A Hierarchical Object-oriented Urban Land Cover Classification Using WorldView-2 Imagery and Airborne LiDAR data

M F Wu [1, 2]**, Z C Sun** [2, 3]**, B Yang** [1]**, S S Yu** [2, 4]

[1] Hunan Normal University, Changsha Hunan 410006, China

[2] Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China

[3] Hainan Key Laboratory Earth Observation, Sanya Hainan 572029, China

[4] Shandong University of Science and Technology, Qingdao Shandong 266510, China

Email: sunzc@radi.ac.cn

Abstract. In order to reduce the "salt and pepper" in pixel-based urban land cover classification and expand the application of fusion of multi-source data in the field of urban remote sensing, WorldView-2 imagery and airborne Light Detection and Ranging (LiDAR) data were used to improve the classification of urban land cover. An approach of object-oriented hierarchical classification was proposed in our study. The processing of proposed method consisted of two hierarchies. (1) In the first hierarchy, LiDAR Normalized Digital Surface Model (nDSM) image was segmented to objects. The NDVI, Costal Blue and nDSM thresholds were set for extracting building objects. (2) In the second hierarchy, after removing building objects, WorldView-2 fused imagery was obtained by Haze-ratio-based (HR) fusion, and was segmented. A SVM classifier was applied to generate road/parking lot, vegetation and bare soil objects. (3) Trees and grasslands were split based on an nDSM threshold (2.4 meter). The results showed that compared with pixel-based and non-hierarchical object-oriented approach, proposed method provided a better performance of urban land cover classification, the overall accuracy (OA) and overall kappa (OK) improved up to 92.75% and 0.90. Furthermore, proposed method reduced "salt and pepper" in pixel-based classification, improved the extraction accuracy of buildings based on LiDAR nDSM image segmentation, and reduced the confusion between trees and grasslands through setting nDSM threshold.

Keywords: object-oriented; classification; land cover; WorldView-2; LiDAR

1. Introduction

Very high resolution (VHR) remote sensing images are often used for urban land cover classification at a local scale, because detailed information in urban complex scene is presented in VHR images. But there are lots of challenges for deriving urban land cover types from single source VHR images [1,2], such as different objects having spectral similarity, same objects having spectral difference, shadows of tall buildings and large tree crowns, etc. Single source images can only provide a partial information, e.g. spectral information, of urban objects, however, more comprehensive information of urban area is not presented. Therefore, many researches are attempting to integrate multi-source data for urban land cover mapping. Authors reported that the fusion of moderate resolution images and VHR images was helpful for reducing mixed pixels and improving classification accuracy in large area [3,4]. The combined use of radar data and optical high resolution image were valuable for distinguishing high albedo roads and bare soils, reducing mixed pixels, and identifying objects in the

shadows correctly [2,5,6]. Furthermore, the combination of Light Detection and Ranging (LiDAR) data and VHR optical imagery has been applied to urban land cover mapping in the emergence [7-9]. Because of adding derived height information from LiDAR data, the integration of LiDAR data and VHR images improved the separation of ground objects and non-ground objects, and LiDAR data was effective for extracting buildings and trees more completely and accurately [10,11].

Except for the synergistic use of multi-source data, some advanced approaches were proposed and applied to improve the classification accuracy of urban area. In order to reduce the "salt and pepper" noise in pixel-based classification, object-oriented classification method was used, based on a suitable image segmentation scheme. Object-oriented approach can make use of more comprehensive image information, such as spectral, spatial geometry and texture information and can effectively present the characters of urban objects. Many researches were attempting to use the parametric and non-parametric classifier, hierarchical rule-based technique in object-oriented approach to improve classification accuracy [12,13]. Object-based classification reduced the misclassification from spectral variation of individual objects, spectral similarity from different objects and the "salt and pepper" noise [14].

In our study, a hierarchical object-oriented classification approach based on the combine used of WorldView-2 imagery and airborne LiDAR data was proposed and applied to urban area. A comparison analysis of three approaches, including pixel-based from only WorldView-2 image, non-hierarchical object-oriental from fused data, and proposed hierarchical object-oriented from fused data, was carried out. The objective of this study is to expand the application of object-oriented approach using VHR optical image and airborne LiDAR data in the field of urban land cover classification.

2. Datasets and Method
In this study, an urban area was selected as the study area, located in Pingdingshan City, Henan Province, China. Collected WorldView-2 imagery was acquired on August 21, 2014, consisting of one panchromatic (PAN) band with the 0.5 meter resolution and eight multispectral (MS) bands (including Costal Blue, Blue, Green, Yellow, Red, Red Edge, NIR1, NIR2) with 2 meter resolution. Collected airborne LiDAR data was acquired on August 19, 2013 and was provided in ASCII format, including the X, Y, and Z coordinates and delivered first return intensities. The ground sample distances of LiDAR point clouds ranged from 0.12 to 0.4 meter, and the mean point density was 23.0 points/m^2. LiDAR intensity point clouds provided the height information of image objects, including derived digital surface model (DSM) and digital elevation model (DEM). Normalized DSM (nDSM) image, with 0.5 meter spatial resolution, which was generated by subtracting DSM with DEM, represented the above-ground feature and reduced the impact of the topography (Figure 1).
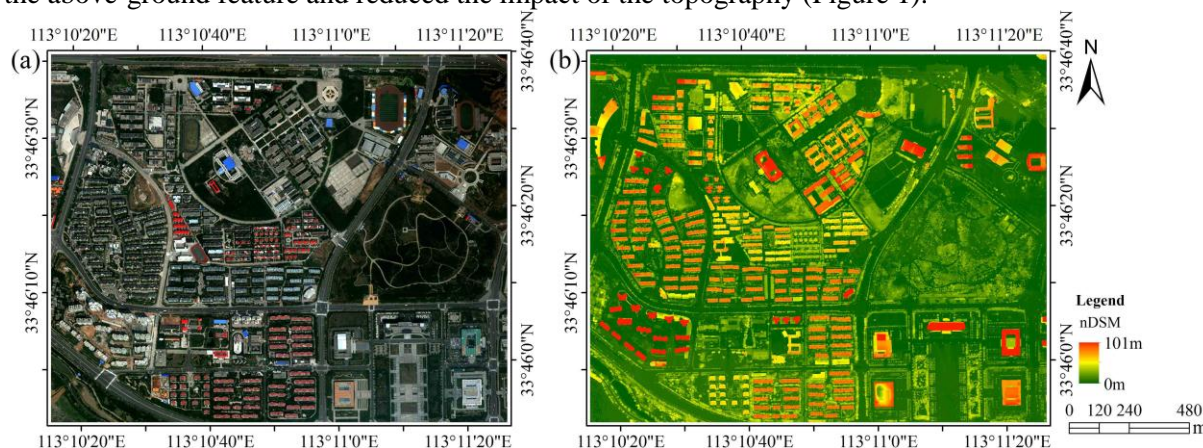

Figure 1. The WorldView-2 and LiDAR nDSM image of the study area.
(a): WorldView-2 fused imagery (band5,band3,band2 compound); (b): LiDAR nDSM data

Figure 2 showed the data procession for classifying urban land cover in this study, including data pre-processing, hierarchical object-oriented classification, and accuracy assessment and classification results analysis.
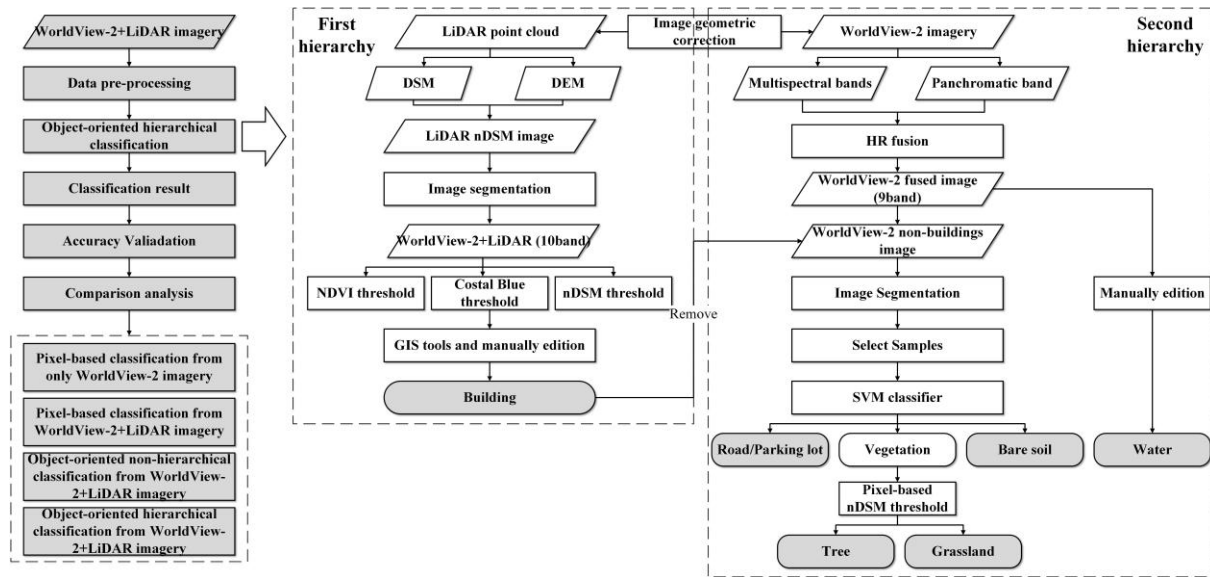
Figure 2. The data processing of urban land cover classification in this study.

2.1 Data pre-processing

During the data pre-processing, WorldView-2 MS and PAN ortho-rectified images were fused using a Haze-and-ratio-based (HR) fused scheme. The HR fused method, proposed by Jing [15], belonged to a PAN modulation (PM) based fusion method and took into account haze. HR fused scheme was presented as Equation (1).

$$\frac{M_{i,f} - h_i}{M_i - h_i} = \frac{PAN - h_p}{P_L - h_p} \tag{1}$$

Where $M_{i,f}$ $(i = 1..n)$ denoted the fused i th MS band, M_i was the corresponding i th MS band, *PAN* represented the high-resolution PAN image. h_p and h_i were related to atmosphere effects in PAN band and in the i th MS band, respectively. P_L was a spatially degraded PAN image by averaging the high-resolution PAN pixels. A fused eight WorldView-2 MS image, with 0.5 meter resolution, was generated through HR fusion.

WorldView-2 MS fused bands and PAN band were projected to a Universal Transverse Mercator (UTM) coordinate system with the corresponding rational polynomial coefficients (RPC) sensor camera model, and were registered to the LiDAR nDSM image (UTM projection), with less than 1 pixel of the Root Mean Square (RMS).

2.2 Hierarchical object-oriented classification

In this study, according to visual survey and the experience of predecessors [8,9], the land cover types in our study area were classified into six classes, including building, tree, road/parking lot, grassland, bare soil and water. Building and tree belong to non-ground object, and road/parking lot, grassland, bare soil and water belong to ground object. The water was masked before classification due to the little area proportion of the study area. The processing of classification included two hierarchies.

2.2.1 First hierarchy: LiDAR nDSM image segmentation and building extraction. Suitable image segmentation was a key factor of object-oriental classification and information extraction. In this study, the multi-resolution segmentation method, which was wide-applied in the image segmentation, was applied to objects classification. Different scale parameter, color/elevation criterion and shape criterion (including compactness and smoothness) were set for different land cover types. The building extraction was implemented, because of significant height and shape difference between buildings and other objects in LiDAR nDSM image. The multi-resolution segmentation was used in LiDAR nDSM image, and selected parameters were set through lots of experiments (scale: 30, elevation: 0.9, shape:

0.1, compactness: 0.5, smoothness: 0.5). Segmented objects were overlaid on 10band images (including eight fused WorldView-2 MS bands, one WorldView-2 PAN band and LiDAR nDSM data) for extracting the building's characters. Then a threshold-based approach was applied to building extraction. We recognized that the nDSM of buildings was higher than other ground objects, but the height of some low-rise buildings was similar to trees. The Normalized Difference Vegetation Index (NDVI) value of most buildings was lower than the vegetation (including trees and grasslands), but some buildings with blue roofs have a similar NDVI value to vegetation. However, the Costal Blue value of buildings with blue roofs was higher than vegetation significantly. Therefore, the intersection of nDSM, NDVI and Costal Blue thresholds were acquired and a little manually edition was carried out to generate the result of building extraction.

2.2.2 Second hierarchy: WorldView-2 image segmentation and objects classification. In the second hierarchy, building objects were masked. Although the height of ground objects (including road/parking lot, grassland, bare soil and water) were similar, it was considerable to discriminate those ground objects based on spectral and spatial properties. Therefore non-building image was classified based on the WorldView-2 image segmentation. The multi-resolution segmentation was used to WorldView-2 9band fused image (including eight fused MS bands and one PAN band) and selected parameters were defined (scale: 50, color: 0.9, shape: 0.1, compactness: 0.6, smoothness: 0.4). The Support Vector Machine (SVM) classifier was carried out in the segmented WorldView-2 9band image, with the kernel of Radial Basis Function (RBF). Based on Sun's conclusions, it was suitable that the Gamma parameter G was set to 0.1, and the penalty parameter C was set to 100 [16]. Selected samples include roads/parking lots, bare soils and vegetation. Trees and grasslands were merged into vegetation due to the complex shape and unsuitable scale of the segmentation. Selected features of objects include mean value of eight MS bands and one PAN band, as well as the length/width (the ratio of length and width of objects) of segmented objects, because the length/width of roads was larger than other objects. After SVM classification, an nDSM threshold was set to separate trees and grasslands from the vegetation based on pixel scale. Finally, the extracted buildings in the first hierarchy and the classified objects in the second hierarchy were merged and recoded to generate final classification result. The data processing was carried out on the platform of eCognition software (version 8.7).

2.3 Accuracy assessment and classification results analysis
As for accuracy validation, 510 random points were selected as the reference data in the WorldView-2 fused image. The number of created random points was 72 of buildings, 117 of trees, 160 of roads/parking lots, 141 of grasslands, and 20 of bare soils. The confusion matrix was constructed and the producer's accuracy (PA), the user's accuracy (UA), the overall accuracy (OA) and the overall kappa coefficient (Kappa) were calculated to analyse the classification accuracy. A visual comparison and accuracy analysis were carried out to four classification results: pixel-based classification from only WorldView-2 imagery (9band), pixel-based classification from WorldView-2 imagery and LiDAR nDSM data (10band), object-oriented non-hierarchical classification from 10band image, and proposed object-oriented hierarchical classification from 10band image. SVM classification was applied to all classification, and the scale parameters of object-oriented non-hierarchical classification was the same as segmented WorldView-2 imagery in the second hierarchy.

3. Results and analysis
According to the former methods, in the first hierarchy, the NDVI threshold was set as 0.2, Costal Blue threshold was set as 500, and nDSM threshold was set as 2.4 meter based on a lots of manual experiments. The intersection of NDVI greater than 0.2 and Costal Blue less than 500 were removed. The part of nDSM greater than 2.4 meter were extracted from removed intersection, and a little manual edition was carried out to generate building objects. In the second hierarchy, nDSM threshold was set as 2.4 meter to separate trees (≥2.4 meter) and grasslands (<2.4 meter) from classified vegetation. Table 1 listed the accuracy validation of classification results.

Table 1. The accuracy validation of urban land cover classification in this study

Classification Method	Accuracy	Non-ground		Ground		
		Building	Tree	Road/Parking lot	Grassland	Bare soil
Pixel-based classification from only WV-2 imagery	PA(%)	44.44	70.94	68.75	66.67	70.00
	UA(%)	56.14	65.35	73.83	60.65	63.64
	OA(%)	65.29				
	OK	0.54				
Pixel-based classification from WV-2 and LiDAR image	PA(%)	90.28	88.89	76.88	88.65	65.00
	UA(%)	80.25	83.87	92.48	80.65	76.47
	OA(%)	84.31				
	OK	0.79				
Object-oriented non-hierarchical classification from WV-2 and LiDAR image	PA(%)	84.72	68.38	89.38	68.09	50.00
	UA(%)	91.04	67.23	83.63	67.13	100.00
	OA(%)	76.47				
	OK	0.68				
Proposed object-oriented hierarchical classification from WV-2 and LiDAR image	PA(%)	94.44	94.02	88.75	95.74	90.00
	UA(%)	95.77	95.65	94.04	91.22	72.00
	OA(%)	92.75				
	OK	0.90				

PA: the producer's accuracy; UA: the user's accuracy; OA: the overall accuracy; OK: the overall kappa coefficient; WV-2: WorldView-2

Table 1 showed that the lowest overall accuracy was yield in the Pixel-based classification from WorldView-2 imagery alone (OA: 65.29%, OK: 0.54). As for the classification from WorldView-2 imagery alone, because the confusion between building and road/parking lot was serious, PA and UA of the building were only 44.44% and 56.14%; PA and UA of road/parking lot were only 68.75% and 73.83%, respectively. In addition, the mixture between tree and grassland was also obvious due to the similar reflectance of near-infrared band.

Adding LiDAR nDSM data, the PA and UA of the building (PA: 90.28%, UA: 80.25%), road/parking lot (PA: 76.88%, UA: 92.48%), tree (PA: 88.89%, UA: 83.87%) and grassland (PA: 88.65%, UA: 80.65%) improved significantly in the pixel-based classification, because of height difference from building and road/parking, tree and grassland. However, the accuracy of bare soil increased inconspicuous and the PA of the road/parking lot was lower than the building. The phenomenon indicated that the problem of confusion between bare soil and road/parking lot was not solved efficiently.

In terms of object-based non-hierarchical classification from WorldView-2 and LiDAR image, the overall accuracy (OA: 76.47%, OK: 0.68) was lower than pixel-based classification from two data. Non-hierarchical classification based on object scale reduced the noise of classification results, however, a scale parameter of the segmentation were unsuitable to all classes. Much tiny area including different classes were missed, which lead to the decreasing of classification accuracy, especially the tree (PA: 68.38%, UA: 67.23%) and grassland (PA: 68.09%, UA: 67.13%).

The proposed object-oriented hierarchical classification from WorldView-2 and LiDAR data was the best accuracy, compared to other three approaches (OA: 92.75%, OK: 0.90). The PA and UA of the tree (PA: 94.02%, UA: 95.65%) and grassland (PA: 95.74%, UA: 91.22%) improved significantly because of setting nDSM threshold. Hierarchical segmentation for different land cover types was helpful for improving detailed land cover classification.
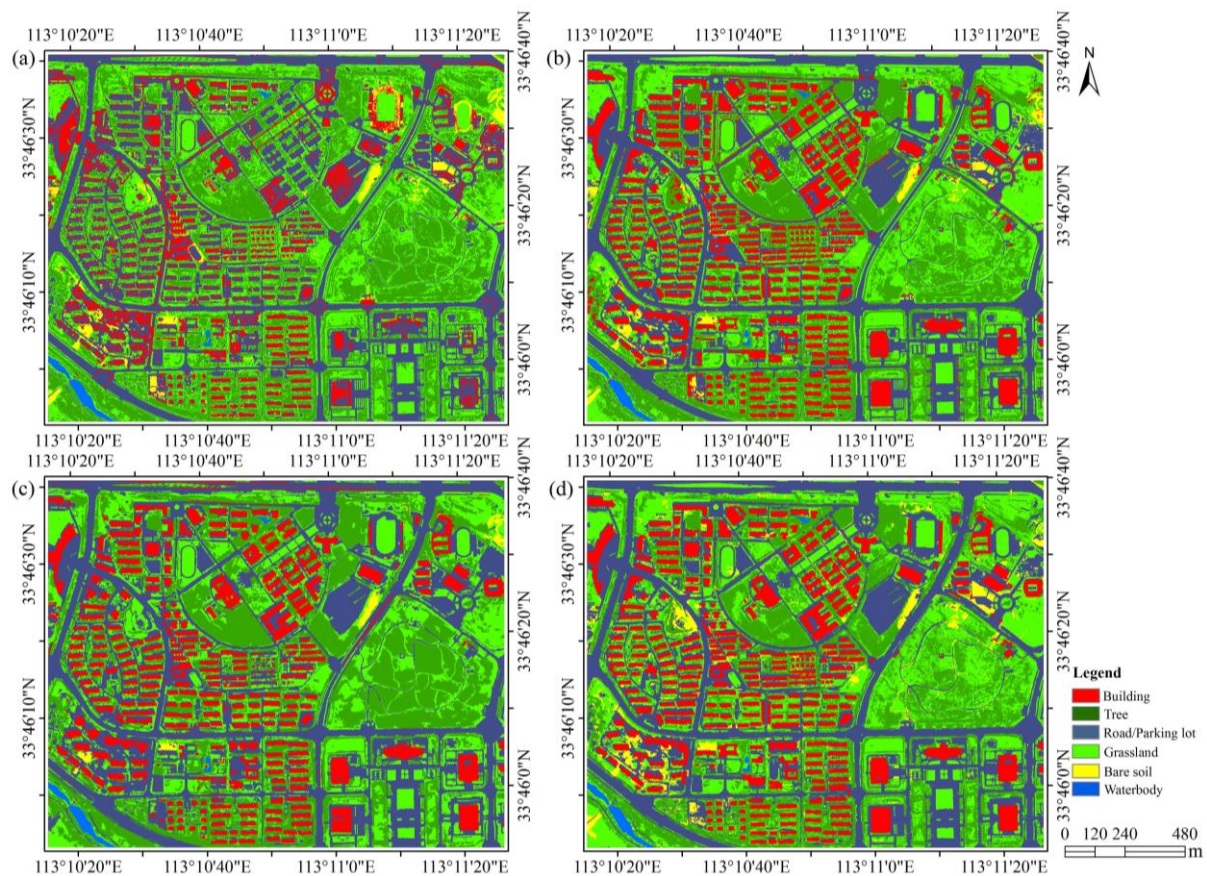
Figure 3. The classification results of urban land cover in this study.

(a): pixel-based from only WV-2 imagery; (b): pixel-based from WV-2 and LiDAR image; (c): object-oriented non-hierarchical from WV-2 and LiDAR image; (d): proposed object-oriented hierarchical from WV-2 and LiDAR image

Figure 3 showed the four classification results of our study area. Non-LiDAR classification cannot separate ground and non-ground objects with similar spectral characters effectively, such as building and road/parking lot, tree and grassland (Figure 3(a)). After adding LiDAR height information, the misclassification between building and road/parking lot was decreased obviously, and buildings were extracted completely (Figure 3(b)). Object-oriented classification reduced much "salt and pepper" noise but some detailed information was ignored. The non-hierarchical classification caused the extracted buildings incomplete and irregular, then many trees and grassland were merged and indistinguishable (Figure 3(c)). Proposed object-based hierarchical classification had the best visual presentation (Figure 3(d)).

Furthermore, the detailed comparison of four classification results was performed in Figure 4. As for extracting buildings (Figure 4a), the confusion between building and road was the most serious in the pixel-based classification from WorldView-2 image alone (Figure 4 a(2)). After adding LiDAR nDSM data, the building's extraction had a pleasant performance, however the some noise made extracted buildings incomplete (Figure 4 a(3)). Object-based non-hierarchical classification from two data decreased the noise, many irregular buildings were occurred in classification result. Because different scale parameters of image segmentation was suitable for different land cover types, single scale of segmentation cannot fit closely the boundary of objects (Figure 4 a(4)). Proposed method solved following problems effectively and extracted more regular and complete buildings (Figure 4 a(5)). Similarly, for roads extraction (Figure 4 b(1)-b(5)), Proposed method reduced the uncertainties and errors, including the misclassification from spectral similarity to road and building, speckle noises, and unsuitable scale parameters of segmentation. In terms of classified vegetation (Figure 4c), due to non-height information, classified trees and grasslands had an inferior performance in the pixel-based classification from WorldView-2 image alone (Figure 4 c(2)). Because of the overlarge scale of

segmentation, many detailed presentation of trees and grasslands was incorporated (Figure 4 c(4)). Trees and grasslands were identified accurately using nDSM feature based on pixel scale (Figure 4 c(3), c(5)). Noticeably, for shaded objects, proposed method detected accurately the vegetation in the shadows. Other three methods had the mixture between shaded vegetation and low albedo roads (Figure 5 d(1)-d(5)). In summary, proposed method in the study was valuable for improving the accuracy of urban land cover classification.
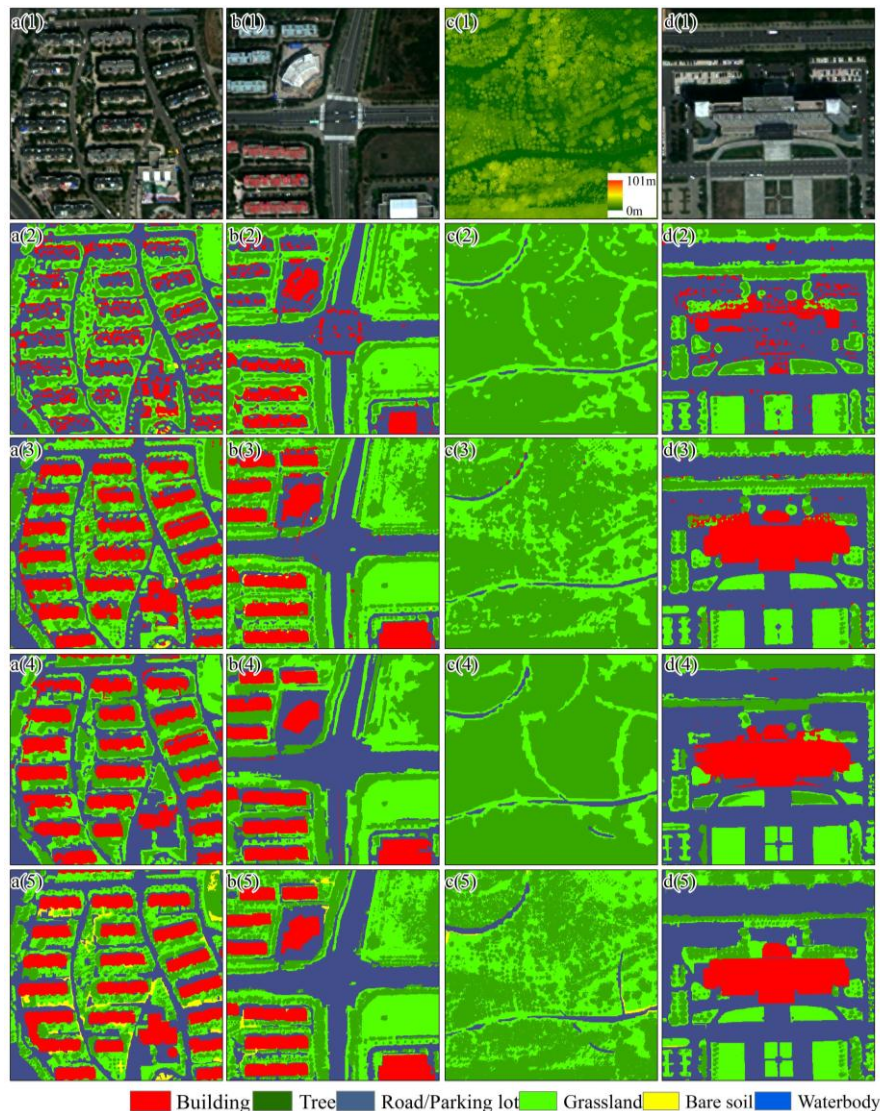

Figure 4. The detailed performance of the four classification methods.
WorldView-2 image: a(1), b(1), d(1); LiDAR nDSM data: c(1); Pixel-based classification from WorldView-2 imagery alone: a(2), b(2), c(2), d(2); Pixel-based classification from WorldView-2 and LiDAR image: a(3), b(3), c(3), d(3); Object-oriented non-hierarchical classification from WorldView-2 and LiDAR image: a(4), b(4), c(4), d(4); Proposed object-oriented hierarchical classification from WV-2 and LiDAR image: a(5), b(5), c(5), d(5)

4. Conclusions

In this study, based on the combined use of WorldView-2 imagery and airborne LiDAR data, an object-oriented hierarchical method was proposed and applied to urban land cover classification. Primary finds were obtained: (1) Proposed method yield the best accuracy and visual performance compared to other three approaches (OA: 92.75%, OK: 0.90). (2) Buildings extraction was complete and regular based on LiDAR nDSM image segmentation due to height and shape difference from buildings and other objects. Adding LiDAR height information reduced the confusion between buildings and roads/parking lots. (3) Compared with pixel-based method, object-oriented classification

decreased the noise but ignored some detailed information. Unsuitable scale parameters of the segmentation decreased the accuracy of classified trees and grasslands. Correct and detailed discrimination between trees and grasslands was presented based on nDSM threshold in the vegetation. (4) Proposed method reduced the misclassification between shaded vegetation and low-albedo roads, which was helpful for identifying the objects in the shadows in the high resolution image. In conclusion, proposed method explored the potential of LiDAR data and hierarchical object-based classification in the urban land cover classification. However, proposed method increased the accuracy of bare soils non-clearly. How to improve the performance of classified bare soils will be further studied.

Acknowledgments

References

[1]   Dengsheng L, Scott H and Emilio M 2011 Impervious surface mapping with Quickbird imagery. *Int. J. Remote Sens.* **32** 2519-2533

[2]   Huadong G, Huaining Y, Zhongchang S, Xinwu L and Cuizhen W 2014 Synergistic use of optical and PolSAR Imagery for urban impervious surface estimation *Photogramm. Eng. Remote Sens.* **80** 91-102

[3]   Mohapatra R P and Changshan W 2010 High resolution impervious surface estimation: an integration of IKONOS and Landsat-7 ETM+ imagery *Photogramm. Eng. Remote Sens.* **76** 1329-1341

[4]   Tammy E P and James B C 2013 Comparing urban impervious surface identification using Landsat and high resolution aerial photography. *Remote Sens.* **5** 4942-4960

[5]   Gaia Vaglio L, Veraldo L, Qi C, Leila G, Fabio Del F, Antonio B, David C, Beccy W, Jeremy L and Riccardo V 2013 Optical and SAR sensor synergies for forest and land cover mapping in a tropical site in West Africa. *Int. J. Appl. Earth Obs.* **21** 7-16

[6]   Yuanzhi Z, Hongsheng Z and Hui L 2014 Improving the impervious surface estimation with combined use of optical and SAR remote sensing images. *Remote Sens. Environ.* **141** 155-167

[7]   Yunhao C, Wei S, Jing L and Zhongping S 2009 Hierarchical object oriented classification using very high resolution imagery and LiDAR data over urban areas. *Adv. Space Res.* **43** 1101-1110

[8]   Jungho I, Zhenyu L, Jinyoung R and Lindi J Q 2012 Impervious surface quantification using a synthesis of artificial immune networks and decision/regression trees from multi-sensor data. *Remote Sens. Environ.* **117** 102-113

[9]   Shezhou L, Cheng W, Xiaohuan X, Hongcheng Z, Dong L, Shaobo X and Pinghua W 2016 Fusion of airborne discrete-return LiDAR and hyperspectral data for land cover classification. *Remote Sens.* **8** 1-19

[10]   Wai Yeung Y, Ahmed S and Nagwa E 2015 Urban land cover classification using airborne LiDAR data: A review. *Remote Sens. Environ.* **158** 295-310

[11]   Lei C, Shuhe Z, Wenquan H and Yun L 2015 Building detection in an urban area using LiDAR data and QuickBird imagery *Int. J. Remote Sens.* **33** 5135-5148

[12]   Xiaoxiao L, Soe W M, Yujia Z Chritopher G, Xiaoxiang Z and Billie L Turner I 2014 Object-based land-cover classification for metropolitan Phoenix, Arizona, using aerial photography *Int. J. Appl. Earth Obs.* **33** 321-330

[13]   Yuanrong H, Xinxin Z and Lizhong H 2016 Object-based distinction between building shadow and water in high-resolution imagery using fuzzy-rule classification and artificial Bee Colony optimization. *J. Sens.* **2016** 1-10

[14]   Soe W M, Patricia G, Anthony B, Susanne G and Qihao W 2011 Per-pixel vs.object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens. Environ.* **115** 1145-1161

[15]   Linhai J and Qiuming C 2009 Two improvement schemes of PAN modulation fusion methods for spectral distortion minimization. *Int. J. Remote Sens.* **30** 2119-2131

[16]   Zhongchang S, Huadong G, Xinwu L and Linlin L 2011 Estimating urban impervious surfaces from Landsat-5 TM imagery using multilayer perceptron neural network and support vector machine. *J. Appl. Remote Sens.* **5** 1-17