Frontiers of Information Technology & Electronic Engineering
www.zju.edu.cn/jzus; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Urban landscape classification using Chinese advanced high-resolution satellite imagery and an object-oriented multi-variable model*

Li-gang MA[†1], Jin-song DENG[†1,2], Huai YANG[1], Yang HONG[2,3], Ke WANG[†‡1]

([1]*Institute of Applied Remote Sensing & Information Technology, Zhejiang University, Hangzhou 310058, China*)

([2]*School of Civil Engineering and Environmental Sciences and School of Meteorology, University of Oklahoma, Norman, OK 73019, USA*)

([3]*State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China*)

[†]E-mail: 11114054@zju.edu.cn; jsong_deng@zju.edu.cn; kwang@zju.edu.cn

Received Mar. 7, 2014; Revision accepted May 30, 2014; Crosschecked Feb. 4, 2015

Abstract: The Chinese ZY-1 02C satellite is one of the most advanced high-resolution earth observation systems designed for terrestrial resource monitoring. Its capability for comprehensive landscape classification, especially in urban areas, has been under constant study. In view of the limited spectral resolution of the ZY-1 02C satellite (three bands), and the complexity and heterogeneity across urban environments, we attempt to test its performance of urban landscape classification by combining a multi-variable model with an object-oriented approach. The multiple variables including spectral reflection, texture, spatial autocorrelation, impervious surface fraction, vegetation, and geometry indexes were first calculated and selected using forward stepwise linear discriminant analysis and applied in the following object-oriented classification process. Comprehensive accuracy assessment which adopts traditional error matrices with stratified random samples and polygon area consistency (PAC) indexes was then conducted to examine the real area agreement between a classified polygon and its references. Results indicated an overall classification accuracy of 92.63% and a kappa statistic of 0.9124. Furthermore, the proposed PAC index showed that more than 82% of all polygons were correctly classified. Misclassification occurred mostly between residential area and barren/farmland. The presented method and the Chinese ZY-1 02C satellite imagery are robust and effective for urban landscape classification.

Key words: ZY-1 02C satellite, Classification, Urban, Multi-variable model
doi:10.1631/FITEE.1400083 **Document code:** A **CLC number:** TP751.1

1 Introduction

Remote sensing technology has been widely applied in various fields and is now one of the most efficient means of studying urban land cover, structure, landscape, and the eco-environment. With the fast-pace changes in urban landscapes, particularly in rapidly expanding metropolitan regions, the acquirement of frequently updating land-use datasets has become a routine assignment. Consequently, deriving timely and accurate land-use information to keep pace with urban development is a critical challenge confronting urban planners (Hu and Wang, 2013). Requirements of land-use mapping and monitoring by means of automated classification on remote sensing imagery have played an increasingly important role in decision-making and urban management.

In terms of remote sensing technology, Earth observing satellites can acquire abundant image data, ranging from high resolutions (i.e., 0.5–1 m, such as GeoEye, WorldView, IKONOS, and QuickBird) to moderate resolutions (i.e., 10–30 m, such as SPOT,

‡ Corresponding author

* Project supported by the Chinese Ministry of Environmental Protection (No. STSN-05-11), Zhejiang Key Scientific and Technological Innovation Team Projects (No. 2010R50030), and the National Natural Science Foundation of China (No. 31172023)

ORCID: Li-gang MA, http://orcid.org/0000-0001-9377-2189

ALOS, ASTER, and LANDSAT) for mapping urban landscapes. However, an overwhelming majority of satellite images are dominated by the United States and European space agencies. In recent years, China has launched a series of high-resolution satellites, among which ZY-1 02C is one of the most advanced. This satellite, however, is equipped with only three spectral bands (green, red, NIR), which may increase the difficulty in discriminating urban targets. There is an urgent need to examine its utility and capabilities in detecting subtle changes in urban land use and land cover. Table 1 shows more in-depth information about this ZY-1 02C satellite imagery.

Table 1 Basic parameters of the ZY-1 02C satellite's P/MS camera

Parameter	Value
Wavelength	
B1 (green)	0.52–0.59 μm
B2 (red)	0.63–0.69 μm
B3 (NIR)	0.77–0.89 μm
Spatial resolution	10 m
Swath	60 km
Side-sway ability	±32°
Revisiting period	3–5 d
Coverage period	55 d
Phase	2012-2-28

Information extraction and classification techniques are other crucial aspects of image analysis. A large number of classification algorithms have been developed during the past decades. From the perspective of image processing, they can be grouped as traditional single pixel-based classification methods and object-oriented classification methods. It is generally stated that object-oriented classification methods outperform pixel-based methods in high-resolution imagery and could be used to extract tangible information which is well-suited to vector geographic information system (GIS) data. Chen *et al.* (2007) demonstrated the potential of object-based image analysis (OBIA) in mapping urban land cover for the city of Beijing from ASTER data with a relatively high accuracy. Durieux *et al.* (2008) proposed precise monitoring of building construction using an object-based classification method applied to SPOT5 images. Jacquin *et al.* (2008) assessed the ability of an object-based approach in classifying urban objects at multiple spatial scales with SPOT5 images. The images acquired by the ZY-1 02C satellite have similar spatial resolution to SPOT and ASTER; urban mapping with this data through an object-oriented approach could probably yield acceptable results. Based on the characteristics of this kind of imagery (relatively high spatial resolution and low spectral resolution), we attempt to explore a method, which integrates five categories of variables (fraction of impervious surface, vegetation index, texture derived from the gray level co-occurrence matrix, spatial autocorrelation, and shape related features) for urban classification with an object-based approach.

2 Methods

2.1 Study area

The study area is the Hangzhou City, capital of Zhejiang Province, China. Hangzhou lies in the southern wing of the Yangtze Delta (Fig. 1), and covers an area of 3068 km^2 (720 km^2 for the city proper) with a population of 3.932 million (Deng *et al.*, 2008). This city has experienced a rapid expansion during the past decade, while its historical and modern buildings coexist due to heritage policy on historical and cultural protection. Since 1978, industrial zones and commercial buildings have sprung up extensively across the Hangzhou City after China's economic reform. The diversity of land-cover makes it an extremely challenging task for classification and an ideal place for urban classification study.

2.2 Data sets used in current research

A multispectral image with 10 m spatial resolution acquired in February, 2012 by the P/MS sensor of the ZY-1 02C satellite was examined in this study. A fine land-use map derived from a visually interpreted high-resolution aerial photo in 2010 was used as the reference data. Field survey data, including GPS positions and photos acquired in 2012, was adopted as additional reference for an assessment of land-use classification accuracy.

2.3 Image processing

Using the referencing on the high-resolution image and the in-field GPS survey, geometric correction was first conducted. Altogether 30 ground control points (GCP) were selected and an RMSE of 0.5 was achieved to guarantee geometric fidelity.

As shown in Fig. 2, our classification scheme is composed of five steps: (1) variable calculation; (2) feature selection; (3) correlation and variable importance analysis; (4) object-oriented classification; (5) accuracy assessment.

2.4 Variable calculation and selection

2.4.1 Fraction images

Ridd (1995) pointed out that urban areas can be divided into three parts, namely impervious surfaces, vegetation, and soil. Theoretically, accurate identification of the percentage of impervious surfaces may contribute to the improvement of urban image classification since this feature should be different for different land-use types. In practice, linear spectral mixture analysis (LSMA) has shown its potential in estimating impervious surfaces and improving urban classification (Lu and Weng, 2006).

Mixed pixel problems, as well as extraction of quantitative sub-pixel information in satellite imagery with medium or coarse spatial resolution, are usually solved by applying the spectral mixture analysis (SMA) technique (Smith *et al.*, 1990). It has been hypothesized that a linear combination of the spectra of all components within a pixel constitutes the spectrum measured by a sensor (Adams *et al.*, 1995). The principle and detailed description can be found in Settle and Drake (1993). In this research, a minimum noise transform (MNF) was applied to the image, and three end-members (impervious surface, vegetation, and shade) were selected from the feature space of MNF components. The residential area was identified for the impervious surface end-member selection, while vegetation was selected from trees and dense grass. The shade end-member was selected from rivers and lakes. The original three-band-multispectral image was unmixed into three fraction images with a least square solution. Impervious surface fractions were finally extracted.

2.4.2 Vegetation indices

Researchers have developed various vegetation indices based on the fact that plants reflect more in the near-infrared band and less in the red band compared with non-vegetated surfaces (Li and Weng, 2005). Despite their limitations, vegetation indices such as NDVI and EVI are highly useful measurements for vegetation identification (Carlson and Ripley, 1997). NDVI is used in this study for urban classification.

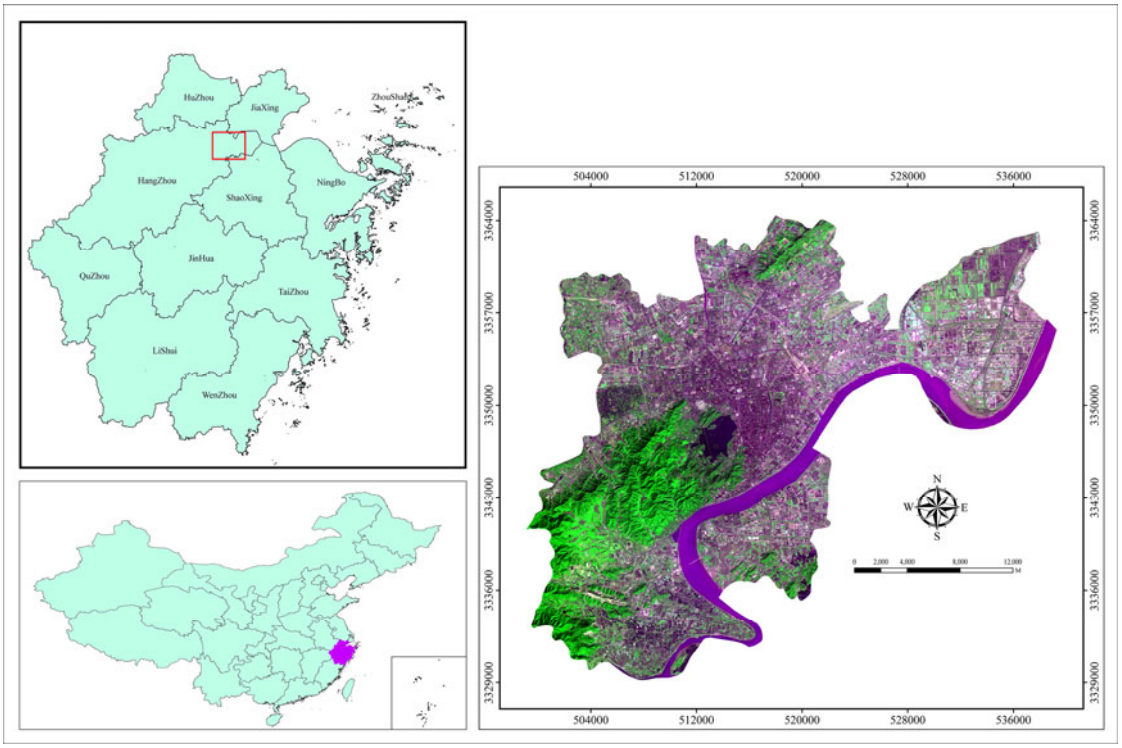

Fig. 1 Geographic location of the Hangzhou City, Zhejiang Province, China

This figure shows the location of Hangzhou in Zhejiang Province, China. The main urban area includes six districts

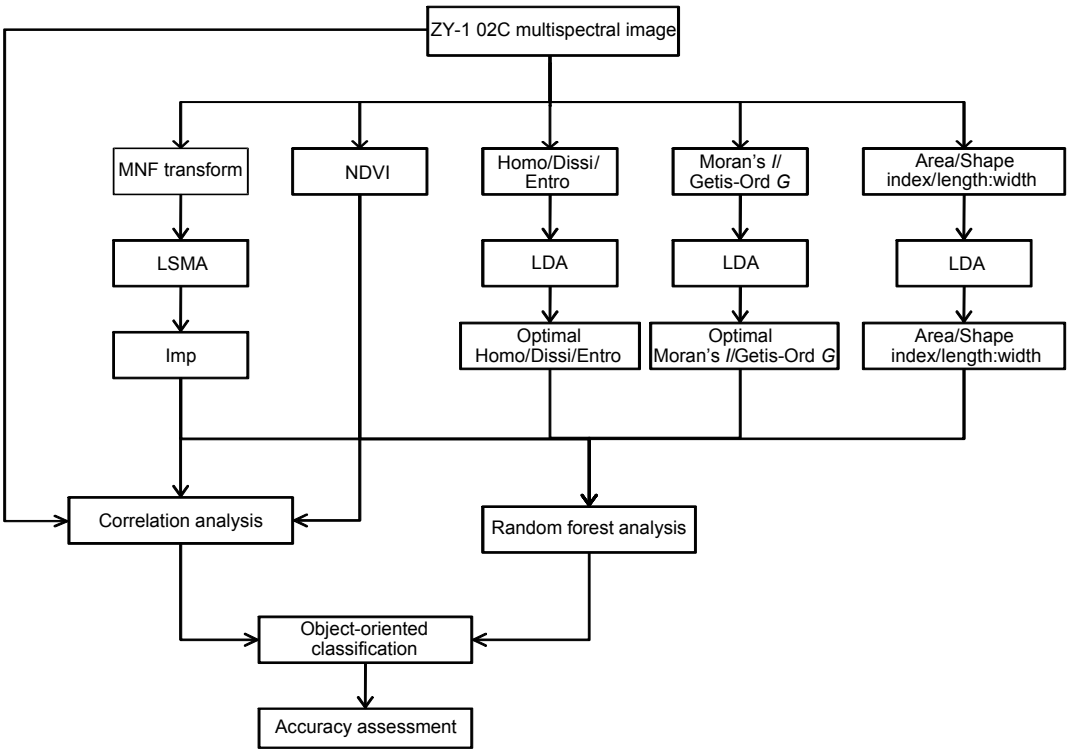

Fig. 2 Flowchart of the proposed method consisting of impervious surface (Imp) extraction, homogeneity/dissimilarity/ entropy (Homo/Dissi/Entro) and Moran's *I*/Getis-Ord *G* calculation, linear discriminant analysis (LDA), etc.

2.4.3 Texture images

Lu and Weng (2007) believed that the most significant variables for remote sensing image classification were spectral information. However, texture and context information should never be ignored, especially when high-resolution remote sensing images are used. Texture measures such as the gray level co-occurrence matrix (GLCM) and spatial statistic based signal analysis have already been developed and applied in remote sensing image classification. Some researchers determined that the most useful features for analyzing the content of remote sensing imagery were those variables calculated using a GLCM approach (Shanmugan *et al.*, 1981; Pacifici *et al.*, 2009). So, three texture features derived from the GLCM have been considered in this work, namely homogeneity, dissimilarity, and entropy. Homogeneity measures the composition of similar images while entropy indicates the disorder or heterogeneity in an image. Dissimilarity measures the difference among elements of the co-occurrence matrix from each other. Their formulations are shown in the following:

$$\text{Homogeneity}=\sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\frac{p(i,j)}{1+(i-j)^2},$$

$$\text{Dissimilarity}=\sum_{i=1}^{N-1}\sum_{j=1}^{N-1}p(i,j)\cdot|i-j|,$$

$$\text{Entropy}=-\sum_{i=1}^{N-1}\sum_{j=1}^{N-1}p(i,j)\cdot\log(p(i,j)),$$

where i and j are indexes of the gray tones in the windows or the elements of the co-occurrence matrix, $p(i, j)$ indicates the normalized frequency at which two neighboring resolution cells separated by a fixed shift occur on the image (one with gray tone i and the other with gray tone j), and N represents the dimension of the co-occurrence matrix.

In Earth sciences, many properties exhibit spatial clustering of similar values around an individual location. The local patterns of spatial association or spatial autocorrelation can be identified by local indicators of spatial association (LISA) measures. Local Moran's I and Getis-Ord Local G are two indices of spatial autocorrelation and they are computed for the three spectral bands as additional textural

information in our classification. They are defined as (Han *et al.*, 2012)

$$\text{Local Moran's } I_i(d)=(x_i-\bar{x})\sum_{j=1}^{n}w_{ij}(d)(x_j-\bar{x}),$$

$$\text{Local Getis } G_i(d)=\frac{\sum_{j\neq i}w_{ij}(d)x_j}{\sum_{j\neq i}x_j},$$

where x_i and x_j are the values of variable x at positions i and j respectively, $\bar{x}$ is the mean of variable x, $w_{ij}(d)$ is the weight between positions i and j within distance d, and n is the sample size.

A positive value for Local Moran's I indicates that the feature is surrounded by features with similar values; such a feature is part of a cluster. A negative value for it indicates that the feature is surrounded by features with dissimilar values; such a feature is an outlier. Getis-Ord Local G is useful for determining clusters of similar values where clusters of high values result in a high G value and clusters of low values result in a low G value (Shahtahmassebi *et al.*, 2014).

As for the window size, Shaban and Dikshit (2001) concluded that high-resolution images usually need larger window sizes than low-resolution images. Considering the 10 m spatial resolution of our image, all three bands with window sizes of 3×3, 5×5, 7×7, 9×9 are calculated. A lag distance from 1 to 9 is also calculated using Rook's case.

2.4.4 Shape related features

The urban area is a complicated entity composed of many different types of constructions and natural objects. Different targets may share the same reflectance but have distinct shapes, for example, a building and road. Concerning its nature in external appearance, we choose area, shape index, and length/width as three variables for an object-based classification. Detailed descriptions of the three indexes can be found in the reference book of eCognition Developer 8.7 (Trimble Germany GmbH, 2012).

2.4.5 Feature selection

Wu *et al.* (2012) pointed out that the classification accuracy may be significantly decreased when all possible features are used in a classification procedure. Feature selection seems to be inevitable. Forward stepwise linear discriminant analysis (LDA) is applied to assess usefulness and the influences of the proposed descriptive features for our classification problem. In this method, the features that contribute most to land cover classification are determined by reviewing and evaluating all variables at each step. The model then includes that variable and the process is iterated (Hermosilla *et al.*, 2012).

2.4.6 Variable importance measure

To better understand the contribution of each variable in the classification, a random forest algorithm is adopted. The random forest machine learner is a meta-learner, which means it consists of many individual learners (trees). The random forest uses multiple random tree classifications to vote on an overall classification for the given set of inputs (Livingston, 2005). It provides a ranking of variable relevance by comparing classification accuracies obtained with, and then without, each of the features (Novack *et al.*, 2011). The reader is referred to Livingston (2005) for a detailed description of this algorithm.

2.5 Classification

Object-oriented classification is usually deemed to be superior to a per-pixel classification approach because image objects can represent meaningful information while single pixels cannot (Mathieu *et al.*, 2007). Therefore, better classification results can be acquired, especially for fine spatial resolution data. We use eCognition Developer 8.7 to perform object-based classification. With this software, appropriate values need to be assigned to the three key parameters, namely shape, compactness, and scale, in the object-based paradigm. The most crucial parameter for image segmentation is scale, which controls the object size. By modifying the shape criterion, the color criterion is indirectly defined (color=1−shape) to change the relative weighting of the reflectance and shape in defining the segments. The shape criterion is composed of two parameters, compactness and smoothness. The smoothness criterion is used to optimize image objects with regard to smoothness of borders, while the compactness criterion is used to optimize image objects with regard to compactness. To pay more attention to spectrally homogeneous pixels for image segmentation, we give a smaller weight to shape and set the shape parameter to be 0.1.

Compactness and smoothness parameters are both set to 0.5 for a trade-off. Based on a trial-and-error analysis, we find that a scale parameter of 5 is appropriate for this study. After segmentation, the support vector machine (linear kernel) algorithm is applied for classification of the whole image. Note that the image is firstly classified into 26 classes according to the reflectance of each land cover. Then all classes are merged into seven classes with reference to our classification system. The system is built based on a U.S. geological survey land-use/land-cover classification system for use with remote sensor data modified for the national land cover dataset and an NOAA coastal change analysis program. According to the specific situation of the Hangzhou City, some revisions have been made to make it more suitable for a study in China. Our classification system is composed of seven classes: Residential, Commercial/Industrial/Transportation, Forest, Groves, Water, Barren, and Farmland.

2.6 Accuracy assessment

To perform the object-based accuracy assessment, a total of 760 (no less than 50 for each class) land-cover polygons or objects are randomly selected from the studied site in the Hangzhou City. The polygons are visually interpreted with reference to the land-use map and field survey data and compared with the classification results derived from the object-based approach. Note that the 26 original classes are grouped into seven classes before computing the accuracy indexes.

The predicted classes derived from the classifier and visually interpreted classes with reference to the land-use map are compared by means of the error matrix. Several indexes such as producer's accuracy (PA), user's accuracy (UA), overall accuracy (OA), and the kappa statistic are calculated from this matrix. UA implies commission errors (i.e., when an object is committed to an incorrect class), PA details omission errors (i.e., when a segment is incorrectly classified into another category and thus omitted from its correct class), and OA is the ratio of the number of correctly classified objects to the total number of objects. Finally, the off-diagonal observations of the rows and columns and diagonal values of the error matrix are incorporated by the kappa statistic so that a more robust accuracy assessment can be made (Aguilar *et al.*, 2013).

Besides pixel-based accuracy assessment, the area residual approach can also be used to evaluate the classification accuracy. Forty-five 50×50 sample plots (i.e., 500 m×500 m) are selected from the classified image. We select plots from intersections between straight lines and cycles, with a distance of 200 pixels between two cycles (Fig. 3). Reference plots are acquired by visual interpretation of each square with reference to the land-use map in which several classes are included. Finally, classified plots and reference plots are compared using the polygon area consistency (PAC) index. A formula description of this index is

$$W_s=A_i/B_i\times 100\%,$$

where W_s represents the polygon area consistency, A_i the common area between the sample and the reference, and B_i the reference area.

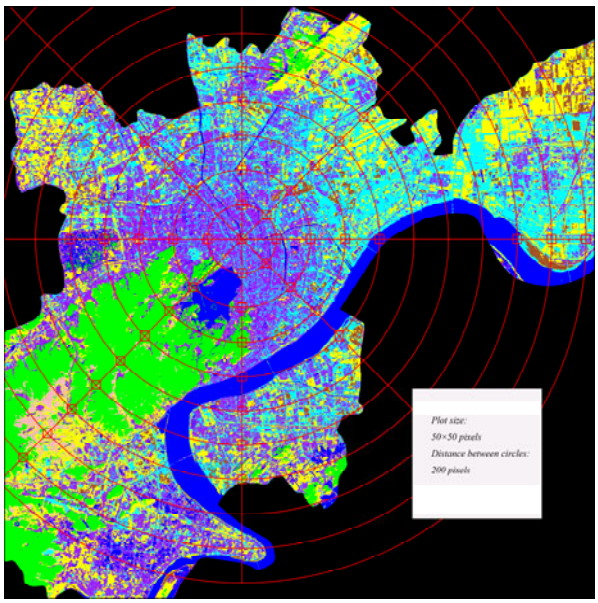

Fig. 3 Strategy for collection of reference data for accuracy assessment of the area, illustrating the approach for allocating sample plots on the classified image

3 Results and discussion

3.1 Feature analysis

For the GLCM and spatial autocorrelation group, the variables selected were: (1) Homo_R_3×3, Dissi_G_9×9, Entro_R_9×9; (2) Moran_G_9, Moran_

NIR_9, Getis_NIR_3 (the first word means texture, the second character implies band, and the number represents the window size). As shown in Fig. 4, 9×9 was suitable for dissimilarity and entropy, but 3×3 was needed for homogeneity. A lag distance of 9 was optimal for Moran's *I*, but 3 was necessary for Getis *G*. Each multispectral band had its own advantages in texture calculation. All variables selected by LDA displayed their contribution to accuracy improvement according to Fig. 4. It was clear that the classification accuracy increased steadily when GLCM, spatial autocorrelation, and three shape indexes were gradually included. Fig. 4 shows that there was a dramatic increase when three multispectral bands were included one by one. However, no obvious increase can be found when NDVI and the fraction of impervious surface images were added. This could be explained as follows: (1) Information contained in NDVI and the fraction images almost overlapped with the original bands or they were useless in the classification; (2) Importance of the two variables was far below that of the three original bands, so the influence cannot be illustrated even if they made some tiny contribution; (3) They did have their function and that influence was really large, but it cannot be answered by linear projection no matter how many discriminant models were built. To understand the exact reasons, a correlation analysis of the first 11 variables is presented in the heat map (Fig. 5).

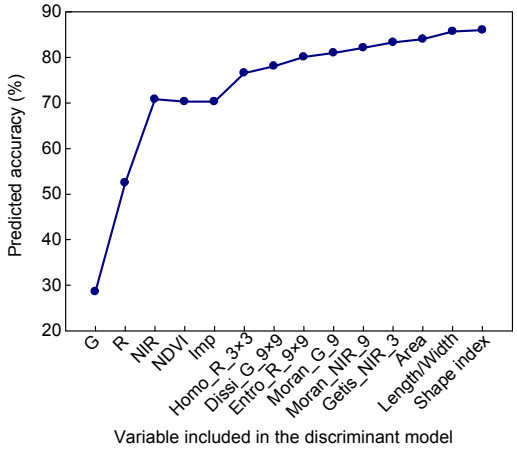

Fig. 4 Predicted cross-validated classification accuracy when the 14 features were progressively included in the discriminant model

As can be seen from Fig. 5, there was a high correlation between the green and red bands, and the correlation coefficient can be higher than 0.8. Correlation between the red band and the fraction image of the impervious surface (Imp) was similar to that of the multispectral bands. NDVI had little correlation with the near-infrared band or low correlation with most of the other variables. Therefore, we deemed that information overlap was not a core problem even though correlation between the red band and the impervious surface fraction reached a high level, not to disregard the low correlation between NDVI and the other variables. An assumption of information overlap can be excluded. In addition, spatial autocorrelation variables (Moran_NIR_9 and Moran_G_9) exhibited little correlation with almost all variables.

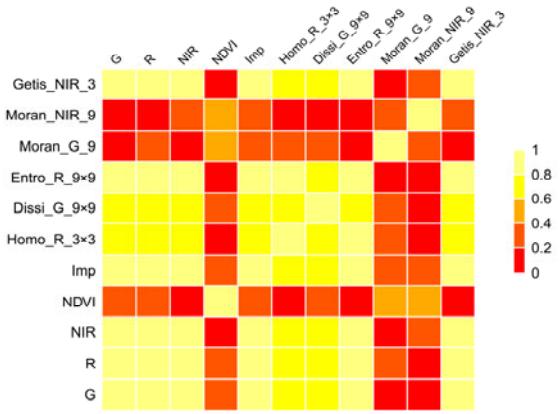

Fig. 5 Correlation of 11 variables indicated by the correlation coefficient

It was evident in Fig. 6 that the spatial autocorrelation measures are the most important variables, especially for Moran's *I* calculated from the near-infrared band with a lag distance of 9. GLCM based variables not obtaining a high rank may be attributed to the relatively high correlation between these variables and the original spectral bands. NDVI and the impervious surface fraction were two variables that had almost the same significance as the original three bands according to this algorithm. This was easily understood because NDVI was originally designed for vegetation detection, and impervious surface fraction indicates the percentage of the impervious surface in every pixel. Every land cover was normally characterized by its own impervious surface distribution. Normally, the impervious surface of the industrial section should be higher than that of the residential area to gain as much profit as possible. As

expected, the three shape-related features appeared to be the least significant among the whole data set. The reason for spatial autocorrelation outperforming other variables may be due to the fact that the city area was mostly covered by residential buildings, which in fact were auto-correlated in every block. Another possible reason would be that the land-use pattern in Hangzhou was highly policy oriented; in other words, every construction project should be consistent with land-use planning, which was usually spatially auto-correlated in a small area.

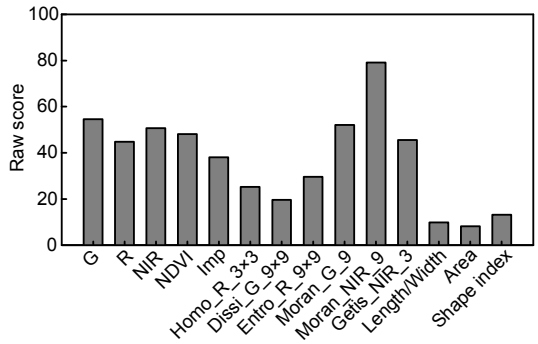

Fig. 6 Variable importance of the 14 features provided by the random forest algorithm

As mentioned above, NDVI and impervious surface faction were essential variables that had low correlation with the three original bands. Although a high correlation was found between the green and red bands, none of them were abandoned because a small discrepancy may imply a huge power of discrimination for the land cover. Thus, all of the 14 variables (multispectral bands included) were applied in subsequent object-oriented classification.

3.2 Classification analysis

Feature analysis was based on statistics of samples which depended on accurate selection of the variables. Classification provided an approach for demonstrating the usefulness of that method. As shown in Fig. 7 and Table 2, the overall classification accuracy was satisfactory. Most classes were correctly classified owning to those selected important variables. However, as mentioned before, Hangzhou is a historical, developing city. Ancient architectures, modern commercial buildings, modern residential apartments, private houses, and rural housing mixed together generated a great challenge. Second, since

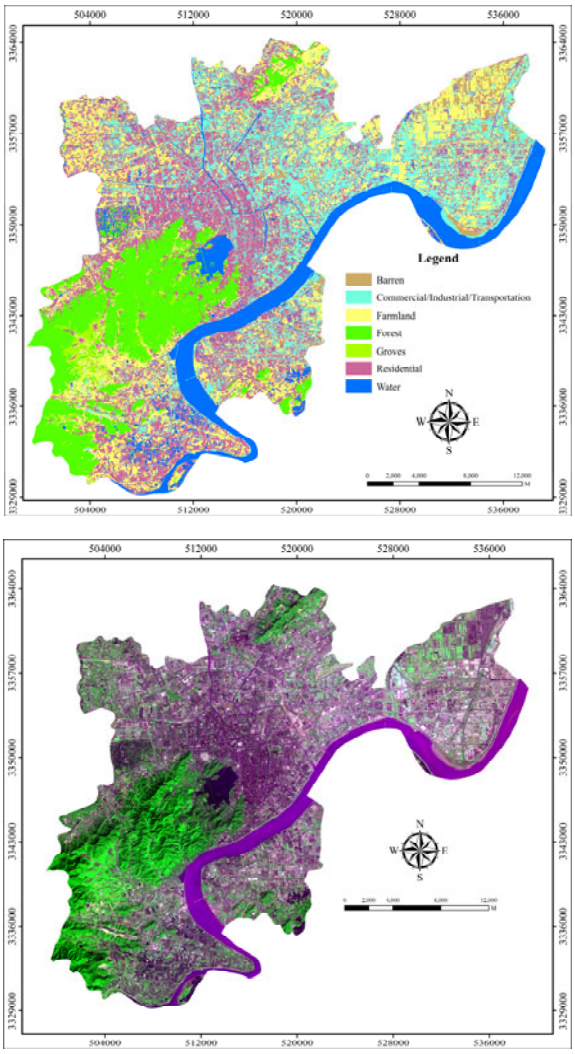

Fig. 7 The original image (R: 2, G: 3, B: 1) and land use/cover map of the Hangzhou City produced by object-oriented classification

this image was captured in February 2012, a shadow casted by high buildings and mountains was also a factor that cannot be ignored. We merged shadow into residential buildings in this study, which contributed to part of the errors. Moreover, water is widely spread throughout the southern part of the city with various forms like rivers, lakes, ponds, and paddy fields. A large discrepancy could be found between different forms of water due to diversities of reflectance from the water surface. Similarly, for shadow, different forms of water and dark roof buildings led to erroneous classification.

Examining the error matrix (Table 2) derived from the object-oriented classification approach, we

Table 2 Overall accuracy, producer's accuracy (PA), user's accuracy (UA), and the kappa statistic produced by the object-oriented classifier

Classified data	Reference								PA (%)	UA (%)
	Residential	Commercial/Industrial/ Transportation	Forest	Groves	Water	Barren	Farmland	Total		
Residential	153	7	2	0	5	2	2	171	94.44	89.47
Commercial/Industrial/ Transportation	1	131	0	0	2	3	2	139	90.34	94.24
Forest	0	0	110	3	1	0	1	115	94.02	95.65
Groves	0	0	0	56	0	0	0	56	91.80	100.00
Water	0	0	0	0	91	0	0	91	91.92	100.00
Barren	3	3	0	0	0	55	0	61	87.30	90.16
Farmland	5	4	5	2	0	3	108	127	95.58	85.04
Total	162	145	117	61	99	63	113	760		

Overall classification accuracy=92.63%; overall kappa statistic=91.24%

found that water and groves had the highest user's accuracy while farmland the highest producer's accuracy. This could be attributed to its characteristic of high NDVI value. Next to vegetation and water land-use, user's accuracy of commercial/industrial/ transportation was the second highest. This could be explained by the high response characteristic in the impervious surface fraction, area differences, and autocorrelation distinction. Also, industrial, civic, and office land use usually tend to cluster, to gain the maximum benefit, a situation that is expected to occur more in the future. In contrast to office and industrial land use, civic land use is a type of non-profit land use (Hu and Wang, 2013). Here, the producer's accuracy and user's accuracy of barren were the lowest. Table 2 shows that it is difficult to distinguish the barren area from the residential area and commercial/industrial/ transportation area. The reason can be found from Fig. 7, where the densely distributed residential buildings were easily misclassified into barren area for their resemblance in both texture and reflectance. Roads located in industrial area resulted in barren misclassification. Some confusion also existed between residential area and commercial/industrial/ transportation area. The low user's accuracy of the residential class was indicated by the misclassification of commercial/industrial/transportation into residential. This problem occurs between narrow roads with low digital numbers and dark roof residential buildings.

The error matrix shows that farmland parcels had a high degree of confusion with residential and forests. As shown in Fig. 7, this problem emerged in the northern mountain area in which human planted trees and natural ones grow together, forming its own texture. Other possible reasons for the low accuracy of farmland could be the error of the land-use map or land cover change from the fall of 2010 to the spring of 2012. A proper method that may solve this problem is to use topographic data. Forests usually show their appearance on mountains while farmland normally occurs on the planes. Moreover, as Hangzhou is a tourist city, vegetation coverage is a significant factor for government decision-making. Some residential blocks featured a high level of vegetation coverage and were reckoned as farmland by the classifier.

Water was a third major error (omission error represented by relatively low producer's accuracy) in the classification. In other areas, water could not be a problem. However, as discussed before, Hangzhou is a typical southern city covered with various kinds of water including river, lake, pond, and wetland. The reasons for the confusion between water and residential could be summarized as follows:

1. From Fig. 7, south Hangzhou was primarily covered by ponds. The DN value and the shape index of the dark residential resemble that of ponds to a large extent.

2. Rivers usually develop their own branches, and some of these branches are so narrow that they cannot be segmented even with a very small scale parameter. Average reflectance of these objects looked like those of dark residential segments.

3. The shadow, whose reflectance was similar to water in an urban area, was merged into residential, which influences accuracy.

3.3 Area analysis

The highest accuracy was achieved for identification of forest according to the area evaluation index (Fig. 8). This was in agreement with the point-based accuracy assessment, indicating the robustness of the proposed method in forest detection. According to the PAC index, farmland attained the second highest accuracy. This was opposite to its low user's accuracy in the error matrix. A probable reason could be that this class was overestimated in the final classification. Overestimation of this class resulted in a larger common area, which improved the performance of this index and reduced the classification accuracy that could be illustrated by a point-based accuracy assessment. Residential and water were two classes that displayed the lowest accuracy due to their complicated forms of distribution and shortage of the blue band, because many researchers believed that the normalized difference of the built-up index (NDBI) was useful for extraction of the built-up area in which the blue band was included. Barren was not a widely distributed class. The area evaluation index of this class easily suffered from its small total sample area. Therefore, its final accuracy will be affected by several large segments. For future work we could collect more barren samples for area analysis. The superiority of the sampling scheme adopted in our study was reflected by its randomness and representativeness of all the eight directions.

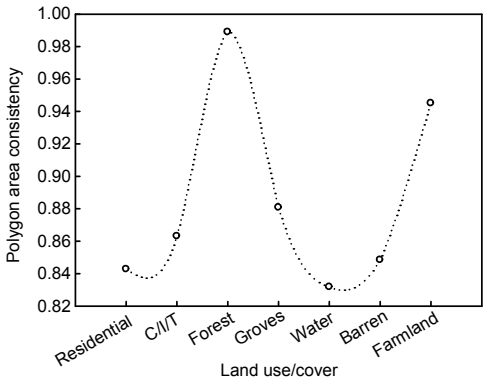

Fig. 8 Area evaluation results (C/I/T means commercial/industrial/transportation)

Overall, the area evaluation result was satisfactory though the accuracies of a very few classes were not high enough. The area precision of most classes exceeded 82%, which was deemed acceptable for most applications.

4 Conclusions

This research is, to our knowledge, one of the few investigations of urban landscape classification using China's ZY-1 02C satellite multispectral image in the Hangzhou City. Although this satellite has only three bands, a relatively satisfactory result is obtained using an object-oriented approach with the combination of various categories of variables selected by LDA. Variable importance analysis illustrated that the spatial autocorrelation is of the most significance during the classification, which could be attributed to land-use planning and construction characteristics of this city. Shortcomings of this research include barren, farmland, and residential misclassifications, which need to be overcome using data mining techniques.

Note that this study focuses mainly on exploring the feasibility of integration of different categories of variables in obtaining an acceptable classification result by means of the Chinese ZY-1 02C imagery. Thus, a random forest was used instead of comparing different classification results with disparate variables. Area analysis with the PAC index proved that an acceptable area accuracy could be attained; meanwhile, overestimation of farmland and the relatively low accuracy of residential or water need to be solved using other approaches. Statistical analysis with various combinations of variables and their role in land cover recognition is left to our future work.

In conclusion, we have proposed a method that integrates five categories of variables using the combination of LDA and the random forest algorithm. Results derived from this method illustrated its effectiveness for urban landscape classification, and revealed the high quality of the Chinese ZY-1 02C multispectral image. Variable importance extracted from GLCM, spatial autocorrelation, and a shape related index proved the reliability of data in texture calculation. Future work can focus on inclusion of other variables for detailed classification or for comparison of the classification accuracy between

this data and other similar spatial/spectral resolution data such as SPOT.

References

Adams, J.B., Sabol, D.E., Kapos, V., *et al.*, 1995. Classification of multispectral images based on fractions of endmembers: application to land-cover change in the Brazilian Amazon. *Remote Sens. Environ.*, **52**(2):137-154. [doi:10.1016/0034-4257(94)00098-8]

Aguilar, M.A., Saldaña, M.M., Aguilar, F.J., 2013. GeoEye-1 and WorldView-2 pan-sharpened imagery for object-based classification in urban environments. *Int. J. Remote Sens.*, **34**(7):2583-2606. [doi:10.1080/01431161.2012.747018]

Carlson, T.N., Ripley, D.A., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sens. Environ.*, **62**(3):241-252. [doi:10.1016/S0034-4257(97)00104-1]

Chen, Y., Shi, P., Fung, T., *et al.*, 2007. Object-oriented classification for urban land cover mapping with ASTER imagery. *Int. J. Remote Sens.*, **28**(20):4645-4651. [doi:10.1080/01431160500444731]

Deng, J.S., Wang, K., Deng, Y.H., *et al.*, 2008. PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *Int. J. Remote Sens.*, **29**(16):4823-4838. [doi:10.1080/01431160801950162]

Durieux, L., Lagabrielle, E., Nelson, A., 2008. A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. *ISPRS J. Photogr. Remote Sens.*, **63**(4):399-408. [doi:10.1016/j.isprsjprs.2008.01.005]

Han, N., Wang, K., Yu, L., *et al.*, 2012. Integration of texture and landscape features into object-based classification for delineating Torreya using IKONOS imagery. *Int. J. Remote Sens.*, **33**(7):2003-2033. [doi:10.1080/01431161.2011.605084]

Hermosilla, T., Ruiz, L.A., Recio, J.A., *et al.*, 2012. Assessing contextual descriptive features for plot-based classification of urban areas. *Landscape Urban Plan.*, **106**(1):124-137. [doi:10.1016/j.landurbplan.2012.02.008]

Hu, S., Wang, L., 2013. Automated urban land-use classification with remote sensing. *Int. J. Remote Sens.*, **34**(3):790-803. [doi:10.1080/01431161.2012.714510]

Jacquin, A., Misakova, L., Gay, M., 2008. A hybrid object-based classification approach for mapping urban sprawl in periurban environment. *Landscape Urban Plan.*, **84**(2):152-165. [doi:10.1016/j.landurbplan.2007.07.006]

Li, G., Weng, Q., 2005. Using Landsat ETM+ imagery to measure population density in Indianapolis, Indiana, USA. *Photogr. Eng. Remote Sens.*, **71**(8):947. [doi:10.14358/PERS.71.8.947]

Livingston, F., 2005. Implementation of Breiman's random forest machine learning algorithm. ECE591Q Machine Learning Journal Paper.

Lu, D., Weng, Q., 2006. Use of impervious surface in urban land-use classification. *Remote Sens. Environ.*, **102**(1-2):146-160. [doi:10.1016/j.rse.2006.02.010]

Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.*, **28**(5):823-870. [doi:10.1080/01431160600746456]

Mathieu, R., Freeman, C., Aryal, J., 2007. Mapping private gardens in urban areas using object-oriented techniques and very high-resolution satellite imagery. *Landscape Urban Plan.*, **81**(3):179-192. [doi:10.1016/j.landurbplan.2006.11.009]

Novack, T., Esch, T., Kux, H., *et al.*, 2011. Machine learning comparison between WorldView-2 and QuickBird-2-simulated imagery regarding object-based urban land cover classification. *Remote Sens.*, **3**(10):2263-2282. [doi:10.3390/rs3102263]

Pacifici, F., Chini, M., Emery, W.J., 2009. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sens. Environ.*, **113**(6):1276-1292. [doi:10.1016/j.rse.2009.02.014]

Ridd, M.K., 1995. Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. *Int. J. Remote Sens.*, **16**(12):2165-2185. [doi:10.1080/01431169508954549]

Settle, J.J., Drake, N.A., 1993. Linear mixing and the estimation of ground cover proportions. *Int. J. Remote Sens.*, **14**(6):1159-1177. [doi:10.1080/01431169308904402]

Shaban, M.A., Dikshit, O., 2001. Improvement of classification in urban areas by the use of textural features: the case study of Lucknow City, Uttar Pradesh. *Int. J. Remote Sens.*, **22**(4):565-593. [doi:10.1080/01431160050505865]

Shahtahmassebi, A., Pan, Y., Lin, L., *et al.*, 2014. Implications of land use policy on impervious surface cover change in Cixi County, Zhejiang Province, China. *Cities*, **39**:21-36. [doi:10.1016/j.cities.2014.02.002]

Shanmugan, K.S., Narayanan, V., Frost, V.S., *et al.*, 1981. Textural features for radar image analysis. *IEEE Trans. Geosci. Remote Sens.*, **GE-19**(3):153-156. [doi:10.1109/TGRS.1981.350344]

Smith, M.O., Ustin, S.L., Adams, J.B., *et al.*, 1990. Vegetation in deserts: I. A regional measure of abundance from multispectral images. *Remote Sens. Environ.*, **31**(1):1-26. [doi:10.1016/0034-4257(90)90074-V]

Trimble Germany GmbH, 2012. eCognition Developer 8.7.1: Reference Book. Documentation, T. Munich, Germany.

Wu, B., Wang, X., Shen, H., *et al.*, 2012. Feature selection based on max-min-associated indices for classification of remotely sensed imagery. *Int. J. Remote Sens.*, **33**(17):5492-5512. [doi:10.1080/01431161.2012.663111]