

Synergistic Use of WorldView-2 Imagery and Airborne LiDAR Data for Urban Land Cover Classification

M F Wu^{1,2}, Z C Sun^{2,3}, B Yang¹, 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. There are lots of challenges for deriving urban land cover types for high resolution optical imagery because of spectral similarity of different objects, mixed pixels, shadows of buildings and large tree crowns. In order to reduce these uncertainties, recently, it's a trend of the classification of urban land cover from multi-source sensors in the field of urban remote sensing. In this study, a hierarchical support vector machine (SVM) classification method was applied to the urban land cover mapping, using the WorldView-2 imagery and airborne Light Detection and Ranging (LiDAR) data. The results showed that: (1) The overall accuracy (OA) and overall kappa (OK) were 72.92% and 0.66 for WorldView-2 imagery alone; while the OA and OK were improved up to 89.44% and 0.87 for the synergistic use of the two types of data source. (2) Buildings and road/parking lots extracted from fused data were more precision and well-shaped. The two classes from fused data were optimally classified with higher producer's accuracy and user's accuracy than WorldView-2 imagery alone. The trees were also easily separated from the grasslands when the airborne LiDAR data was added. (3) The fused data could reduce the phenomenon of different spectral character of the complex and detailed objects. It was also helpful to address the problem of shadows from the high-rise buildings. The results from this study indicate that the synergistic use of high resolution optical imagery and airborne LiDAR data can be an efficient approach to improving the classification of urban land cover.

1. Introduction

Sustainable and cost effective urban sprawl dynamic monitoring is desirable, but the development and implementation of accurate urban land cover classification remain a significant challenge. It is often difficult to accurately classify urban land cover from Very High Resolution (VHR) optical images alone. Some limitations include: (a) some materials of urban objects have similar spectral properties. For example, some buildings with the grey roofs share similar reflectance characteristics with the asphalt roads; and the spectral properties of high albedo objects such as bare soils, cloud, sand and the buildings with white roofs are similar [1]. (b) Because of the heterogeneity in urban landscapes, the mixed pixel has been recognized as a challenge to improve classification accuracy [2]. (c) Shadows casted by buildings or large tree crowns have become a constraint for the classification [3]. Therefore, recent advances in Earth Observation allow using multi-source high resolution data to conduct the more detailed urban land cover classification.

The fusion of Light Detection and Ranging (LiDAR) data and VHR optical imagery has been used for urban land cover mapping in the emergence. Chen et al. combined Quick Bird image with LiDAR



In this study, synergistic use of the WorldView-2 image and airborne LiDAR data, a hierarchical support vector machine (SVM) classification method was used to map the urban land cover types. A compared analysis between WorldView-2 imagery alone and the fusion of this two data source was conducted. The objective of this study is to expand the application of fusion of VHR optical imagery and LiDAR data in the field of urban land cover.

The study area is located in the Pingdingshan City, Henan Province, China. Collected WorldView-2 imagery was acquired on 21 August 2014. The 11-bit WorldView-2 imagery consists of one panchromatic (PAN) band with 0.5×0.5 meter pixel size and eight multispectral bands across wavelengths ranging from 400 to 1040 nm, with the spatial resolution of 2 meter. Collected airborne LiDAR data was acquired on 19 August 2013 and were provided in ASCII format, including the X, Y, and Z coordinates and delivered first return intensities. The two dimensional mean point density was 23.0 points/m². The DSM and the DEM were derived from airborne LiDAR point clouds. The nDSM image was generated by subtracting the DSM with DEM, with 0.5 meter spatial resolution (Yan, 2015).



Figure 1 shows the data procession for classifying the urban land cover from WorldView-2 imagery and airborne LiDAR data, Including data pre-processing, a hierarchical SVM classifier construction, accuracy assessment and classification result analysis.

2.1 Data pre-processing

During the data pre-processing, the WorldView-2 imagery was rectified to a Universal Transverse Mercator (UTM) coordinate system with the corresponding rational polynomial coefficients (RPC) sensor camera model. The LiDAR nDSM image, with 0.5 meter resolution, was registered to the WorldView-2 imagery (UTM), and the Root Mean Square (RMS) value was less than 1 pixel

A Haze-and-ratio-based (HR) fused scheme was used by WorldView-2 MS and PAN bands. The HR fused method is a PAN modulation (PM) based fusion method that takes into account haze [10]. Equation (1) shows that $M_{i,f}$ ($i = 1..n$) denotes the fused i th MS band, M_i is the corresponding i th MS band, PAN denotes the high-resolution PAN image. h_p and h_i are related to atmosphere effects in PAN band and in the i th MS band, respectively. P_L is a spatially degraded PAN image by averaging the high-resolution PAN pixels. A 10 bands image, including a WorldView-2 MS image with 0.5 meter spatial resolution by HR fusion method, WorldView-2 PAN image and LiDAR nDSM image, were generated to conduct following classification.

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

2.2 Hierarchical support vector machine (SVM) classifier

A pixel-based SVM method was used in the classification. SVM is supervised, non-parametric statistical learning techniques designed to solve classification and regression problems [11]. SVM has shown great capabilities to model complex multi-modal, nonlinear data distributions in high-dimensional features spaces [12]. In our study, a SVM method with the kernel of Radial Basis Function (RBF) was used to perform the classification. The kernel parameter G was set to 0.1, the reciprocal of band number (10), and the penalty parameter was set to 100 [13].

Based on the traditional SVM model, a hierarchical classification was added. An nDSM threshold determined by the Decision Tree for splitting ground and non-ground objects was 2.4 meter. After mask, the SVM classifier was used to ground and non-ground images respectively. The land cover types in our study area were classified into six classes, including buildings, trees in non-ground objects, and road/parking lots, grasslands, bare soils, and waterbody in ground objects. The waterbody was manually clipped before classification because of the proportion of area of waterbody was little. The pure samples were manually selected by cross-validating geo-referenced field sites and aerial photos. Selected samples polygons were randomly split to training sample datasets and test sample datasets according to 2:1 ratio. The training samples were used to the classification processing of SVM, and the test samples were used to build the confusion matrix and validate SVM modal accuracy. A “Majority/Minority” filtering for each class, with 3×3 window, was employed to reduce the “salt and pepper” phenomenon in pixel-based classification.

2.3 Accuracy validation

As for accuracy validation, the random points of every class (the number of random points was Buildings: 200, Trees: 150, Roads/Parking lots: 200, Grasslands: 100, Bare soils: 100) were selected based on aerial photos and field truth data and the producer’s accuracy (PA), the user’s accuracy (UA), the overall accuracy (OA) and the overall kappa coefficient (Kappa) were calculated. The confusion matrix was built to analyse the classification accuracy.

3. Results and analysis

Table 1. The classification accuracy of the study area land cover/land use

WorldView-2	Buildings	Trees	Roads/Parking lots	Grasslands	Bare soils	Total	UA (%)
Buildings	114	11	60	3	10	198	57.58
Trees	0	128	0	22	0	150	85.33
Road/Parking lots	47	5	133	10	4	199	66.83
Grasslands	2	10	2	84	1	99	84.85
Bare soils	4	0	11	0	85	100	85.00
Total	167	154	206	119	100	746	
PA (%)	68.26	83.12	64.56	70.59	85.00		
OA (%)	72.92						
OK	0.66						

WorldView-2+LiDAR nDSM	Buildings	Trees	Roads/Parking lots	Grasslands	Bare soils	Total	UA (%)
Buildings	169	6	16	4	4	199	84.92
Trees	0	141	2	5	1	149	94.63
Road/Parking lots	4	9	181	4	2	200	90.50
Grasslands	1	6	3	90	0	100	90.00
Bare soils	1	1	8	2	88	100	88.00
Total	175	163	210	105	95	748	
PA (%)	96.57	86.50	86.19	85.71	92.63		
OA (%)	89.44						
OK	0.87						

PA: the producer's accuracy; UA: the user's accuracy; OA: the overall accuracy; OK: the overall kappa coefficient

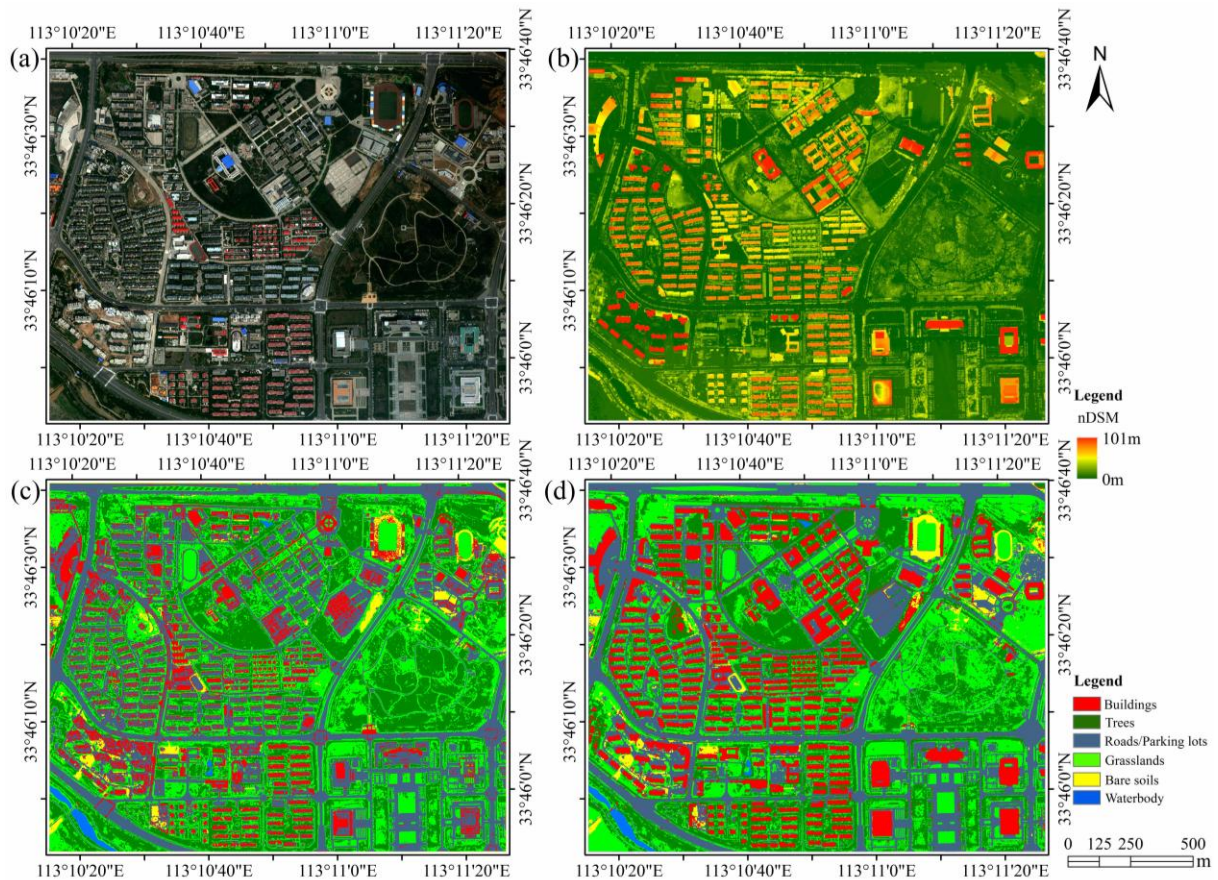


Figure 2. The land cover classification results in the development area.

(a) WorldView-2 imagery; (b) LiDAR nDSM image; (c) classification result of WorldView-2 image alone; (d) classification result of the synergistic use of WorldView-2 image and LiDAR nDSM data

The hierarchical classification results, of our study area, from fusion of WorldView-2 imagery and LiDAR nDSM data were generated to compare with WorldView-2 imagery alone using a non-hierarchical classification (Figure 2). Then the results of accuracy evaluation are listed in Table 1.

According to the Figure 2 and Table 1, we generated the finer classification result from fusion of WorldView-2 imagery and LiDAR nDSM data, compared with WorldView-2 imagery alone. Buildings and roads/parking lots, extracted from fused data, were more precision and well-shaped. Only 68.26% of buildings could be classified correctly from WorldView-2 imagery alone. 96.57% of buildings could be extracted accurately after integrating LiDAR height information. Similarly, the producer's accuracy and user's accuracy of roads/parking lots improved from 64.56% to 86.19%, and from 66.83% to 90.50%. What's more, adding height information, the PA of grasslands was from 70.59% up to 85.71%, and the UA of trees was from 85.53% up to 96.63%. The fused data reduce the misclassification between grasslands and trees. In summary, for the synergistic use of WorldView-2 imagery and airborne LiDAR data, the overall accuracy improved from 72.92% to 89.44%, overall Kappa coefficient improved from 0.66 to 0.87.

A visual comparison is showed in Figure 3. The height information improved the buildings extraction. For WorldView-2 imagery alone, much grey roofs misclassified as roads and parking lots (Figure 3(a)). Then as for the spectral difference in an individual roof, the buildings extracted were well-shaped from fused data (Figure 3(b)). Furthermore, due to the non-orthographic phenomenon in WorldView-2 high spatial resolution data, the roofs and floors of tall-buildings were classified as buildings and roads respectively, and the shadows on the roads, from high-rise buildings, were misclassified as buildings. After adding nDSM data, the high-rise buildings were extracted more precision and the misclassification between shadow and buildings was reduced (Figure 3(c)). In addition, hierarchical fused data classification reduced the mixture of some grasslands and trees. Therefore, height information can contribution to classification (Figure 3(d)).

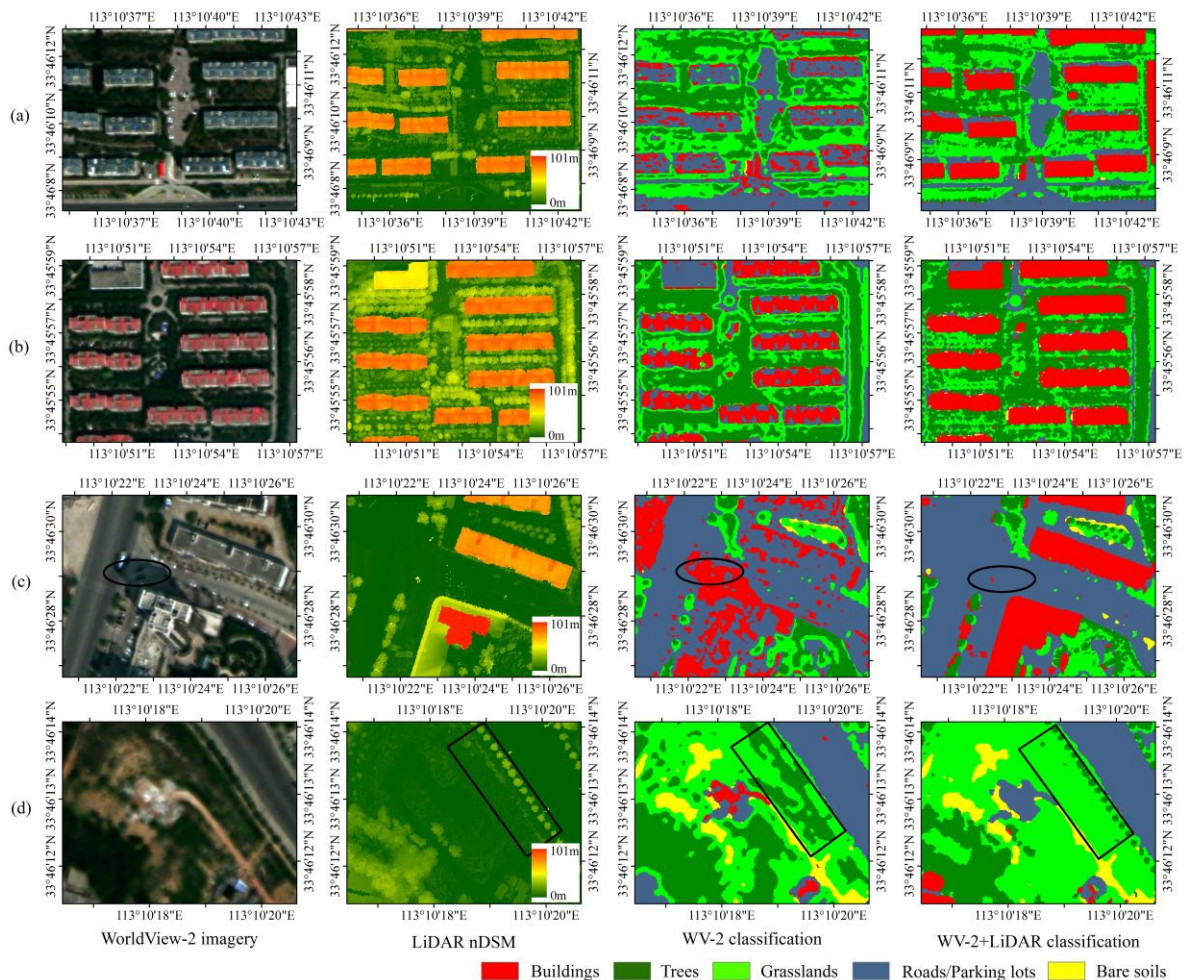


Figure 3. The detail of land cover classification results in study area.

(a) shows WV-2+LiDAR classification reduced the mixture between grey roofs and asphalt roads; (b) presents WV-2+LiDAR classification was valuable for solving the problem of spectral difference in individual roof; (c) illustrates WV-2+LiDAR classification improved the extraction of high-rise buildings and shaded roads; (d) explains WV-2+LiDAR classification classified trees and grasslands more accurately.

4. Conclusions

In the study, based on the fusion of WorldView-2 imagery and airborne LiDAR data, a hierarchical SVM classification method was applied to urban land cover classification, and compared with the non-hierarchical classification using WorldView-2 imagery alone. Three results were obtained:

- (1) Compared with WorldView-2 image alone, the overall accuracy and overall Kappa coefficient of classification from fused data improved from 72.92% to 89.44%, and from 0.66 to 0.87 respectively. The major reason was improving the classification accuracy of buildings and roads/parking lots.
- (2) Due to adding the height information, fused data reduced the confusion of white roofs with high albedo roads, and grey roofs with low albedo roads. Fused data also reduced the imprecise buildings extraction from spectral difference in a roof due to complex detail of urban objects.
- (3) Moreover, fused data could extract the high-rise buildings more precisely due to the LiDAR nDSM orthographic image, and fused data was helpful for addressing the shadows from buildings to some extent.

Acknowledgments

The paper is supported by the National Natural Science Foundation of China (No.41201357). The authors wish to acknowledge the experts providing the WorldView-2 imagery and airborne LiDAR data used for the study, and the anonymous reviewers for providing helpful suggestions that greatly improved the manuscript.

References

- [1] Hangqiu X and Meiya W 2016 Remote sensing-based retrieval of ground impervious surfaces *J. Remote Sens.* **20** 1270-1289
- [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] Dengsheng L, Scott H and Emilio M 2011 Impervious surface mapping with Quickbird imagery. *Int. J. Remote Sens.* **32** 2519-2533
- [4] 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
- [5] Hartfield K A, Landau K I, and Leeuwen W J D V 2011 Fusion of high resolution aerial multispectral and lidar data: land cover in the context of urban mosquito habitat. *Remote Sens.* **3** 2364-2383
- [6] Im J, Lu Z, Rhee J, and Quackenbush L J 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
- [7] Zhou Y and Qiu F 2015 Fusion of high spatial resolution worldview-2 imagery and lidar pseudo-waveform for object-based image analysis *ISPRS J. Photogramm.* **101** 221-232
- [8] Längkvist M, Kiselev A, Alirezaie M and Loutfi A 2016 Classification and segmentation of satellite orthoimagery using convolutional neural networks *Remote Sens.* **8** 1-21
- [9] Yushi C, Chunyang L, Pedram G, Chunyu S and Yanjeng G 2016 Deep fusion of hyperspectral and LiDAR data for thematic classification *IGARSS* 3591-3594
- [10] 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
- [11] Schölkopf B, Smola A 2002 Learning with kernels-Support vector regularization, optimization, and beyond, MIT Press
- [12] Huang C, Davis L S and Townshend J R G 2002 An assessment of support vector machines for land cover classification *Int. J. Remote Sens.* **23** 725-749
- [13] Zhongchang S, Huadong G, Xinwu L, Linlin L and Xiaoping D 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