HAL open science

Study on Methods of Extracting New Construction Land Information Based on SPOT6

Lei Guo, Dongling Zhao, Rui Zhang, Meng Du, Zhixiao Li, Xiang Wang, Yaru Wang

To cite this version:

Lei Guo, Dongling Zhao, Rui Zhang, Meng Du, Zhixiao Li, et al.. Study on Methods of Extracting New Construction Land Information Based on SPOT6. 9th International Conference on Computer and Computing Technologies in Agriculture (CCTA), Sep 2015, Beijing, China. pp.94-106, 10.1007/978-3-319-48357-3_10 . hal-01557848

HAL Id: hal-01557848

https://inria.hal.science/hal-01557848v1

Submitted on 6 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Study on Methods of Extracting New Construction Land Information Based on SPOT6

Lei Guo[1,a], Dongling Zhao[2,b,*], Rui Zhang [2,c], Meng Du[2,d], Zhixiao Li [2,e], Xiang Wang[1,f], Yaru Wang[1,g]

[1]Hebei Provincial Seismological Bureau, Shijiazhuang 050021, China;

[2]College of Information and Electrical Engineering,China Agricultural University, Beijing 100083, China）

[a]guolei0430@126.com, [b] zhaodongling@cau.edu.cn, [c]zhangrui2063@126.com, [d]1165028791@qq.com,[e]lizhixiao91@cau.edu.cn, [f] wangx@eq-he.an.cn, [g] wangyr24@163.com,

Abstract. SPOT6 is a new remote sensing satellite launched in 2012, with high spatial resolution and strong data acquisition ability. However, a complete data preprocessing technology for the regulation of land resources has not yet been formed. According to the characteristics of SPOT6 satellite images, four different image fusion methods -- Gram-Schmidt, HPF, PanSharpand PanSharpening were selected to conduct the comparison experiment by using the software platforms of ENVI, ERDAS and PCI. We evaluate the results' performances from 3 different aspects. First, evaluating the image quality of experiment results qualitatively, then assessed quantitatively by establishing evaluation indexes including mean, standard deviation, information entropy, average gradient and correlation coefficient. Finally, evaluating the applicative effect of fused images based on the classification accuracy. The analysis results shows that the method of PanSharp is best to extract construction land information. Based on the PanSharp fusion image, in order to obtain the texture information under different scales, the authors screened the texture features according to Shannon entropy, and then used distance-based approach J-M to calculate the separation for choosing the optimal texture window. Once got the texture information, combining it with the original image to participate in the multi-scale image classification. The research result showed that multi-window texture participation in classification can improve separation of objects. Finally we extract construction land information with the method of SVM .This study may provide the technical support for application of SPOT6 image in the land resources management.

Keywords: SPOT6, image fusion, SVM, multi-window texture, new construction land

1 Introduction

Construction land refers to basic premises of all social and economic activities people engaged in, including residential land, roads, land use, public service area (excluding Greenland and water) [1]. With China's rapid development of industrialization and urbanization construction land most prone to change, especially in the last 30 years, with the rapid expansion and he sprawl phenomenon of construction land, causing decreased utilization of land resources, reduction of arable land resources, environmental pollution and ecological destruction and other issues. Therefore, it has important significance for the rational development of urban planning programs, socio-economic and resource use and sustainable development of ecological environment to obtain new construction information timely and accurately.

For high-resolution image information extraction, it can't rely solely on spectral characteristics, but also greater use of spatial information of remote sensing images. In recent years, the use of auxiliary image texture information of spectral classification in order to enhance the classification accuracy is becoming a hot topic of remote sensing information extraction areas [2]. Based on the research of suitable fusion methods of extracting construction land information, this paper tries to use the object-oriented SVM classification combined with multi-window texture and spectral information to extract new construction land information.

2 Materials and Methods

2.1 Study Area and Data

The study area was located in spa town, Haidian District, Beijing of China, with the controlled area of about 33.32km^2, with an average elevation of about 50 meters. Plain area is 17.79 km^2, accounting for 55%, mountainous area is 14.53 km^2, accounting for 45%.The town contains a wealth of land use / land cover types, is an ideal test area of land use information extraction.

The data sources of study area included SPOT 6 satellite data obtained on January 25th, 2013 and a topographic map with the size 1/2000 obtained on 2008. The vegetation coverage of image is low, some places have snow cover.

2.2 The Principle of Support Vector Machine

SVM method is based on the VC dimension theory and structural risk minimization principle of statistical learning theory. Based on the limited sample information, finding the best compromise between complexity of the model and learning ability, in order to obtain the best generalization ability [3,4].

By learning algorithm, SVM can automatically find the support vector that have a greater ability to distinguish, then construct a classifier that can maximize the interval between classes, so there is a good promotion and a higher classification accuracy.

For simple binary classification problem, SVM is to find an optimal split plane that can be adjusted as correctly as possible to separate two kinds of data with wider distance. The basic idea is illustrated by the following figure.

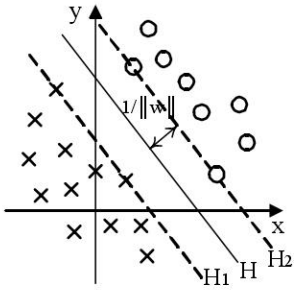


Fig.1. optimal hyperplane

Solid point and hollow point of figure s represent two categories of training samples; H is classified line; the distance between H1 and H2 are classified intervals. According to the maximum interval method, we require not only the classification lines separate the two classes of samples correctly, but also should make the interval of classification is max; extended to high-dimensional space, the optimal classification line can be extended to the optimal hyperplane.

2.3 Test flow

The test procedure of this research mainly includes pre-processing of images (orthorectification, image fusion and image clipping), extracting of texture feature, combination of multi-source information, classification of images based on SVM (multi-scale segmentation ,parameter setting of the value of C and γ, etc.) and accuracy of evaluation. The test flow chart is expressed as Fig. 2.

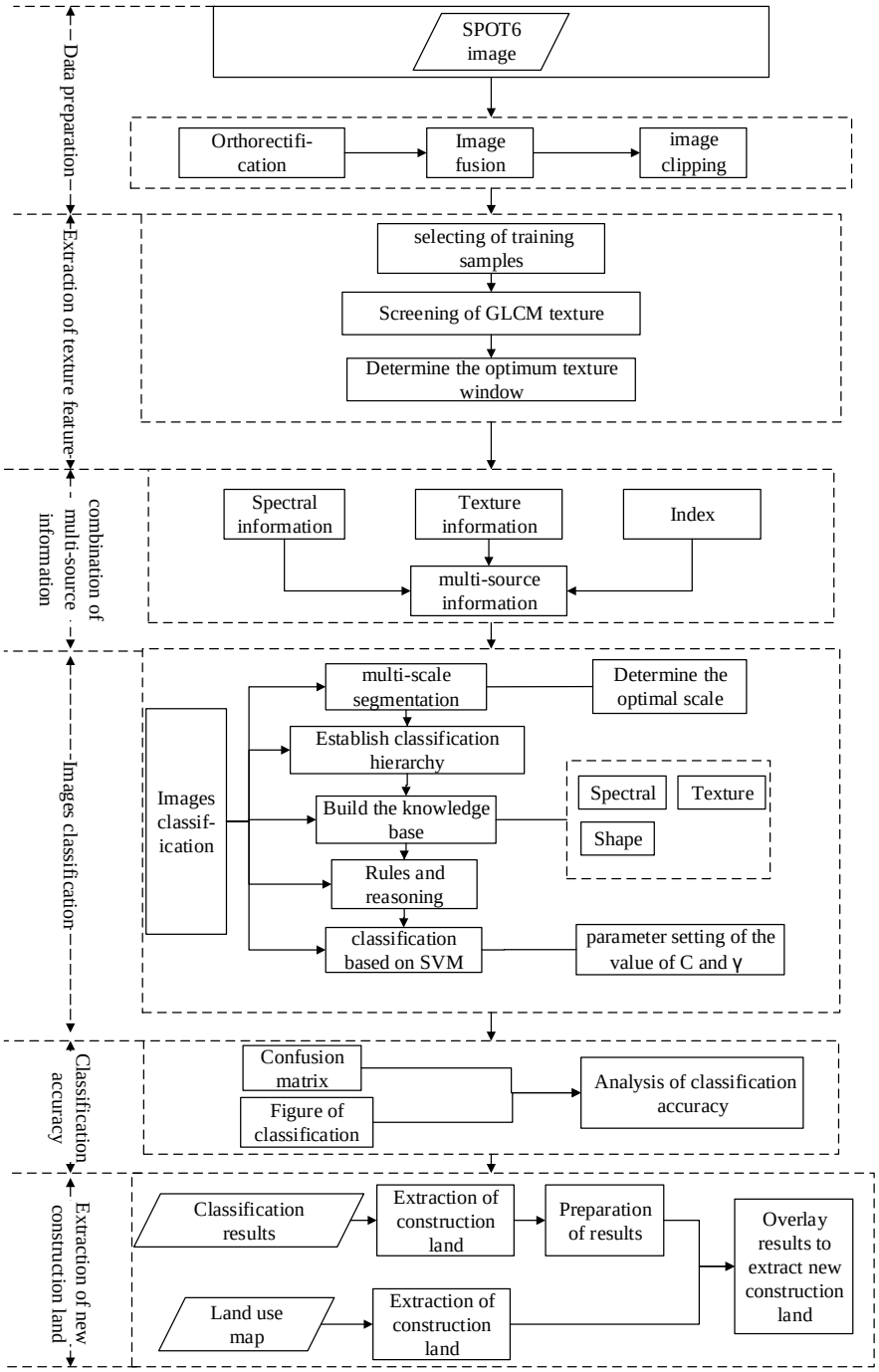


Fig. 2. The flow chart of experimental design

2.4 Data Preparation

Data preprocessing including orthorectification, image fusion into and image clipping. In this paper, data preprocessing process is mainly focused on the choice of fusion methods.

Image fusion can compensate for the lack of information in a single image, take advantage of a variety of images, get multi-faceted feature information, is a crucial step of information extraction [5]. According to the characteristics of SPOT6 satellite images, four different image fusion methods of Gram-Schmidt, HPF, PanSharp and PanSharpening were selected to conduct the experiment of comparison. For evaluating the results' performances, we compare them from 3 aspects. The image quality of experiment results was evaluated qualitatively, and also assessed quantitatively by establishing evaluation indexes including mean, standard deviation, information entropy, average gradient and correlation coefficient. The applicative effect of fused images was also evaluated based on the evaluation of the classification accuracy, the results are shown in figure 3. The analysis results show that it is better to choose the image fusion methods of PanSharp to extract construction land information.

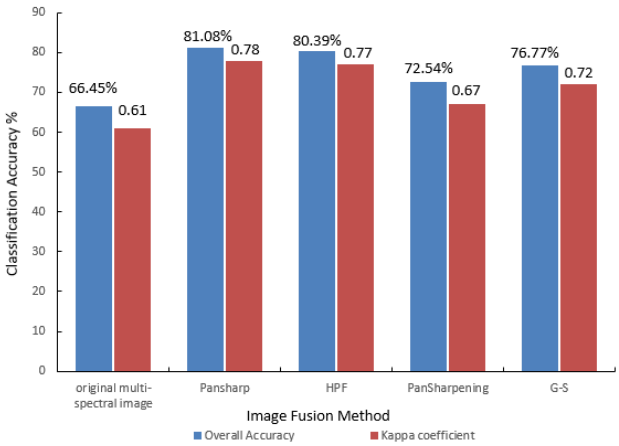


Fig. 3. Image classification overall accuracy and Kappa coefficient

2.5 Texture Feature Selection

Accuracy and precision of analysis will improve coupled with texture information on the basis of spectral information on the original image in the remote sensing thematic information extraction. GLCM can describe spatial distribution and structural characteristics of grayscale of each pixel in the image. There were advantages in improving the effect of ground objects classification in images by using texture of images. Then the texture feature of the first principal component can be extracted

according to gray-level co-occurrence matrix (GLCM) which is proposed by Haralick, et al. [6]. These texture statistics can be classified combined with multi-spectral remote sensing image. And each texture statistics can be involved in the calculation along with gray value of each band, etc, as a basis for classification [7]. There are 8 GLCM texture statistics which were often used in the 14 kinds , including mean, standard deviation, homogeneity, contrast, dissimilarity, entropy, angle second moment, correlation.

By using principal component analysis (PCA) while extracting the texture feature of SPOT 6 image, the result shows that the cumulative contribution of the first principal component is 95.73%,which represents the high and low frequency part of 4 bands of these images. The outcome not only has the effect on descending dimension and descending dimension, but also can incarnate the texture feature of land objects and basic tonality feature [8]. Then the 8 GLCM texture statistics of the first principal component can be extracted. Based on Shannon entropy and visual judgment, we compare and analysis eight kinds of texture features. In this paper, we calculate the information entropy of texture features in 3 × 3 window as example, the results are shown in Table 1.

Table 1 .Information entropy eight texture features

Mean	Standard deviation	Homogeneity	Contrast	Dissimilarity	Entropy	Angular Second Moment	Correlation
2.21	0.25	2.18	0.23	1.77	1.99	1.88	0.47

As can be seen from Table 1, the mean of the highest entropy, provide the most abundant information; homogeneity degree and entropy followed, can also provide a wealth of information. From visual judgment, although the mean entropy is high, but the edge of the texture image is not clear, there is no prominent feature of the border (figure 4), therefore excluded. So we chosen homogeneity and entropy involved in classification.

（a）Mean （b）Homogeneity （c）Entropy

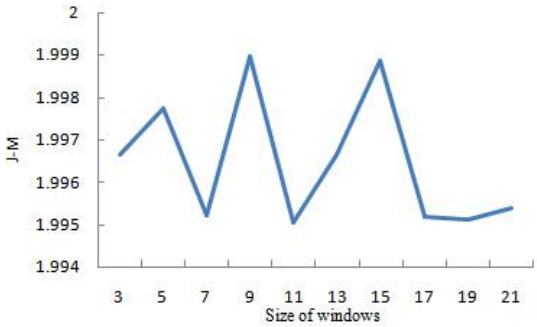
Fig. 4. Texture features of SPOT6 image (partial)

2.6 The Determination of Texture Window

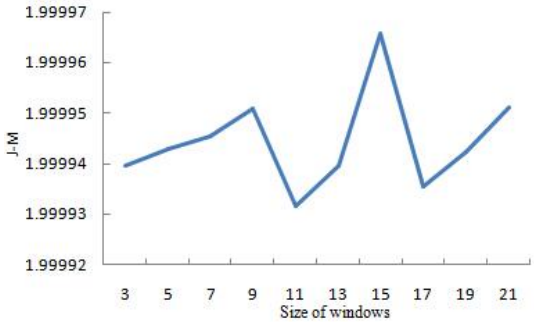
Research shows that, when the sliding window is larger than 21 x 21, it will be difficult to reflect object properties using the texture features we extracted, so the window size of 21 x 21 is chosen as the largest sliding window [9]. The moving

window size is set to 3 x 3, 5 x 5, 7 x 7, 9 x 9, 11 x 11, 13 x 13, 15 x 15, 17 x 17, 19 x 19, 21 x 21 in succession, in order to compare and analyze texture feature at different scales of windows.

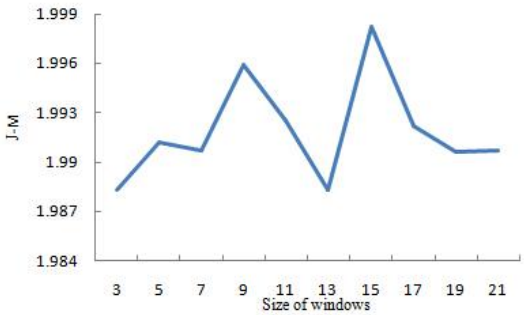
As different land objects feature extraction has different optimal texture scale, therefore need to filter the most suitable texture window [10]. In this paper, we choose the index J-M distance to calculate the separability. This paper, we are mainly for extract newly-added construction land, the types of land use in the study area is: vegetation covered area (including vegetation cover cultivated land, woodland), dark place (including water, shadow), construction land (including buildings, roads), bare land and snow all 5 categories. Figure 5, figure 6 shows the curve of class separability (J-M distance) with different texture windows.



（a）Dark place

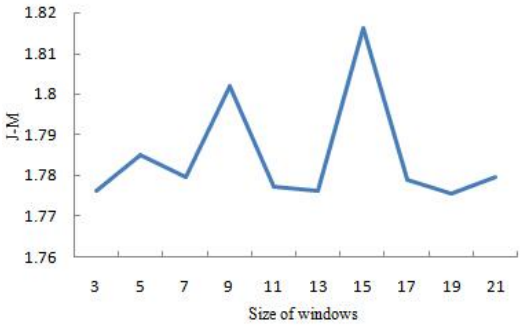


（b）Vegetation



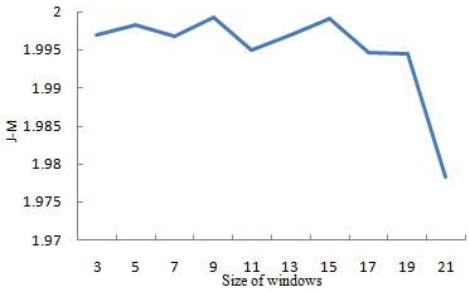
（c）Bare land

（d）Construction land

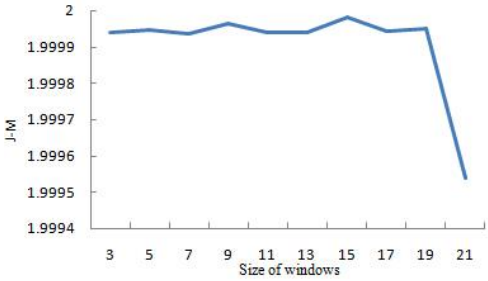


（e）Snow

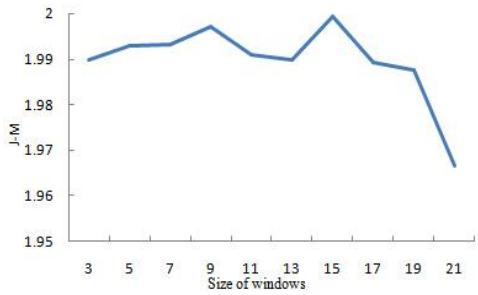
Fig.5. Curve of J-M distance with different size windows (Homogeneity)



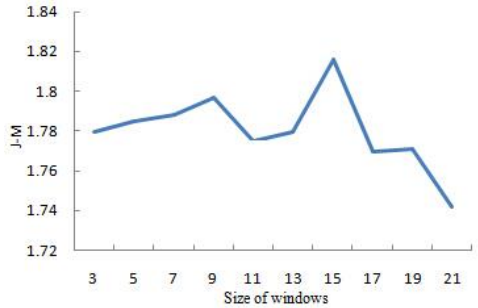
（a）Dark place



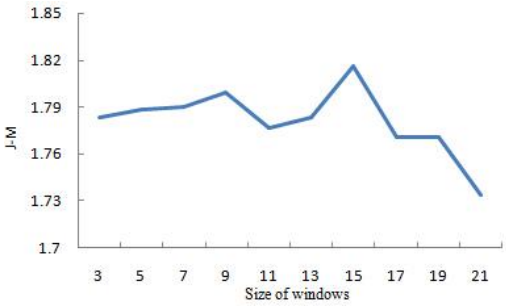
（b）Vegetation



（c）Bare land



（d）Construction land



（e）Snow

Fig. 6. Curve of J-M distance with different size windows (Entropy)

Figure 5 and Figure 6 show that for the texture features include entropy and homogeneous of SPOT6 images, all classes were not presented in increasing or decreasing trend. For the homogeneous, vegetation, bare land, snow, construction land can be divided good in the 15 × 15 window, the separability of dark place in 9 × 9 window reached the highest. For the texture features entropy, the separability of all classes in 15 × 15 window is highest, and reached the minimum in the window 21 × 21. So for the SPOT6 image, selected texture feature homogeneous in 9 × 9 and 15 × 15 window and entropy in 15 × 15 window take part in the classification. Multi-window texture take part in classification can improve the separability of different types of objects effectively.

2.7 The Determination of SVM Kernel Function and Parameters

This study selected the RBF kernel function as SVM kernel function, because the RBF kernel can map samples into a higher dimensional space, can deal with case when the relationship between class label and feature is nonlinear ; and is applied for low dimension, high dimension, small sample, large sample. It has wide convergence domain, is kind of ideal classification function [11]. The kernel parameter and penalty factor C are two necessary parameter of RBF kernel function. This paper we use cross validation algorithm to determine the two parameters, it means the selected training samples will be divided into an equal number of N subset, and make the N - 1 groups as a model of training samples, the rest of the group as the test sample, make use of the test sample to test and verify the classification results accuracy of n-1 part data, by changing C and γ, we looking for higher sample classification accuracy [12]. This paper we choose the parameters selection model Grid. py which provided by the software libsvm 2.83 to search C and γ. The final determination of C=32, γ =0.0001.

3 Results and Analysis

3.1 The classification results

Combining the optimal texture characteristics and index characteristics (normalized differential vegetation index, NDVI, normalized differential water index and normalized construction index) as the band and then merge with multi spectral image, participate in image classification. After repeated experiments, determining the parameters of multi-scale segmentation are: segmentation scale 100, shape parameter 0.4 compactness parameter 0.6, the segmentation results can basically ensure most object boundary extraction correctly. Using visual judgment and GPS select site of typical sample. Finally we collected 305 samples altogether, 175 of which were used as training samples, the other 130 as the test sample. According to the training samples, select the SVM RBF kernel function make object-oriented classification, the classification results are shown in figure 7.

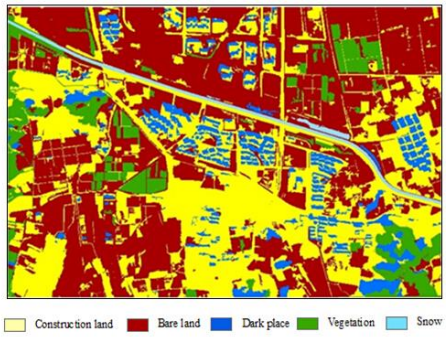


Fig.7. SVM classification results of SPOT6 images

To evaluate the classification results accuracy effectively, classification accuracy as shown in the table2:

Table 2 .The form of SVM classification accuracy evaluation

	Dark place	Vegeta-tion	Bare land	Constru-tion land	Snow	Total	Customer accuracy（%）
Dark place	21	0	1	1	0	22	95.45
Vegetation	0	17	4	0	0	21	80.95
Bare land	0	2	17	6	1	26	65.38
Construction land	4	1	6	45	5	61	73.77
Snow	0	0	0	0	2	2	100.00
Total	25	20	24	51	8		
Producer accuracy（%）	84.00	85.00	70.83	88.24	25.00	84.00	
Overall accuracy：79.70%							
Kappa coefficient：0.717							

Image recognition ability for dark place and vegetation are better than that of bare land and the construction land, construction land and bare land have a certain degree of mixing. SVM classification capability of roads and other linear features is good, even smaller rural road can also be identify effectively.

3.2 The extraction results of new construction land

Before extracting new construction land, we need deal with classification results. There were some shadows of building existing in class of dark place. Using the neighborhood characteristics of eCognition, set the threshold, those can be classified into construction land; and merge adjacent segmentation map spot. According to the minimum on map spot, those witch area less than 666.67 square meters will be deleted. Finally, make differential operation to classification results which were treated and base land use map. The extraction results are shown below.

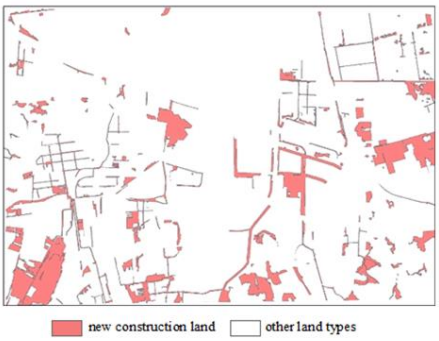


Fig.8. new construction extraction results

Statistical results of new construction land are shown in Table3.

Table 3. Statistics of new construction

Types	The number of polygons	Are of polygons(km^2)	Percentage of area（%）
Construction land	38	117.62	88.61
Bare land	12	10.35	7.80
Dark place	5	4.17	3.14
Snow	2	0.60	0.45
Total	57	132.74	
Misclassification rate			11.39
Leakage rate			3.77

The accurate rate of extraction is more than 80%, achieved superior extraction effect, but there is a certain degree of error and omission phenomenon. Among those new construction land which extracted , the bare land divided into construction land had high misclassification rate, which may be related to the image itself , cause the SPOT6 image as the winter image, vegetation coverage is low, may result in the impact on the classification results.

4 Conclusion and Discussion

(1) Image fusion can compensate for the lack of information in a single image, take advantage of a variety of images, get multi-faceted feature information, is a crucial step of information extraction. The most suitable fusion method for the extraction of construction land was selected by quality evaluation system of fusion image. For SPOT6 image, select the fusion method of PanSharp can improve the extraction accuracy of construction land.

(2) Multi-window texture features take part in image classification can improve the class separability and classification accuracy effectively. SPOT6 image, choose the texture features homogeneity of 9×9 window and 15×15 window, choose the texture features entropy of 15×15 window.

(3) The classification method of SVM can solve the small sample, nonlinear and high dimensional problems, obtain global optimal solution, it shown the superiority in the application of road and building.

(4) The SPOT6 image in this paper was winter data, vegetation coverage rate was low, and there are snow covered, this may cause certain effect on the extraction of construction land information. We should select more data to verify the reliability of this method.

Acknowledgment

This research was founded by the Beijing Municipal Science and Technology Projects (Z141100000614001), and the projects of Hebei Provincial Seismological Bureau (DZ20150428102, DZ2015017012).

References

1. Bouziani M., K. Goïta, D.-C. He. Automatic change detection of buildings in urban environment from very high spatial resolution images using existing geodatabase and prior knowledge [J]. ISPRS Journal of Photogrammetry and Remote Sensing. 2010, 65(1): 143-153.
2. Huang Yan, Zhang Chao, Su Wei, Yue An-zhi. A study of the optimal scale texture analysis for remote sensing image classification [J]. Remote sensing for land & resource. 2008, (4):14-19.
3. Huang X, Zhang L P and Li P X. Classification of high spatial resolution remotely sensed imagery based upon fusion of multi-scale features and SVM [J]. Journal of Remote Sensing, 2007, 11(1): 48–54.
4. Zhang Lei, Shao Zhen-feng, Ding Lin. Semi-supervised Collaborative Classification for Hyperspectral Remote Sensing Image with Combination of Cluster Feature and SVM[J].Acta Geodaetica et Cartographica Sinica, 2014,43(8): 855–861.
5. Huang He, Feng Yi, Zhang Meng. Research on fusion of mapping satellite-1 image and its evaluation [J]. Bulletin of Surveying and Mapping, 2013, 2013(1): 6–9.
6. Mariz C., D. Gianelle, L. Bruzzone, et al. Fusion of multi-spectral SPOT-5 images and very high resolution texture information extracted from digital orthophotos for automatic classification of complex Alpine areas [J]. International Journal of Remote Sensing. 2009, 30(11):2859-2873.
7. Liu Meng-meng, Liu Yalan, Sun Guoqing. SVM land Cover classification based on spectral and textural ferture using stratified samples [J]. Remote Sensing Technology and Application, 2014, 29 (2):315-323.
8. Li Jin-lian, Liu Xiaomei, Li Hengpeng. Extraction of Texture Feature and Identification Method of Landuse Informati on from SPOT5 Image [J]. Journal of Remote Sensing, 2006, 10(6): 926 – 931.
9. Huang Xi, Zhang Liangpei, Li Pingxiang. Classification of High Spatial Resolution Remotely Sensed Imagery Based Upon Fusi on of Multiscale Features and SVM [J]. Journal of Remote Sensing, 2007, 11(1): 48-54.
10. Zhang Feng, Xue Yan-li, Li Ying-cheng. Object-oriented Building Extraction of Multi-source Remote Sensing Imagery Based on SVM [J]. Remote Sensing For Land &Resource. 2008, 25(2): 27-29.
11. Wu B, Xiong Z G, Chen Y Z, Zhao Y D. Classification of quickbird image with maximal mutual information feature selection and support vector machine [J]. Procedia Earth and Planetary Science, 2009, 1(1):1165-1172.
12. He Ling-min, Shen Zhang-quan, Kong Fan-sheng, Liu Zhen-ke. Study on Multi-source Remote Sensing Images Classification with SVM [J]. Journal of Image and Graphics, 2007, 12(4): 648-654.