

## OBJECT BASED IMAGE ANALYSIS COMBINING HIGH SPATIAL RESOLUTION IMAGERY AND LASER POINT CLOUDS FOR URBAN LAND COVER

Xiaoliang ZOU<sup>a, b, c</sup>, Guihua ZHAO<sup>c</sup>, Jonathan LI<sup>b</sup>, Yuanxi YANG<sup>a, c</sup>, Yong FANG<sup>a, c, \*</sup>

<sup>a</sup> State Key Laboratory of Geo-Information Engineering, Xi'an, China 710054 – (zhz200271, yuanxi\_yang)@163.com

<sup>b</sup> Department of Geography and Environmental Management, Faculty of Environment, University of Waterloo, Waterloo, ON, Canada N2L 3G1 – (x26zou, junli)@uwaterloo.ca

<sup>c</sup> Xi'an Institute of Surveying and Mapping, Xi'an, China 710054 – rockdavidzou@gmail.com

### Commission III, WG III/4

**KEY WORDS:** Object Based Image Analysis (OBIA), Image Segmentation, Classification, Accuracy Assessment, nDSM, Parcel Map, TrueOrtho Photo, Lidar Point Clouds

### ABSTRACT:

With the rapid developments of the sensor technology, high spatial resolution imagery and airborne Lidar point clouds can be captured nowadays, which make classification, extraction, evaluation and analysis of a broad range of object features available. High resolution imagery, Lidar dataset and parcel map can be widely used for classification as information carriers. Therefore, refinement of objects classification is made possible for the urban land cover. The paper presents an approach to object based image analysis (OBIA) combining high spatial resolution imagery and airborne Lidar point clouds. The advanced workflow for urban land cover is designed with four components. Firstly, colour-infrared TrueOrtho photo and laser point clouds were pre-processed to derive the parcel map of water bodies and nDSM respectively. Secondly, image objects are created via multi-resolution image segmentation integrating scale parameter, the colour and shape properties with compactness criterion. Image can be subdivided into separate object regions. Thirdly, image objects classification is performed on the basis of segmentation and a rule set of knowledge decision tree. These objects imagery are classified into six classes such as water bodies, low vegetation/grass, tree, low building, high building and road. Finally, in order to assess the validity of the classification results for six classes, accuracy assessment is performed through comparing randomly distributed reference points of TrueOrtho imagery with the classification results, forming the confusion matrix and calculating overall accuracy and Kappa coefficient. The study area focuses on test site Vaihingen/Enz and a patch of test datasets comes from the benchmark of ISPRS WG III/4 test project. The classification results show higher overall accuracy for most types of urban land cover. Overall accuracy is 89.5% and Kappa coefficient equals to 0.865. The OBIA approach provides an effective and convenient way to combine high resolution imagery and Lidar ancillary data for classification of urban land cover.

### 1. INTRODUCTION

With the rapid developments of the sensor technology, high spatial resolution imagery and Light Detection And Ranging (Lidar) cloud points can be captured nowadays, which makes classification, extraction, evaluation and analysis of a broad range of object features available. High spatial resolution imagery contains rich spectral, spatial and texture information. The laser scanner can rapidly acquire high-precision 3D information which is a more suitable and auxiliary method for urban classification. The combination of high resolution imagery and Lidar point clouds based on Object Based Image Analysis (OBIA) method can be produced results at greater classified accuracy than the contributions of either field alone. Many researchers have been attempting to combine multi-data to identify or extract object feature, to detect elevated objects such as building and tree, to detect changes of vegetation or to estimate for vegetation parameters, and to classify or segment properties in land cover and land use. Classification accuracy can be improved by accessing to accurate ancillary information and other types of modelling such as DEM, DSM or Lidar point clouds. The earliest method in land cover mapping using remotely sensed data incorporating topographic information from digital terrain models and other GIS data was presented (Franklin, S.E. et al., 1992). The concepts of pixel-based image processing were developed in the 1970s and pixel-based image

classification was applied in land cover and land use. The main drawback of per-pixel classification in multi-dimensional feature space is that it doesn't make use of any spatial concept (Blaschke and Strobel, 2001). Object-oriented image processing overcomes these difficulties by segmenting the image into meaningful objects based on both spectral and spatial characteristics. Automatic classification of land cover feature with high resolution imagery and Lidar based on object-oriented approach was used to detect individual building and tree crown (Sohel Syed et al., 2005). An object-oriented approach for analysing and characterising the urban landscape structure at the parcel level was presented using high-resolution digital aerial imagery and Lidar data for the Baltimore area (Zhou and Troy, 2008). The development overview about OBIA methods was given, which aim to delineate usable objects from imagery while at the same time combining image processing and GIS functionalities in order to utilize spectral and contextual information in an integrative way (Blaschke, 2010). High point density Lidar data and colour-infrared orthoimages were applied for the classification of shrubs and trees in a study area in Denmark (Thomas, et al. 2013). A graph based algorithm was promoted, which combines multispectral imagery and airborne Lidar information to completely delineate the building boundaries in urban and densely vegetated area (Gilani et al., 2015). High resolution imagery, Lidar datasets and parcel map data have been

\* Corresponding author

widely used in classification as information carriers. The paper presents an approach to use object based image analysis (OBIA) combining high spatial resolution imagery and Lidar cloud points in order to refine objects classification results for urban land cover, and the advanced workflow is designed.

## 2. STUDY AREA AND DATA PREPROCESSING

### 2.1 Study Area

The research focused on the photogrammetric test site Vaihingen/Enz of Institut für Photogrammetrie, Stuttgart University in Germany ([http://www.ifp.uni-stuttgart.de/forschung/photo/test\\_site/index.en.html](http://www.ifp.uni-stuttgart.de/forschung/photo/test_site/index.en.html)). The test site was used for the independent in-flight evaluation of new digital airborne sensors. The test site itself is located about 28km north-west of Stuttgart/Germany. The city of Vaihingen/Enz is situated on the river Enz. The overall test area is about 7.5 km (east-west) x 4.7 km (north-south). The terrain heights differ between 193m and 356m above mean sea level. It belongs to a hilly area providing several types of vegetation such as steep vineyards and quarries.

### 2.2 Data Sets of Vaihingen/Enz

The paper selects a subset of datasets, which comes from ISPRS WG III/4 test project on urban classification and 3D building reconstruction (Franz Rottensteiner et al., 2013). These datasets were provided by ISPRS WG III/4 working group. This test datasets were captured over Vaihingen/Enz using Intergraph/ZI DMC digital aerial cameras carried out by the German Association of Photogrammetry and Remote Sensing (DGPF) in 2008 (ISPRS Report, 2013). The aim of this project is to provide state-of-the-art datasets which can be used by interested researchers in order to test own methods and algorithms on urban object classification and building reconstruction. Figure 1 shows a patch of trueortho photo (TOP) in Vaihingen/Enz Block for the paper study.



Figure 1. A patch of trueortho photo in Vaihingen/Enz block

The patch of test area belongs to urban area among datasets for various object classes, which is characterized by many residential buildings, many detached houses and multi-story buildings surrounded by trees. The river Enz flows through this area. The variety of urban and suburban land cover types makes it ideal for testing urban object classification. Meanwhile, high resolution colour-infrared TrueOrtho Photo (TOP), original airborne Lidar point clouds, digital surface model (DSM) and other ancillary data are provided by ISPRS WG III/4 work group.

**2.2.1 Images Data:** Working group provides two state-of-the-art airborne image datasets, consisting of very high resolution colour-infrared TOP tiles and corresponding digital surface models (DSM). The DSM is generated via dense image matching with Trimble INPHO 5.3 software, and the TOP mosaic is generated with OrthoVista module in Trimble INPHO software (ISPRS Report, 2013). The TOPs are 8 bit TIFF World format with three bands: near infrared, green and red bands. The DSM is a TIFF World format with one band and the grey levels are encoded as 32 bit float values. The ground sampling distance (GSD) of both the TOP and the DSM is 9cm with the same grid.

**2.2.2 Airborne Laser Scanning Dataset:** Working group provides the original airborne laser scanning datasets from Vaihingen/Enz. These datasets are captured with a Leica ALS50 system (ISPRS Report, 2013). Inside an individual strip the average point density is 4 points/m<sup>2</sup>. The laser datasets are LAS 1.2 file format and each point with XYZ coordinate, intensity, elevation, number of returns, returns number, scan angle, class and source ID information. These laser datasets are made available as ancillary data for classification.

## 3. WORKFLOW DESIGN

The paper designs the workflow of Object Based Image Analysis (OBIA), which can be separated in four main steps. Figure 2 illustrates a designed workflow of OBIA classification with ancillary data of point clouds.

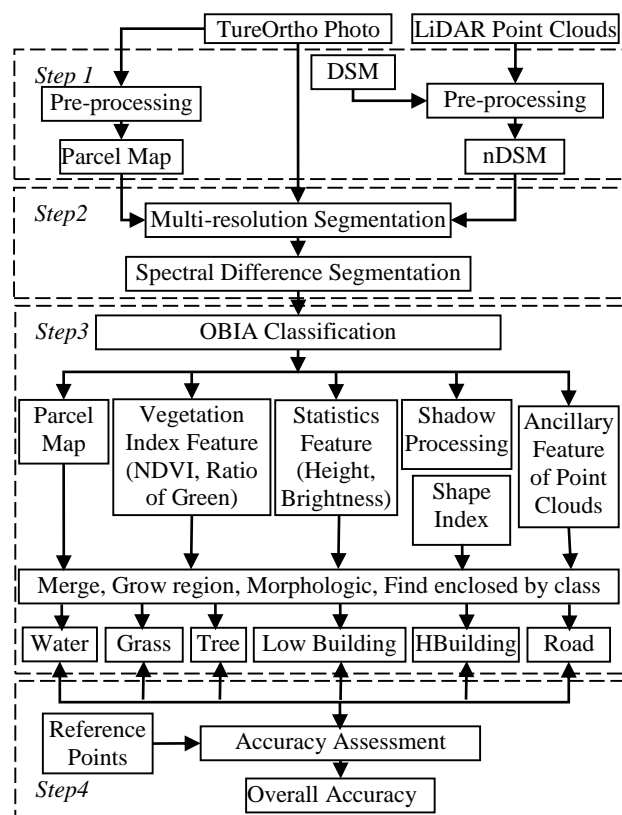


Figure 2. Workflow of OBIA classification with ancillary data

The first step is to pre-process in order to obtain parcel map of water bodies and to create a normalised digital surface model (nDSM) from the point clouds. The polygon lines along the river Enz is collected in shape file on trueortho photo using ArcGIS 10.3.1 for Desktop software (ESRI, 2015) and assigned to the thematic attributes ID of water bodies for parcel map. And then,

the digital elevation model (DEM) is generated by de-noising, filtering and interpreting processing from laser scanning point clouds. The nDSM is produced by subtracting DEM from the existed DSM, forming difference in height using image calculator tools on ArcGIS software (ESRI, 2015). Meanwhile, buildings are classified roughly into a standard class from Lidar point clouds.

The second step is multi-resolution image segmentation. Image objects primitives are created via multi-resolution image segmentation integrating scale parameter, the colour and shape properties and compactness criterion. Image can be subdivided into separate regions of objects primitives. And then, the spectral difference segmentation is used to merge neighbouring image objects with a spectral mean below the threshold.

The third step is image objects classification. The image objects classes of interest typically discerned in urban classification are assigned to six different classes: water bodies, low vegetation/grass, tree, low building, high building and road. Image objects are attributed with a high-dimensional feature space, such as object vegetation index feature (NDVI, ratio of green), object shape feature (length/width, shape index, area), object statistics feature (difference in height, brightness) and ancillary feature of point clouds (intensity, class, elevation, number of returns, returns number). So, image object can be linked to a hierarchical network. A rule set of knowledge decision tree is created for classification.

The fourth step is to assess the validity of classification results. Accuracy assessment is performed by comparing randomly distributed sampled points with the classification results.

#### 4. OBJECT BASED IMAGE ANALYSIS APPROACH

The paper uses eCognition Developer software (eCognition, 2014) to execute the image segmentation and image classification based on an OBIA approach.

##### 4.1 Image Segmentation

Image segmentation is the subdivision of colour-infrared true-ortho image into separated object regions, which meet certain criteria of homogeneity and heterogeneity. Image objects primitives are created via multi-resolution segmentation based on the homogeneity criterion, which is a combination of the spectral and shape properties in smoothness and compactness criterion. Multi-resolution segmentation algorithm is a bottom up region-merging technique starting with one-pixel objects and a local optimization procedure (Benz et al., 2004). This algorithm locally minimizes the average heterogeneity of image objects and maximizes their respective homogeneity for a given variance threshold of image objects (eCognition, 2014). The multi-resolution segmentation approach allows for nested segmentation at different scale parameters. A hierarchical network of image objects can be constructed by repeating the segmentation with different scale parameters. In the hierarchy, smaller image objects are merged into larger ones. Each image object is encoded with its hierarchical position, its membership and its neighbours. Higher values for the scale parameter result in larger image objects, smaller values in smaller ones (Zhou, et al., 2008; eCognition, 2014).

Firstly, we use the approach of multi-resolution segmentation to create image objects primitives from image pixels. We select a group of parameters, which scale parameter is equal to 30, shape parameter 0.3 and compactness parameter 0.5. The weighting of

NIR image layer has been increased to 2, and the weightings of other image layers such as red, green are increased to 1. The thematic layer of water bodies with parcel boundary is used to be candidates for segmentation. The weighting of water body layer is increased to 1. The difference of height in nDSM derived from the point clouds are used for the classification. The weighting of nDSM layer is increased to 1. The Lidar point clouds were not considered in segmentation process, and the weighting of Lidar layer is ignored and setup to 0. The multi-resolution segmentation creates image objects using an iterative algorithm, whereby image objects are grouped until a threshold representing the upper object variance is reached. The scale parameter is weighted with separation of shape and compactness parameters to minimize the fractal borders of the objects. Secondly, we use the spectral difference segmentation to merge neighbouring image objects with a spectral mean below the threshold given. The threshold of maximum spectral difference is increased to 12. The neighbouring image objects are merged to produce the final objects if the difference between their layer mean intensity is less than 12. These parameters can distinguish between image objects of building, shadow, road features and other objects primitives. Figure 3 and Figure 4 illustrate the results of multi-resolution segmentation and spectral difference segmentation.

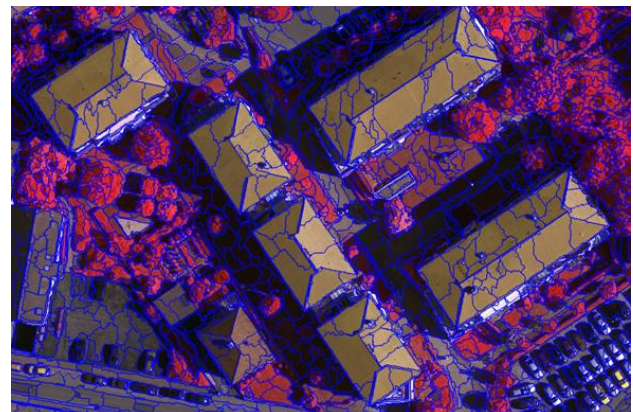


Figure 3. Results of multi-resolution segmentation using the parameters of scale =30, shape=0.3 and compactness=0.4

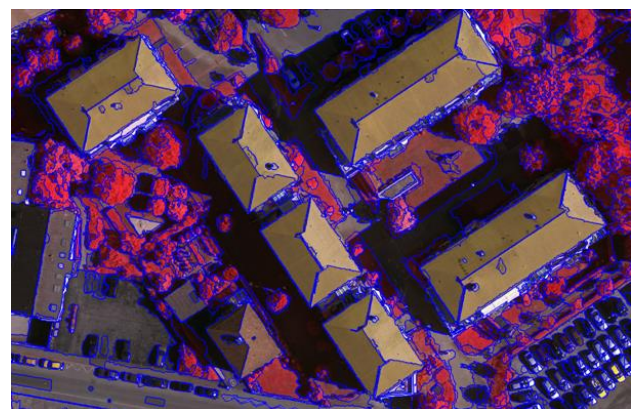


Figure 4. Results of spectral difference segmentation using the parameter of maximum spectral difference=12

##### 4.2 Imagery Classification

The aim of classification is to identify the image object as certain class. According to the currently available knowledge and terrain analysis of test area, six categories in urban classification are defined as water bodies, low vegetation/grass, tree, low building, high building and road. These object primitives can first be

assigned to class after initial segmentation. A knowledge base of classification rules is constructed to classify each object into certain class. To choose the relevant image objects, the threshold of objects are determined by quantitative analysis using image object related features information. Objects characteristics such as NDVI, ratio of green, thematic attribute, point cloud-related attribute, spatial relation and statistical measures are utilized to create the knowledge rules for classification. Ancillary data such as water bodies with parcel boundary and height information of building or tree derived from nDSM is also used to create the knowledge rules, forming process tree. Figure 5 illustrates the class hierarchy and classification rules based on OBIA approach.

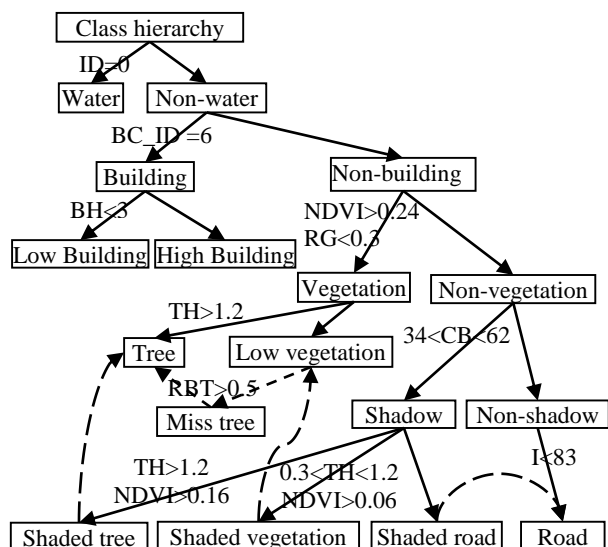


Figure 5. Class hierarchy and classification rules based on OBIA approach

The class hierarchy, the object class and the rule sets of process tree are described briefly and the value of thresholds for the object class is listed in Figure 5. Thematic attribute ID of water bodies means the attribute table value of parcel boundary assigned to 0. BC\_ID means a standard class (ID=6) of building class from point clouds. BH means building height. NDVI is the normalized difference vegetation index. RG means ratio of green band. TH means tree height. CB means the customized brightness value for processing shadow. The symbol of I means the intensity value of Lidar datasets.

**4.2.1 Water class:** Image objects primitives are divided into water and non-water via a thematic layer of water bodies with parcel boundary. The ID attribute of thematic layer is 0. The non-water area is classified into building and non-building.

**4.2.2 Low building and high building classes:** In pre-processing phase, the buildings of Lidar datasets are classified and stored in the standard class (0-9). The 6th class belongs to building category of point clouds. The image object is assigned to the building class with relational point filter by class if the value of BC\_ID is equal to 6. Low building is classified if the value of building height (BH) is less than 3 meter, and the others are assigned to high building class.

**4.2.3 Low vegetation and tree classes:** The Normalized Difference Vegetation Index (NDVI) is one of the most widely useful features for differentiation of vegetation and non-vegetation in urban areas. NDVI is derived from the red and near infrared bands (Rouse, et al., 1973; Carleer, et al., 2006; Nathalie Pettorelli., 2013). Ratio of Green (RG) is an important index

which is a percentage of the green band to the whole bands. The customized feature is created to calculate NDVI and ratio of green as equation (1) and (2) respectively.

$$NDVI = \frac{Mean(NIR) - Mean(Red)}{Mean(NIR) + Mean(Red)} \quad (1)$$

$$RG = \frac{Mean(Green)}{Mean(NIR) + Mean(Red) + Mean(Green)} \quad (2)$$

where NDVI = normalized difference vegetation index  
NIR = near infrared band  
Red = red band  
Green = green band  
Mean = mean value of image pixel layer for objects  
RG = ratio of green band

A measurement value of NDVI is an effective threshold to differentiate vegetation and non-vegetation. Based on the available knowledge of NDVI and the ratio of green, vegetation is first identified if NDVI is greater than 0.24 and RG is less than 0.3. Tree can be distinguished from vegetation if TH is greater than 1.2 meter using height information in the ancillary nDSM dataset. Other vegetation is assigned to grass/low vegetation.

To missing tree or vegetation, we use class-related feature with related border to tree or vegetation and grow the border boundary from vegetation, tree or unclassified respectively. Missing tree is assigned to tree class if the Related Border of Tree (RBT) is greater than 0.5, and missing vegetation is assigned to low vegetation if the Related Border of Vegetation (RBV) is greater than 0.3. A large proportion of tree or vegetation can be effectively classified.

**4.2.4 Shadow objects:** The shadow area is a dark image made by high buildings or tree blocking rays of light. Shadows may vary the spectral properties of image objects. Shaded objects are classified by height, NDVI and neighbour-object relation to identify the corresponding classes. In study area, parts of road and low vegetation objects are covered by shadow. We define as customized brightness feature to detect shadow region. The customized feature is created to calculate Customized Brightness (CB) as equation (3).

$$CB = \frac{Mean(NIR) + Mean(Red) + Mean(Green)}{3} \quad (3)$$

where CB = customized brightness of shadow

The non-vegetation area is classified into shadow and non-shadow via pointing to the values of CB. Shadow objects are labelled from unclassified class if the CB values are greater than 34 and less than 62. Shaded objects are further classified into shaded tall objects and shaded short objects based on height information of nDSM data. Shaded tall objects were classified into shaded tree if a height value TH is greater than 1.2 and NDVI greater than 0.16. Shaded short objects are divided into shaded vegetation and shaded road. Shaded vegetation were those objects with a NDVI value greater than 0.06 and a height value TH of nDSM data is greater than 0.3 and less than 1.2. The other shadow area is assigned to shaded road. Shadow over tree, shadow over vegetation and shadow over road are reassigned the homologous classes. To refine the shadow area, shaded tree, shaded vegetation and shaded road are processed by class-related feature with related border respectively.

**4.2.5 Road class:** The road is classified from non-shadow or unclassified class if the intensity value of Lidar point clouds is less than 83. Missing road in vegetation and shaded road are assigned to road class. Due to only six classes, objects associated with road type are assigned to the road class such as cars, zebra crossing and runway in playground.

**4.2.6 Other processing methods:** To refine the classified results, we select many of other approaches such as merge region, grow region, find enclosed by class, morphologic, shape index features, pixels area and so on. To the same objects class, we use the merge region algorithm on image objects classified as seed and merge these objects region. To smooth the border of objects class, we use morphology approach to create a square mask of 15 x 15 and to add or remove surrounding pixels to an image object by closing or opening operator. To no pixel area which exists in TOP or area pixel which is less than 100 pixels, we use find enclosed by class to assign to the corresponding class. To misclassification of objects, we use class-related feature with related border to tree, vegetation or building, grow the border boundary from related classes and reassign to the classes respectively.

### 4.3 Accuracy Assessment

Accuracy assessment can produce statistical outputs to check the quality of classification results. Two hundred (200) randomly distributed sampled points are created from the study area. Each of these sampled points is labelled to its proper class by visual human interpretation of high resolution TOP imagery, forming the reference dataset. These reference points are compared with classification results at the same locations. The evaluation method is to use these known reference points to assess the validity of the classification results, to compute a confusion matrix based on errors of omission and commission, and to derive the user's accuracy and producer's accuracy for each class, an overall accuracy as well as kappa coefficient.

The confusion matrix is defined in a way that in rows direction the classification dataset is given, while in columns direction the reference dataset. Overall accuracy indicates the percentage of correctly classification objects. The producer's accuracy and error of omission indicate the probability that a classified object actually represents that category in reality, and the user's accuracy and error of commission indicate how well reference dataset objects are classified. Overall accuracy and Kappa coefficient gives an overall assessment to the accuracy of the classification result. User's Accuracy (UA) and Producer's accuracy (PA), Commission Error (CE) and Omission Error (OE), Over Accuracy (OA) and Kappa coefficient (Congalton, R.G., 1991; Jensen, J.R., 1996; Stehman, S.V., 1997) are computed as equations (4), (5), (6), (7), (8) and (9) respectively.

$$UA = x_{ii}/x_{i+} \quad (4)$$

$$PA = x_{ii}/x_{+i} \quad (5)$$

$$CE = 1 - UA \quad (6)$$

$$OE = 1 - PA \quad (7)$$

$$OA = \sum_{i=1}^r x_{ii}/N \quad (8)$$

$$Kappa = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \times x_{+i})} \quad (9)$$

Where  $r$  = the number of rows and columns in the matrix  
 $x_{ii}$  = the number in row  $i$  and column  $i$   
 $x_{i+}$  = the marginal sum of row  $i$ ,  $x_{i+} = \sum_{j=1}^r x_{ij}$   
 $x_{+i}$  = marginal sum of column  $i$ ,  $x_{+i} = \sum_{j=1}^r x_{ji}$   
 $N$  = the total number of observations

## 5. RESULTS AND ANALYSIS

### 5.1 Results of Classification

The study area of Vaihingen/Enz is classified into six objects classes discerned in urban land cover features. The objects classes are water bodies (blue), low vegetation/grass (light green), tree (dark green), low building (purple), high building (red) and road (white). Figure 6 shows the classification results for six classes based on OBIA approach in study area Vaihingen/Enz.



Figure 6. Classification results of the six classes in study area

### 5.2 Results of Accuracy Assessment

The accuracy assessment of the classification results for six classes is performed using a total number of 200 randomly distributed sampling point data. The overall accuracy (OA) of OBIA classification reaches 89.5% in the study area and the Kappa coefficient is recorded as 0.865. The confusion matrix and error statistics of the accuracy assessment for six classes is listed in Table 1. The user's and producer's accuracy for each class is calculated according to equations (4) and (5). Commission error and omission error are measured according to equations (6) and (7). The overall accuracy and the Kappa coefficient are calculated according to equations (8) and (9).

Object Classes	Reference Points						RT	UA (%)	CE (%)
	WA	VE	TR	RD	BU	LB			
WA	16	0	2	0	0	0	18	88.9	11.1
VE	0	31	7	1	1	0	40	77.5	22.5
TR	0	1	36	4	0	0	41	87.8	12.2
Rd	0	1	0	57	1	0	59	96.6	3.4
BU	0	1	0	2	37	0	40	92.5	7.5
LB	0	0	0	0	0	2	2	100	0
CT	16	34	45	64	39	2	200		
PA (%)	100	91.2	80	89.1	94.9	100	OA = 89.5%		
OE (%)	0	8.8	20	10.9	5.1	0	Kappa = 0.865		
Notes	WA=Water, VE= Vegetation, RD=Road, BU= Building, TR= Tree, LB=Low Building, RT= Row Total, CT = Column Total								

Table 1. The confusion matrix and error statistics of accuracy assessment for six classes

### 5.3 Results Analysis

Images objects are generated through multi-resolution segmentation using the different weighting of image layer from high spatial resolution imagery, ancillary nDSM data and thematic data. Neighbouring spectrally similar image objects are merged through spectral difference segmentation. The knowledge base of classification rules are created for the study area. Most of images objects can be well distinguished by using spectral feature or auxiliary data by Lidar datasets and thematic data. The use of ancillary data such as nDSM and Lidar is very helpful for the separation of building and road, tree and low vegetation and shadow over objects using the height information. Parcel map of water bodies as manmade boundaries is very useful in segmentation and classification. It makes the image objects of water bodies identify easily and accurately water class. Parcel map as information carriers can be widely used to the refined classification in urban land cover classification. To identify objects features for classification, quantitative analyses is performed and the threshold values of the chosen class is determined using the customized indexes. To refine the result of classification, object feature, spatial relations, and knowledge rule sets are employed to process the meaningful image objects in this study.

Of course, there exists a few of misclassification objects. Some misclassifications can be observed from the results of classification and the results of accuracy assessment. Most of misclassification happens at spectral similar image area, shadow over objects area and the boundaries of different objects types such as tree and low vegetation, road and vegetation. A possible cause is that the threshold value of customized indexes is lack of sensitivity, especially in shadow over objects area to produce the misclassification between road and vegetation classes. Another possible cause is that there are spectral similarities or spectral overlaps in boundary area. The discrimination of the boundary area has very difficult partly due to nearly identical spectral features and lack of effective spectral and spatial feature to distinguish between low vegetation and tree, building and road classes.

### 6. CONCLUSION

The object based image analysis (OBIA) approach is proved to be an effective and convenient way to combine high spatial resolution imagery and Lidar auxiliary data for classifying urban land cover. The uses of Lidar datasets and nDSM are very helpful for the separation of tree and low vegetation, building and road using the height information. Parcel map of water bodies can be very available for the useful in both classification and segmentation. A few of misclassification area happens at spectral similar image area and the boundaries of different land cover types. For evaluating the classification results, overall accuracy for accuracy assessment reaches 89.5% and Kappa coefficient equals to 0.865 for classification results. To detect single object class like cars in complex urban scenes, it will be further study.

### ACKNOWLEDGEMENTS

First and foremost the authors would like to thank ISPRS Commission III WG/4 work group and German Society for Photogrammetry, Remote Sensing and Geo-information (DGPF) for free provision of the Vaihingen dataset (<http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html>), [Cramer, 2010]). We are very grateful to China Scholar Council for the support to the study. Thanks to the participants in State Key Laboratory of Geo-

Information Engineering for many helpful comments. Thanks Professor Jonathan LI in University of Waterloo for suggestion.

### REFERENCES

- Benz, U. C., Hofmann, P., Willhauck, G. and Lingenfelder, I., 2004. Multi-Resolution, Object-Oriented Fuzzy Analysis of Remote Sensing Data for GIS-Ready Information. *ISPRS Journal of Photogrammetry and Remote Sensing*. Elsevier, Amsterdam, Vol. 58, pp. 239–258.
- Blaschke, T. and Strobel, J., 2001. What's Wrong with Pixels? Some Recent Developments Interfacing Remote Sensing and GIS. *Proceeding of GIS - Zeitschrift für Geoinformations system*, Hühthig GmbH & Co. KG Heidelberg, pp. 12-17.
- Blaschke, T., 2010. Object Based Image Analysis for Remote Sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(2), pp. 1-16.
- Carleer, A.P. and Wolff, E., 2006. Urban Land Cover Multi-level Region-based Classification of VHR Data by Selecting Relevant Features. *International Journal of Remote Sensing*, 27, pp. 1035–1051.
- Congalton, R.G., 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sens. Environ.* 37, pp.35-46.
- Cramer, M., 2010. The DGPF Test on Digital Aerial Camera Evaluation – Overview and Test Design. *Photogrammetrie – Fernerkundung – Geoinformation*, 2(2010), pp.73-82.
- eCognition, 2014. Trimble eCognition Developer User Guide (Version 9.0).pdf, Website: [www.eCognition.com](http://www.eCognition.com) (Accessed 16 February 2015).
- ESRI, 2015. ArcGIS 10.3.1 for Desktop Online Help. Environmental Systems Research Institute, Inc., Redlands, California, USA.
- Franklin, S.E. and Wilson, B.A., 1992. A Three-stage Classifier for Remote Sensing of Mountain Environments. *Photogrammetric Engineering & Remote Sensing*, 58(4), pp. 449-454.
- Gilani, S. A. N., Awrangjeb, M. and Lu, G., 2015. Fusion of Lidar Data and Multispectral Imagery for Effective Building Detection Based on Graph and Connected Component Analysis. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Munich, Germany, Vol. XL-3/W2, pp. 65-72.
- Franz, Rottensteiner., Gunho, Sohn., Markus Gerke. and Jan, Dirk, Wegner., 2013. ISPRS Test Project on Urban Classification and 3D Building Reconstruction. *ISPRS Commission III – Photogrammetric Computer Vision and Image Analysis Working Group III/4 – 3D Scene Analysis*. [http://www.commission3.isprs.org/wg4/complexScenes\\_revision\\_v4.pdf](http://www.commission3.isprs.org/wg4/complexScenes_revision_v4.pdf) (Accessed 16 November 2015).
- Jensen, J. R. 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective (Second edition)*. Prentice Hall, Upper Saddle River, New Jersey 07458, USA, pp.247-252.
- ISPRS Report, 2013. Report on the Joint ISPRS Commission III/IV Workshop “3D Reconstruction and Modelling of Topographic Objects”, Stuttgart, Germany. <http://www.radi.g>

[informatik.tu-muenchen.de/ISPRS/WG-III4-IV2-Report.html](http://informatik.tu-muenchen.de/ISPRS/WG-III4-IV2-Report.html)  
(Accessed 16 February 2016).

Nathalie, Pettorelli., 2013. *The Normalized Difference Vegetation Index, First Edition*, Oxford University Press, Oxford, United Kingdom, pp. 18-29.

Photogrammetric Test Site Vaihingen/Enz, Institut für Photogrammetrie (ifp), Stuttgart University in Germany. [http://www.ifp.uni-stuttgart.de/forschung/photo/test\\_site/index.en.html](http://www.ifp.uni-stuttgart.de/forschung/photo/test_site/index.en.html) (Accessed 16 February 2016).

Rouse, J.W., Haas, R.H., Schell, J.A. and Deering, D.W., 1973. Monitoring Vegetation Systems in the Great Plains with ERTS. *3rd ERTS Symposium, NASA SP-351*, Washington DC, 10-14 December, pp. 309-317.

Sohel, S., Paul, D. and Simon J., 2005. Automatic Classification of Land Cover Features with High Resolution Imagery and Lidar Data: an Object-Oriented Approach. *Proceedings of SSC2005 Spatial Intelligence, Innovation and Praxis: The national biennial Conference of the Spatial Sciences Institute*. Melbourne, pp. 512-522.

Stehman, S. V., 1997. Selecting and Interpreting Measures of Thematic classification Accuracy. *Remote Sensing of Environment*, 62(1), pp. 77-89.

Thomas Hellesen., Leena Matikainen, 2013. An Object-Based Approach for Mapping Shrub and Tree Cover on Grassland Habitats by Use of Lidar and CIR Orthoimages. *Remote Sens.* 5, pp. 558-583.

Zhou, W. and Troy, A., 2008. An Object-oriented Approach for Analysing and Characterizing Urban Landscape at the Parcel Level. *International Journal of Remote Sensing*, 29(11), pp. 3119-3135.