic Information System), in the use of aerial photographs and TM images, has done the classification of the mangrove forest in Malaysia into 14 types and the accuracy respectively are 91.2% and 87.8%.

In recent years, as a large-scale comprehensive mean of earth observation, remote sensing technology reflects the characteristic of high spatial resolution, high spectral resolution and high time resolution (Li D R, 1995, 2002). Traditional image classification based on pixel spectral information is difficult to be applied to high resolution remote sensing image. Therefore, the object-oriented method based on the image began to rise, offsetting the limitations of the traditional method which uses pixels



as the basic classification and processing unit, which used the objects consisting of adjacent pixels containing more semantic information as processing unit. Han S S et al [4] has shown that the object-oriented method is more accurate than the traditional maximum likelihood method based on TM and ETM. He Y H et al [5] has found in the study of land investigation based on CBERS-02 that object-oriented method has higher accuracy than other methods of classification. In the experiment of classification based on IKONOS image, Sun X X et al [6] apply the object-oriented method combined with the hierarchical classification method to improve classification accuracy, which employed the object shape feature. Recently, the object-oriented classification method is applied to the classification and extraction of vegetation. There were a lot of experts and scholars applying the object-oriented method to the classification of vegetation. Harken and Su-gumaranr [7] execute the classification of wetland using nonparametric object-oriented method and the classification accuracy was 92.13%. In China, Zhou C Y [8] and Jiang H [9] have executed the study of the classification of vegetation. Zhang X R et al[10] have done the extraction of scrub vegetation and the accuracy is up to 84.7%. Thus, the information extraction of citrus orchard based on the high resolution remote sensing image with the idea of object-oriented classification has high feasibility.

This paper, based on GF-1 remote sensing image, using the object-oriented classification theory combined with the decision tree classification model, choosing the best segmentation scale for image segmentation after pre-processing the image, doing the data statistics of spectral and object features of different objects on sample selected, has executed the study of the extraction of Gannan orange tree area using different features including the spectral characteristics and the object features, and obtained a satisfactory classification results.

2. General situation of study area

Study area (Xunwu County) is located at the junction in Jiangxi, Fujian and Guangdong provinces, the southeast of Jiangxi Province(Geographical coordinates: N24° 30' 40" -25° 12' 10" , E115° 21' 22" -115° 54' 25"), with the total area of 2311.38 Km², many mountains and hills among them. The mountain area accounted for 75.6% of the total area. The region belongs to the south of the subtropical red soil region. The red soil is widely distributed in mountain and hilly areas, which soil fertility is better. The soil is generally acid, and the main component is iron oxide, aluminium oxide and quartz, which is suitable for kinds of orange, navel orange and other crops.

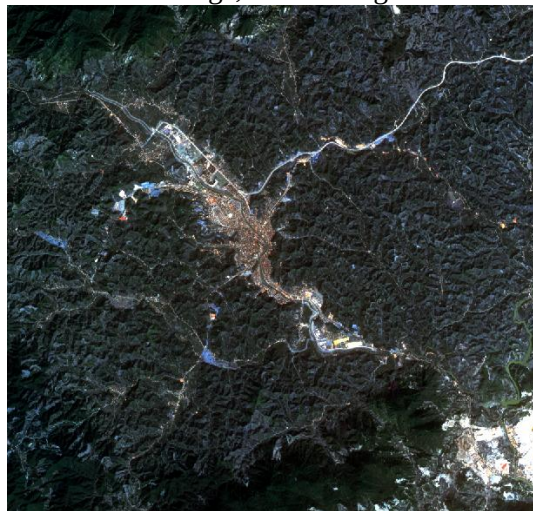


Figure 1. General situation of study area.

Such as shown in Figure 1, in hilly areas, due to the navel orange orchard and woodland, farmland and other background objects are staggered with each other and mosaic in space to constitute a complex mixture, also influenced by the shadow of the mountain, the automatic extraction is considerably difficult. The larger region in the image is the navel orange orchard planting area (abbreviation of orchard below), and the middle narrow part for buildings is the urban area(abbreviation of urban below) , and the white area for mining is the bare area (abbreviation of nudation below), and the road network is complex, no big rivers and lakes. According to the land use classification system of the resource and environment data centre of China Science and Technology

School, also according to the characteristics of the study area and preliminary interpretation, we can ensure the classification system of Xunwu, as shown in Table 1.

Table 1. Land use classification system.

The name	Description
Farmland	The land for planting crops, including cultivated land, new development land, consolidation and reclamation land, leisure land
Orchard	The land for planting fruit trees, widely distributed in the study area on both sides of the valley and hilly terrain
Forest	The forest land of arbors, shrubs, bamboos and other coastal mangrove, including land for herbage growth
Urban	The town for the residence of the various types of housing land and ancillary facilities land, including ordinary residential, apartments, villas and other land, and the rural homestead for life
Nudation	Towns, villages, factories, mines and internal unexploited land; and the surface soil, no vegetation covered, or the rocks and gravels on the surface. The coverage is more than or equal to 70%; surface is covered with sand, basic non-vegetated land
Road	The land of strip shape, its length greater than the width, the width changed relatively small, the smaller curvature

3. Data source and Preprocessing

3.1. Data source

This study selected the GF-1 remote sensing data covered with less than 20% cloud, and the data contains panchromatic image with resolution of 2m (sensor PMS1) and R, G, B, NIR 4 band multiband image with resolution of 8m (sensor PMS2), as shown in Table 2.

Table 2. GF-1 satellite data information.

Satellite name		Spectral range(μm)	Resolution(m)	Sensor	Note	
GF-1	Panchromatic	0.45-0.90	2	PMS1	Cloud cover<20%	
		0.45-0.52				
	Multiband	0.52-0.59	8	PMS2		
		0.63-0.69				
		0.77-0.89				

3.2. Preprocessing

Firstly, execute the radiation correction and geometric correction on the GF-1 remote sensing image data, furthermore, in order to obtain more accurate extraction all kinds of information from high resolution image, do the fusion processing to obtain more abundant image information to enhance the object feature to improve the classification accuracy.

4. Multi-scale segmentation

The segmentation technology is one of the core technologies of the high resolution image information extraction. The image is decomposed into non-overlapping regions as the basic unit of object-oriented analysis to make the analysis and understanding of the higher level possible (Zhang Y J, 2005). The image information extraction and subsequent analysis are determined directly by optimal segmentation scale. It is very important for the extraction of the target feature, the measurement and classification of the target object and the high level processing [11]. Multi-scale segmentation is, from a single pixel, to merge image objects into the high level objects by the bottom-to-up method. Through the calculation of pixel heterogeneity and homogeneity in size, the small image objects can be merged into some large homogeneous image objects after several times of merging, which can be understood as a similar pixel merged to form the image object stepwise optimization process. The result of the segmentation is determined by the segmentation parameter, which is composed of the scale parameter, the shape parameter and the compactness parameter.

Heterogeneity equation f :

$$f = w \times h_{color} + (1 - w) \times h_{shape}$$

$$h_{shape} = w_{compact} \times h_{compact} + (1 - w_{compact}) \times h_{smooth}$$

where w = spectral weight ($0 < w < 1$)

h_{color} = spectral heterogeneity

h_{shape} = shape heterogeneity

$w_{compact}$ = compactness weight ($0 < w < 1$)

h_{smooth} = smooth degree

$h_{compact}$ = compactness

As shown in Equation, heterogeneity equation is determined by spectral heterogeneity and shape heterogeneity, and shape heterogeneity is composed of smooth degree and compactness.

The scale parameter can be evaluated by the reference polygon. If the segmentation results are the same as the reference polygons, the multi-scale parameters can be used to realize feature extraction and can achieve good segmentation results. If the reference polygons are divided into different parts by the segmentation results, it shows that the segmentation scale is smaller. If the reference polygon is formed into the outer covering by the minimum area of the partition, the segmentation scale is too large. The shape parameters and compactness parameters can be determined by analyzing the contour of the segmentation. If the segmentation results divided the object of different attributes into the same object, or the object of the same attributes into different features, illustrating that the shape parameter and the compactness parameter are too small. Through many experiments, the segmentation parameters are selected as follows: scale parameter is 100, shape parameter is 0.1, and compactness parameter is 0.5.

5. Feature analysis and information extraction

5.1. Object semantic feature

Length-width ratio A :

$$A = L/W$$

where L = length W = width

The brightness mean value of the object :

$$\text{Brightness} = \sum_i^n B_i$$

where B = brightness value of the pixel i = pixel number n = sum of the pixels contained in the object

Normalized Difference Vegetation Index (NDVI):

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

where Red = red band mean value of the object NIR = near infrared band mean value of the object

5.2. Statistic and analysis of object feature

Through the statistical analysis of the DN of the 4 bands and NDVI values on all samples, the average brightness, 4 band reflectance and NDVI values of different vegetation types are obtained. Draw spectral curve and NDVI mean curve, as shown in Figure 2, Table 3.

Table 3. Terrain class brightness statistics.

Terrain class	Forest	Orchard	Urban	Road	Nudation	Farmland	Hill shade
Maximum value	191.5	210.3	217.8	311.3	510.4	180.3	154.5
Minimum value	158.8	182	165.8	211	274.2	165.8	109
Mean value	177.3	197.3	192.9	281.7	370.5	174.6	124.5

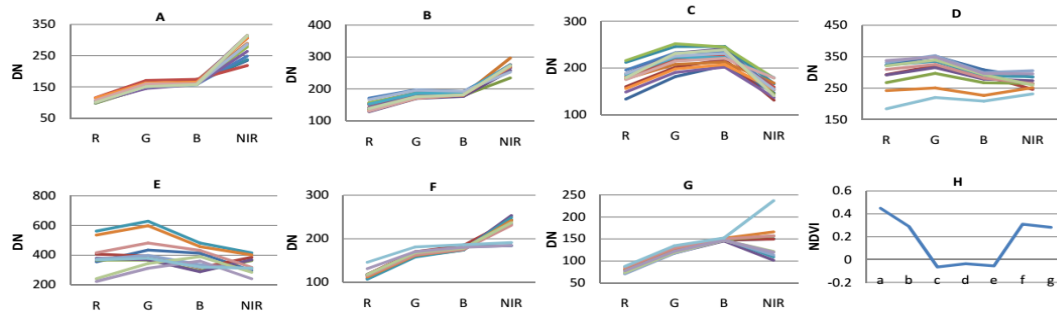


Figure 2. Spectral curve of terrain class and NDVI mean curve.

Where A. Forest B. Orchard C. Urban D. Road E. Nudation F. Farmland G. Hill shade
 H. NDVI (a. Forest b. Orchard c. Urban d. Road e. Nudation f. Farmland g. Hill shade)

5.3. Hierarchical classification of decision tree model

This paper uses the decision tree method, which is not only flexible, intuitive, clear, efficient, robust [12-14], but also can integrate various supplementary information effectively into the decision tree models for classification to obtain the ideal classification result[15-16]. In-depth understanding of the environment of the study area characteristics and spectral, spatial characteristics is the key to the establishment of decision tree classification model. Therefore, the basic indispensable work of establishing the decision tree classification is the analysis of spectral characteristics of typical ground objects and other characteristic variables (Zhao Y S, 2013). According to the above features, combined with the decision tree hierarchical classification, the feature space function threshold rule set model is established to extract the feature information.

1) Distinguish between vegetation and non-vegetation

The NDVI can better distinguish between vegetation and non vegetation. Because of the influence of mountain shadow, the threshold of NDVI 0.22 is smaller than normal vegetation extraction threshold and Mean (NDVI) >0.22 satisfies the classification rules for forest, farmland, orchard and hill shade, does not meet the road, nudation and urban.

2) Distinguish between hill shadow and other vegetation

Through sample statistics analysis, the R band DN mean of hill shadow object is smaller, so select classification rules Mean (R) <87 to distinguish and meet the rules for the hill shadow, otherwise forest, farmland and orchard.

3) Distinguish between forest and farmland, orchard

By sample analysis of forest and farmland, orchard in the B band has a high discrimination, so select classification rules Mean (B) <170 to distinguish and meet the rules for forest, or farmland and orchard.

4) Distinguish between farmland and orchard

Taking it into account that generally there is a certain distance between the orange trees in the orchard to increase fertility absorption and photosynthesis and farmland has higher vegetation coverage degree. Therefore, the brightness of orchard is slightly higher than farmland. Combined with statistical feature, select classification rules Brightness<181 to distinguish, meet for farmland, do not meet for orchard.

5) Distinguish between road and other non-vegetation

Due to that road showed as strip, and the object has a smaller density relative to other segmentation, and the length-width ratio is relatively large, so select classification rules Density<1.1 and Length/width>3.1 to distinguish, satisfy the conditions for the road, not satisfied for other non-vegetation.

6) Distinguish between nudation and urban

Nudation is generally exposed surface, with significantly higher than urban brightness values, so choose classification rules Brightness<223 to distinguish, satisfy the conditions for urban, otherwise it is nudation.

The above classification process can get the decision tree classification model as shown in Figure 3.

- 7) According to the field survey, we found that the shadow of the hill is generally the object around the types of objects, so the shadow of the hill is classified into the objects which have the longest public side with it.

6. Classification results and accuracy evaluation

After spectral analysis of remote sensing image and main object classification and merging, the classification results are derived as shapefile vector format, combined with ArcGIS data processing software to produce the main terrain classification results figure, such as Figure 4.

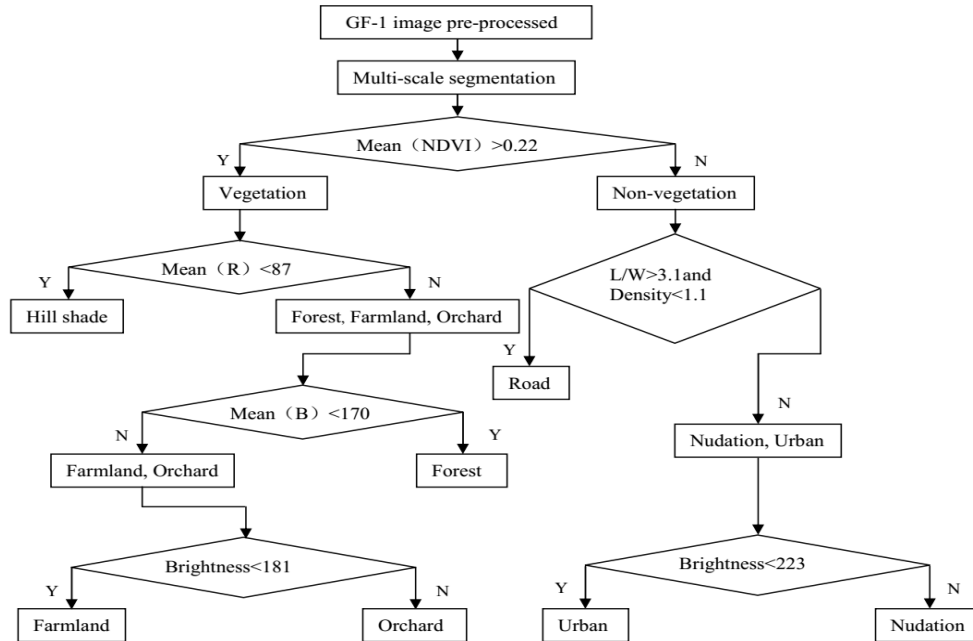


Figure 3. Classification decision tree model.

Accuracy evaluation is usually achieved by comparing the classification results with the reference data, and there are two kinds of data acquisition methods, one is the field survey, the other is to use the higher scale image or other information. This study uses the method of field survey. Through field survey we select 500 sample points, and the samples are generated by the sample points. Classification results extracted according to the experimental classification method have been compared with general classification extracted according to samples, and the quality evaluation confusion matrix is used to calculate the accuracy and the overall accuracy of classification is 91.94%, kappa coefficient is 0.8871, such as Table 4.

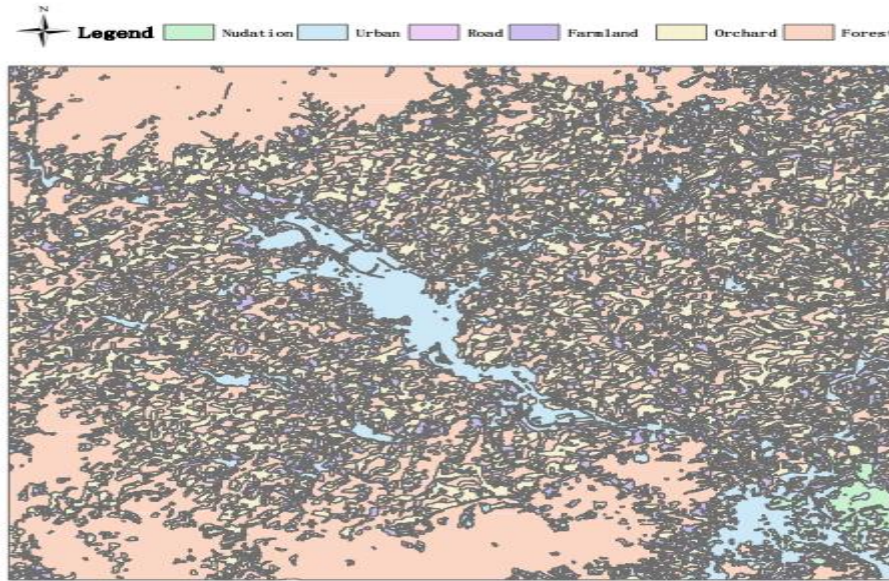


Figure 4. The result of classification.

As seen from Table 4, (1) The classification accuracy of forest is very high because of the obvious spectral feature, (2) The classification accuracy of farmland is very low due to that farmland is mixed in some forest and orchard, and its spectral feature have lower discrimination, (3) By the integration of multi-scale segmentation, object-oriented classification method and decision tree model, we can obtain higher classification accuracy of orchard.

Table 4. Classification accuracy evaluation.

Class name	Forest	Orchard	Urban	Road	Nudation	Farmland	Sum	User(%)
Forest	246	11	0	0	0	4	261	94.25
Orchard	4	86	2	0	0	1	93	92.47
Urban	0	1	240	4	20	1	266	90.23
Road	0	0	1	13	2	0	16	81.25
Nudation	0	0	1	0	53	0	54	98.15
Farmland	6	4	0	0	0	16	26	61.54
Sum	256	102	244	17	75	22	716	
Producer(%)	96.09	84.31	98.36	76.47	70.67	72.73		91.34

Overall Accuracy = 91.34% Overall Kappa = 0.879

7. Conclusions

The orange tree planting is an important agricultural industry of Gannan region, and timely effectively obtaining the relevant information of the orange orchard has important significance to guide the local fruit production and industrial planning.

We choose the object-oriented classification method, based on GF-1 remote sensing image, through multi-scale segmentation experiment, selecting the optimal segmentation parameters, taking not only the statistical spectral characteristics, but also the shape, size and topological relationship into consideration for the information extraction, to establish the decision tree classification model, obtaining the higher information extraction accuracy. The classification results prove fully the feasibility and superiority of object-oriented classification method.

The method also has some limitations and deficiencies. ①Since the data source for imaging is selected on December 29, 2013, non citrus fruit leafy season, the orange trees were sparse causing a certain degree of difficulty of the extraction of citrus fruit. ②For irregular non artificial objects, there is a lot of regular texture information in high resolution remote sensing images, so texture information will be very helpful to extract information of vegetation. Therefore, in our subsequent study of extraction, the main work is to combine texture information with other features of citrus orchard information to obtain a higher accuracy.

8. References

- [1] Yuan J G 1999. Study on Remote Sensing classification of forest vegetation. *Journal of Hebei Normal University (Natural Science)*, 1999(2), pp 274-7.
- [2] Wardlow B D and Egbert S L 2008. Large-area crop mapping using time-series MODIS 250m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment*, 112(3), pp 1096-116.
- [3] Sulong I, Mohd-Lokman H, Mohd-Tarmizi K and Ismail A 2002. Mangrove Mapping Using Landsat Imagery and Aerial Photographs: Kemaman District, Terengganu, Malaysia. *Environment Development & Sustainability*, 4(2), pp 135-52.
- [4] Han S S, Li H T and Gu H Y 2009. The Study on Land Use Change Detection Based on Object-oriented Analysis. *Remote Sense Information*, 2009(3), pp 23-9.
- [5] He Y H, Shi L S and Zhang R H. et al. 2007. Application Research of CBERS-02 Data in Land-use survey. *China Land Science*, 21(2), pp 51-7.
- [6] Sun X X, Zhang J X and Liu Z J 2006. The river and the road extraction from IKONOS panchromatic images using object-oriented classification method. *Science of Surveying and Mapping*, 31(1), pp 62-4.
- [7] Harken J and Sugumaran R 2005. Classification of Iowa wetlands using an airborne hyperspectral image: a comparison of the spectral angle mapper classifier and an object-oriented approach. *Canadian Journal of Remote Sensing*, 31(2), pp 167-74.
- [8] Zhou C Y, Wang P, Zhang Z Y and Qi C T 2008. Urban land use classification based on object-oriented information extraction technology. *Remote Sensing Technology and Application*, 23(1), pp 31-5.
- [9] Jiang H, Zhou W B and Liu Y 2008. Research and Application of the Poyang Lake WetLand Classification Using Remote Sensing. *Remote Sensing Technology and Application*, 23(6), pp 648-52.
- [10] Zhang X R, Liu L S, Zhang Y L, Zhang J P, Nie Y and Zhang H 2010. Extraction of Shrub Vegetation by Object-Oriented Classification Method Based on ENVI ZOOM in High-Altitude Area: A Case of Dingri County. *Geography and Geo-Information Science*, 26(4), pp 104-8.
- [11] Zhang Y J 1996. A Classification and Comparison of Evaluation Techniques for Image Segmentation. *Journal of Image and Graphics*, 1(2), pp 151-8.
- [12] Yu X Y, Zhang A D and Hou X Y 2012. Decision tree classification of orchard information extraction from TM imagery in Jiaodong Peninsula of China. *Science of Surveying and Mapping*, 37(4), pp 57-60.
- [13] Xiong W C, Wei B, Sun Z P, Shen W M and You D A 2010. Australian Forest Fire Disaster Monitoring Based on CCD and IRS Data of HJ-1-B. *Remote Sensing Technology and Application*, 25(2), pp 178-82.
- [14] Guo C H, Xue-Yong Q I, Gong Y L and Zhu L Y 2010. Study on the Technology and Method of Forest Fire Monitoring by Using HJ Satellite Images. *Remote Sense Information*, 2010(4), pp 85-8.
- [15] Vogelmann J E, Sohl T L and Howard S M 1998. Regional Characterization of Land Cover Using Multiple Data Sources. *Photogrammetric Engineering & Remote Sensing*, 64(1), pp 45-57.
- [16] Pielke R A, Lee T J, Copeland J H, Eastman J L, Ziegler C L and Finley C A 1997. Use of USGS-Provided Data to Improve Weather and Climate Simulations. *Ecological Applications*, 7(1), pp 3-21.
- [17] Luo W and Kuang R Y 2014. Orchard information extraction of Dongjiang Source region with HJ satellite data. *Science of Surveying and Mapping*, 39(8).
- [18] Wang Z W and Shi J Z, 2013. Assessment of vegetation by object-oriented classification and integration of decision tree classifier in Yushu. *Acta Pratacul Turae Sinica*, 22(5), pp 62-71.
- [19] Mu F Y and K Han 2012. Object-oriented Remote Sensing Wetland Vegetation Classification and Information Extraction. *Journal of Anhui Agriculture Science*, 2012(12), pp 7574-6.
- [20] Guo Y, Xinfang Y U, Jiang D, Wang S and Jiang X 2012. Study on Forest Classification Based on Object-Oriented Techniques. *Journal of Geo-Information Science*, 14(4), pp 514-22.

- [21] Chang H, Zhan F L, Yang G D, Niu X F, Zhang X Q, Shao P and Tang T Q 2015. Research on Objects-oriented Information Extraction Technology for High Resolution Remote Sensing Image. *Bulletin of Surveying and Mapping*, 2015(1), pp 99-101.
- [22] Zhang X. M, Sheng Y, Nan Z T, Zhao L, Zhou G Y and Yue G Y 2011. Vegetation classification of alpine grassland based on decision tree approach in the Wenquan area of the Qinghai-Tibet Plateau. *Pratacul Tural Science*, 28(12), pp 2074-83.

Acknowledgements

[This work is supported by science and technology projects from Jiangxi Provincial Office of Education Grant #GJJ10483.](#)