www.cafetinnova.org

Indexed in
Scopus Compendex and Geobase Elsevier,
Geo-Ref Information Services-USA, List B of Scientific Journals, Poland, Directory of Research Journals

ISSN 0974-5904, Volume 09, No. 05

International Journal of Earth Sciences and Engineering

October 2016, P.P.2300-2305

Research on the Extraction Method of Apple Orchard Information in Florescent Based on GF-1 Image-A case study In Yiyuan County

YUEFENG LU, SHUO LIU, CHEN FENG AND JIAXIN ZHANG
Shandong University of Technology, Zibo, Shandong, China
Email: *jarry1989ls@163.com*

Abstract: Data are increasingly available from new multi-spectral and high-resolution remote sensors for extracting agricultural information. We conducted a study of extracting apple orchard information from the GF-1 images. Using Yiyuan County in Shandong Province, China as the study area, we compared several different classification methods, including an integrated unsupervised-supervised classification method (improved maximum likelihood), the support vector machine (SVM) method, the nearest-neighbor method, and the decision-tree method. Particularly, we applied the first two methods to the pixel-based classification, and applied the last three methods to object-oriented classification. The comparison finds that the last three methods have higher efficiencies, whereas the first two methods have higher classification accuracy (according to the confusion matrix).

Keywords: *Extraction information, Supervised classification; Decision-tree classification, SVM, Object-oriented classification.*

1. Introduction

There are two mainly methods to extract vegetation information: pixel based extraction method and object-oriented method. Jin compared three supervised classification methods (paralleled method, minimum distance and maximum likelihood method) based on SPOT5 using ENVI, the result shows that draw maximum likelihood classification has highest overall classification accuracy, accuracy and user mapping accuracy [1]. Wang used original bands combination method, principal component analysis method (PCA), soil adjusted vegetation index (SAVI) method and the combination of original bands and derivative bands method to extract Shanghai's green space based on ETM+ images, the results show the combination of original bands and derivative bands method has the best accuracy and reduces soil background effect [2].Yang compared the classification accuracy and kappa coefficient of four kinds of object-oriented methods, including nearest neighbor method, a member function method, support vector machine and decision tree, the result shows the nearest neighbor method performed the best for forest classification in hilly mountain area [3]. Li compared traditional supervised classification, Support Vector Machine and Segmentation Support Vector Machine supervised classification with Object-oriented classification in Classification accuracy, the results show the object-oriented classification method can effectively extract the tobacco planting area of Luobin town [6].

On the basis of the existing vegetation information extraction method, in order to obtain one suitable for Chinese GF-1 remote sensing images, this paper used five existing methods to extract the apple orchard area information, including integrated supervised classification based on pixel (improved maximum likelihood), support vector machine method based on pixel, nearest neighbor method, decision tree based on object-oriented and support vector machine method based on object-oriented. By analysis the classification accuracy and kappa coefficient, the method for information extraction of fruit trees in mountain areas is obtained.

2. The Study Area and Data Characteristics

2.1 Study Area

Yiyuan County is located in the central of Shandong province, belongs to the Yimeng mountain area. Yiyuan County is the highest altitude county in Shandong province, knows as "The Roof of Shandong" (the location is shown in the Figure 1). The geographic coordinate are between 35°55'~36°23'N and 117°54'~118°31'E. The total area is 1636 km2, including cultivated area is 347 km2 and fruit tree area is 467 km2 (apple orchard area is about 173.33km2). The 70% of farmer's income comes from the fruit growing's. Yiyuan Apple is the geography specialty. Dynamic monitoring of dominant apple cultivated area and acquisition of orchard area and distribution is significant to Shandong province apple industry sustainable development.

Figure 1.*The administration and distribution of Yiyuan County*

Analysis the phenophase of apple shown in Table 1, it proved that the remote sensing images of apple florescence can be used to effectively identify the apple orchard in theory.

Table 1: *The Phenophase of Apple*

Name	Phenophase
Germination stage	Late March to early April
Flowering stage	Mid-April to early May
Fruit developing period	To Mid and late October
Flower bud differentiation stage	Mid- September to Mid-October
Deciduous rest period	Late November to early December

2.2 Data Characteristics

The GF-1 satellite carries on two panchromatic/multi-spectral cameras with 2m and 8m resolutions respectively, as well as four 16m resolution multi-spectral sensors. Wide field view (WFV) senor on board the Chinese GF-1, the first satellite of China High-resolution Earth Observation System, is acquiring multi-spectral data with decametric spatial resolution, high temporal resolution and wide coverage, which are valuable data source for environment monitoring. The technical specification of GF-1 WFV cameras is shown in Table 2.

Table 2: *Technical Specification of GF-1 WFV Cameras*

Payloads	Bands No.	Spectral range(μm)	Spatial resolution
WFV	1	0.45-0.52	16m
	2	0.52-0.59	
	3	0.63-0.69	
	4	0.77-0.89	

3. Study Methods and Experiments

The method of extraction information based on the Chinese GF-1 satellite could consult the data source with setting of similar bands (such as SPOT5 images, Quick Bird images, Landsat TM/ETM+ images). We select the following five representative methods of apple orchard information extraction to study the feasibility of methods.

3.1 Extraction method

3.1.1 Improved maximum likelihood classification

The maximum likelihood classification also known as the Bayesian Bayes classification, it is based on the image statistics of supervised classification method, and it is also a typical and applied most widely supervised classification method. And this method is applied to the spectral features of classification objects close to normal distribution, the lower classification error rate, the higher classification accuracy, and the more ideal classification results [4].

This paper uses ISODATA method to do unsupervised classification, using the classification result as the training sample template, and then do the maximum likelihood classification. For one hand, we can using the unsupervised classification results select the uniformed properties region. On the other hand, this method can greatly reduce the manual workload and improve the efficiency of classification, the classification accuracy with respect to the simple maximum likelihood classification is also improved [5].

3.1.2 Support vector machine (SVM)

Support vector machine adopts structural risk minimization learning machine, it is based on the strict theoretical basis, solves the nonlinear, high dimension and local minimum lighting problem well, and it becomes the new research hot spot of machine learning after the research of neural network [17]. This method is a new generation of supervised learning system based on the statistical learning theory.

This paper uses SVM classification method based on pixel and SVM classification method based on object-oriented to extract the information of apple orchard in the study area.

3.1.3Nearest neighbor classification

Nearest neighbor classification method is similar to the traditional supervised classification methods, they all determine the feature space by training samples. But the nearest neighbor classification method process the image objects and not a single pixel, so we can use abundant of characteristic information. What's more, the number of training samples required by the nearest neighbor method is small. Because a sample object can contain a lot of typical pixels as well as the changes between them. If a class needs to be described by a variety of features, it is suitable for the nearest neighbor classification method. This method is easy to deal with the overlapping of multi-dimensional feature space, and can quickly deal with the hierarchical relationship [18].Yi used the nearest neighbor and decision tree method to extract the paddy field in hilly region [6].

3.1.4 Decision Tree classification

As one of the main classification methods in Data Mining, the Decision Tree algorithm does not need

any prior knowledge or preferences and is suitable for exploratory knowledge discovery. With clearly defined structures, fast operation speed, high accuracy, flexibility and robustness, the Decision Tree algorithm can be used to handle high-dimensional data and the knowledge acquired is intuitive and easy to be understood [19]. In this paper, we used the decision tree method based on the object-oriented to extract information of apple orchard in study area.

3.2 Experiments

3.2.1 Technology process

The workflow of the paper for extraction information of Yiyuan County is shown in Figure 2.

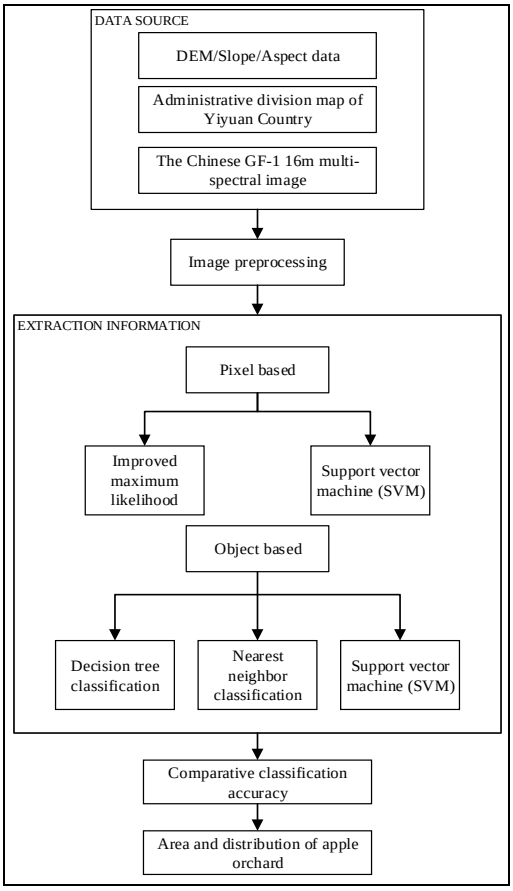

Figure 2: *Work-flow of the study*

3.2.2 Data preparation

Download GF-1 remote sensing image and DEM image. Using the administrative division of Yiyuan County as the border to cut out the remote sensing image. Using ArcGIS software to process the DEM image to obtain the slope image and aspect image. At the time of image data, including GF-1 image, DEM image, Slope image, Aspect image and Yiyuan county administrative divisions, the above as the initial data source.

3.2.3 Data preprocessing

GF-1 satellite grade 1A products are the products of relation radiometric correction. In order to improve the accuracy of classification, we need to preprocess the images, including decompression, radiometric correction, orthorectification, image fusion and image subseting. The workflow is shown in Figure 3.

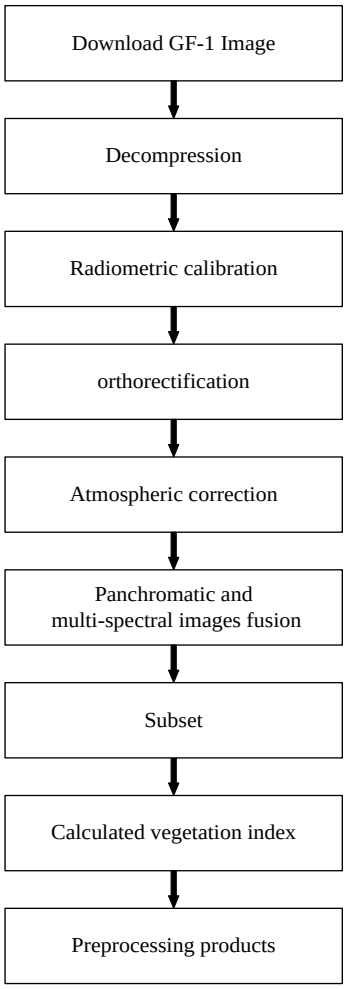

Figure 3: *Image preprocessing*

3.2.4 The Vegetation Index

For complex vegetation remote sensing, only use individual band or a single band data for comparative analysis to extract the information of vegetation has obvious limitations, which tend to use multi spectral remote sensing data by adding, subtraction, multiplication, in linear or nonlinear combinations, resulting in some valuable data of vegetation growth, biomass, and so-called vegetation index. Due to the spectral vegetation itself received, environmental conditions, atmospheric conditions and other factors, so vegetation index (VI) tend to have obvious regional and time domain [7]. The Vegetation index has been widely used to estimate fractional vegetation coverage [8, 9, 10].Gao used the vegetation index to inverse the biomass of vitex negundo shrubs canopy in Beijing Jundu mountainous area [11].

In this paper, by means of vegetation information extraction the most effective vegetation index as the supplementary information, select the normalized planting index (NDVI), the soil adjusted vegetation

index (SAVI), the difference vegetation index (RVI), the ratio of vegetation index (RVI) and the normalized water index (NDWI).

Table 3.*Vegetation indices*

Name	Formula
NDVI	(NIR - Red)/(NIR + Red)
NDWI	(Green - NIR)/(Green + NIR)
DVI	NIR - Red
RVI	NIR / Red
SAVI	[(NIR - Red)/(NIR + Red + 0.5)]*(1 + 0.5)

3.2.5 Accuracy evaluation

The kappa coefficient is a method to calculate the classification accuracy. Its calculated result is between -1 and 1. There are five groups to represent different levels of consistency.

Table 4.*Classification criterion based on Kappa coefficient*

Kappa	Degree of Consistency
0.0-0.2	Low consistency
0.21-0.4	General consistency
0.41-0.6	Medium consistency
0.61-0.8	Height consistency
0.81-1.0	Almost identical(perfect)

The Kappa coefficient based on the error matrix is a multivariate analysis method. The kappa coefficient can reflect the statistical significance in the classification results. There are a lot of scholars have done a lot of work in algorithm and application of Kappa coefficient. This method has been gradually developed into remote sensing classification accuracy evaluation method [12, 13, 14, 15].

4. Experiment

The results of apple orchard information extraction for five methods (improved supervised classification, SVM based on pixel, SVM based on object-oriented, nearest neighbor and decision tree based on object-oriented) based on Yiyuan county GF-1 images. Figure 9 shows As the Figure 4~8 showed.

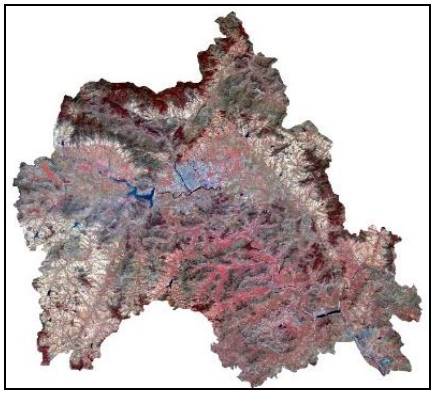

Figure 4.*Visual interpretation*

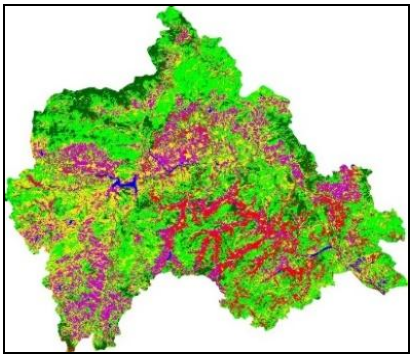

Figure 5. *Improved supervised based on pixel*

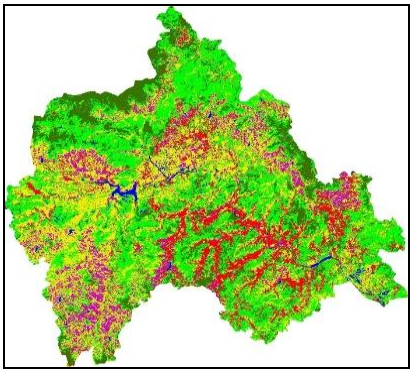

Figure 6. *Support vector machine (SVM) based on pixel*

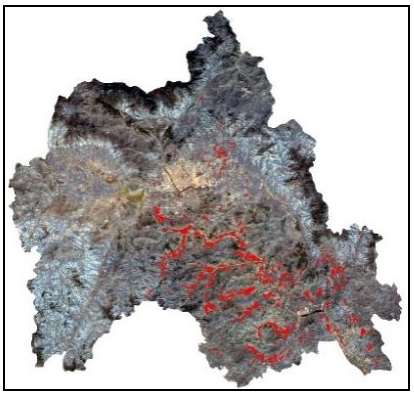

Figure 7. *Nearest neighbor classification based on object-oriented*

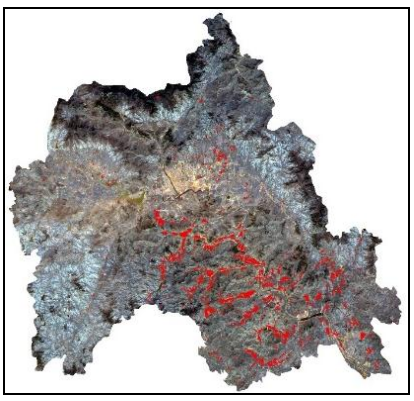

Figure 8. *Support Vector Machine based on object-oriented*

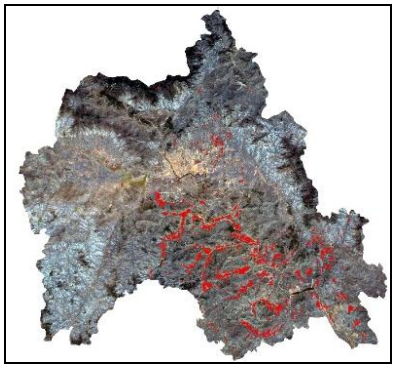

Figure 9. *Decision tree classification based on object-oriented*

Table 5. *Accuracy comparison of Five Classification Methods*

Classification name	Overall accuracy	Kappa coefficient	Area (hm2)	Real Area (hm2)	Area Error
Improve Supervised Classification	98.12%	0.9750	18,865	About 20,000	5.6%
Support Vector Machine(SVM) based on pixel	98.70%	0.9827	20,922		4.6%
Decision tree based on object-oriented	96.12%	0.949	22,643		13.2%
Nearest Neighbor based on object-oriented	99.5%	0.9985	19,567		2.1%
Support vector machine based on object-oriented	94.51%	0.9312	21,569		7.98%

4.1 Analysis the figures of classification results

Through the visual interpretation showing, the apple orchard area is roughly located in the south-east of Yiyuan County. Combining band4 (Red), band3 (Green) and band2 (Blue), the image displays a standard false color image, the vegetation covers occur the sight of red. Due to the time of fluorescence of apple, the most crops haven't growing up and the sight similar to the bare soil. The evergreen trees are located in the higher elevation area, surrounding with the crop area and fruit tree growing.

4.2 Analysis the accuracy of five classification method

The five classification method all extract the apple orchard area efficient, and have the similar accuracy. The nearest neighbor classification based on object-oriented method has the highest accuracy with the overall accuracy is 99.5% and kappa coefficient is 0.9985. The SVM classification based on pixel's accuracy is the higher method with the overall accuracy is 98.70% and the kappa coefficient is 0.9827. The improved supervised classification method is better than the traditional supervised method.

5. Conclusions

The method based on the pixel have a bit of "salt and pepper" in the results, that can result in the appearance of a number of broken polygons, ultimately leading to the low classification accuracy. The methods based on object segment the image and avoid the "salt and pepper", but the classification accuracy is influence in the segmentation scale. In this paper, the best method of the extraction information of apple orchard in Yiyuan County is the nearest neighbor classification. It has the best efficient and high classification accuracy.

The other two object-oriented method can be improved, we can study the different segmentation scales, and increase the classes to obtain the more detail classification results. The methods based on pixel can also improve the classification accuracy, for example, using decomposition of mixed pixels to reduce "synonyms spectrum" and "foreign body in the same spectrum" phenomenon and to improve the single pixel recognition.

The method of maximum likelihood and support vector machine based on pixel have developed used ENVI software. Because of the high special resolution, the processing speed is slow and cost for a long time. What's more, the human factors have influence in selecting samples. In summary, the supervised classification methods of ENVI software is suitable for medium-low resolution images' information extraction in a small range.

Ecognition software is an object-oriented professional classification software. Before the extraction of information, segmentation process should be done at first. The segmentation processing has been fully considered the impact of texture and other spatial features. The processing speed is faster than the methods using ENVI software and the efficiency is high. This paper selects three kinds of commonly used object-oriented classification methods (the nearest neighbor, the decision tree and support vector machine). The segmentation scale is 20. The classification accuracy is in line with standard classification. The results can be reference for economic and agricultural and have the research significance.

6. Acknowledgment

The authors would like to express appreciations to colleagues in our laboratory for their valuable comments and other helps and anonymous reviewers and the editor for the constructive comments and suggestions, all of which have led to great improvements in the presentation of this article. We

also thank the China Centre for Resources Satellite Data and Application for providing the GF-1 dataset.

Supported by the National Natural Science Foundation of China (No. 41501425, No. 41561084), Shandong Provincial Natural Science Foundation of China (No. ZR2014DL001), Shandong Primary Research & Development Plan of China (2015GSF122008, 2016GSF122006), and Anhui Provincial Natural Science Foundation of China (No.1508085SQD207).

References

[1] Jie Jin,Haiyan Zhu,Zixiao Li, "The Comparison of Several Kinds of Supervised Classification Methods in ENVI Remote Sensing Image Processing", *Water Conservancy Science and Technology and Economy*,Vol.20 No.1,2014, pp146-160.

[2] Lishu Wang,Yanan Xie,Hui Geng, "Extraction useful method of Shanghai's green space based on ETM+ image ",*Electronic measurement technology*,Vol.33 No.12,2010,pp 4-7.

[3] Fei Yang,Lifeng Liu,Xuecheng Wang, "Study on Forest Classification Based on Object-oriented Method and SPOT5 Images in Hilly Mountain Area", *Forest Resources Managemen*t,No.5,2014,pp 92-99.

[4] Tao Liu,Zhonglin Sun,Lin Sun, "Research on remote sensing image classification based on maximum likelihood method", *Fujian Computer*,.Vol.1,2010,pp 7-8.

[5] Xin Yang," An Elementary Introduction to Supervised and Unsupervised Classification of Remote Sensing Image", *Acta Geological Sichuan*, Vol.28 (3), 2008, pp251-254.

[6] Fengjia Yi,Rendong Li,Bianrong Chang," Remote sensing identification method for paddy field in hilly region based on object-oriented analysis", *Transactions of the Chinese society of agricultural engineering*,Vol.31 No.11,2015,pp 186-193.

[7] Hailiang Xia,Bo wang, Honhwei Shao," Application of Closest Standard Classification Method Used in Road Extraction", *Surveying and Mapping of Geology and Mineral Resources*,Vol.28(3),2012,pp 27-29.

[8] Lianyu Yu,Huanjie Cai,Fuqi Yao," Applicability of vegetation indices to estimate fractional vegetation coverage", *Transactions of the Chinese society for agricultural machinery*,Vol.46 No.1,2015,pp 231-271.

[9] Hanyue Chen,Zheng Niu,Wenjiang Huang," Estimation of winter wheat LAI using hotspot-signature vegetation indices", *Transactions of the CSAE*,Vol.28 No.1,2012,pp 167-172.

[10] Xinchuan Li,Xingang Xu,Yansong Bao, "Retrieving LAI of winter wheat based on sensitive vegetation index by the segmentation method", *Scientia agricultura sinica*,Vol.17 No.45,2012

[11] Mingliang Gao,Zhaoming Gong,Wenji Zhao, "The study of vitex negundo shrubs canpy biomass inversion in Beijing Jundu mountainous area based on vegetation indices", *Acta ecologica sinica*,Vol.34 No.5,2014,pp 1178-1188.

[12] Feinstein A R,Cicchetti Dv, "High agreement but low Kappa:I. The problems of two paradoxes", *Journal of Clinical Epidemiology*, Vol.43 No.6, 1990, pp 543-549.

[13] Cicchetti D v,Feinstein A R," High agreement but low Kappa:II Resolving the paradoxes", *Journal of Clinical Epidemiology*,Vol.43 No.6,1990,pp 551-558.

[14] Pontius J,"Quantification error versus location error in comparison of categorical maps", *Photogrammetric engineering and remote sensing*, Vol.66 No.8, 2002, pp 1011-1016.

[15] Naesset E, "Use of the weighted Kappa coefficient in classification error assessment of thematic maps", *International Journal of Geographical information system*, Vol.10 No.5, 1996, pp 5591-604.

[16] .Weili Gong, Chuanwei Li, "Current situation and development suggestion of apple industry in Yiyuan of Shandong", *Deciduous Fruits*,Vol.47 No.1,2015,pp 14-16.

[17] Alberto Mirandola , Enrico Lorenzini "Energy, Environment and Climate: From the Past to the Future", *International Journal of Heat and Technology*,Vol.34 No.2,2016,pp 159-164

[18] Tiankun Li," The research of extracted tobacco planting area based on object-oriented classification method", 2013.

[19] Erzhu Li," Hyperspectral remote sensing image classification based on semi-supervised support vector machine", 2014.

[20] Tan Lin, "Application of Object-oriented classification based on SPOT image", 2012.

[21] Xiaohe Zhang," Implement of decision tree classifier and application in remote sensing image classification", 2013.

[22] Yingshi Zhao," Principles and Methods of Remote Sensing Application Analysis (Ⅱ)", 2013, pp 368-378.