

A NEW FRAMEWORK FOR OBJECT-BASED IMAGE ANALYSIS BASED ON SEGMENTATION SCALE SPACE AND RANDOM FOREST CLASSIFIER

A. Hadavand^{a*}, M. Saadatseresht^a, S. Homayouni^b

^a School of surveying and geospatial information engineering, University of Tehran

^b Dept. of Geography, Environmental Studies and Geomatics, University of Ottawa, Ontario, Canada

Commission VII, WG VII/4

KEY WORDS: Segmentation, Scale parameter, Object-based Image Analysis, Land Cover Classification.

ABSTRACT:

In this paper a new object-based framework is developed for automate scale selection in image segmentation. The quality of image objects have an important impact on further analyses. Due to the strong dependency of segmentation results to the scale parameter, choosing the best value for this parameter, for each class, becomes a main challenge in object-based image analysis. We propose a new framework which employs pixel-based land cover map to estimate the initial scale dedicated to each class. These scales are used to build segmentation scale space (SSS), a hierarchy of image objects. Optimization of SSS, respect to NDVI and DSM values in each super object is used to get the best scale in local regions of image scene. Optimized SSS segmentations are finally classified to produce the final land cover map. Very high resolution aerial image and digital surface model provided by ISPRS 2D semantic labelling dataset is used in our experiments. The result of our proposed method is comparable to those of ESP tool, a well-known method to estimate the scale of segmentation, and marginally improved the overall accuracy of classification from 79% to 80%.

1. INTRODUCTION

Land cover information about the earth's surface is critical in most earth and environmental engineering applications (Berger et al., 2013). To name a few examples, one can consider different studies on urban structure (Voltersen et al., 2014), detection of urban objects e.g. trees (Hirschmugl et al., 2007) or roads (Mokhtarzade and Zoej, 2007), 3D modelling (Samadzadegan et al., 2005) and change detection (Lunetta et al., 2006).

Aerial and satellite remote sensing images provide fast, cheap and accurate data source in land cover mapping. Traditionally machine learning methods are employed to produce land cover maps using remote sensing images. Different parametric (e.g. maximum likelihood) and non-parametric (e.g. K-nearest neighbours and support vector machine) methods are developed to improve the accuracy of predicting the proper label for unknown pixels in the image. A comprehensive review on classification methods in remote sensing land cover mapping could be found in (Lu and Weng, 2007).

One of the impressive challenges in land cover mapping comes from the increase in spatial resolution power of imaging sensors. Usually in the images acquired by these sensors, the size of objects is much smaller than the size of a single pixel. Classical machine learning methods consider individual image pixels as independent units and relegate a land cover label to each of them. In these methods, the interrelationship between neighbour pixels, which belong to the same land cover class, is neglected. In addition, the label predicting process should be repeated for similar pixels in a local neighbourhood. Based on (Lu and Weng, 2007), there is a category of classifiers which consider neighbourhood information, called per-field algorithms. In this method which is also known as parcel based or map guided classifier, instead of single pixels, the homogenous patches of the image are classified. Vector data can help to subdivide the image data and produce patches.

Image patches also could be defined using image segmentation algorithms. This solution opens a new area in the classification of high resolution imagery using per-field category which are also known as object-based or object-oriented image analysis methods (Benz et al., 2004). In object-based classification methods an extra pre-processing step is employed to produce the image objects. Image segmentation algorithms are the most widely used methods for this goal. Segmentation is defined as the process of dividing an image scene into homogenous parts which inherently contains similar pixels and completely different from neighbouring parts (Pal and Pal, 1993). Homogenous image patches as the output of segmentation step are known as image objects and considered as the processing units in the object-based classification. The quality of image objects affects directly the final results of image classification. Ideally, the image object's borders should be coincide to the real objects in the image scene. Shape and size of the image segments are important parameters here. In the most of segmentation algorithms, the size and the shape of image segments controlled by some input parameters. In (Wu and Li, 2009) geographical variance, wavelet transform, local variance, semi-variogram and fractals methods are introduced as quantitative methods to deal with the scale issue in the remote sensing imageries. Local variance method has been more frequently considered in the estimation of segmentation scale parameter (Drăguț et al., 2014; Drăguț et al., 2010).

Due to the diversity of objects in the real world, especially in urban areas, it seems difficult to cope with all kind of objects with a single level segmentation. For taking the local changes into consideration, one popular method is to use the hierarchical or the multi-scale segmentation techniques (Johnson, 2013). In this category of methods, the image is segmented at different scales. Then different solutions are used to select the best segmentation. Analysing the probability of belonging objects to a given class in the hierarchy (Johnson, 2013), adding features from coarser level segmentation to finest level and classify finest level objects (Johnson and Xie, 2013), and majority

voting analysis on pixel level classification in different hierarchy levels to choose the best one are some examples.

Three main challenges in hierarchical techniques are to choose the number of levels and their appropriate scales and the method to integrate them and produce the final result. A few researches are only studied the adding a priori knowledge about the image scene in making the hierarchy of segmentation levels. In addition, there is a lack of methods that deal with the integrating multi-scale segments in the existence of multi-source data. In this paper, we proposed a new supervised method to build a scale-space and hierarchical segmentation, and a rule based method to integrate the image objects in different levels in order to reach to the optimized segments.

2. METHODS

2.1 Image and DSM dataset

ISPRS 2D semantic labelling dataset provides airborne high resolution image and DSM for scientific researches on urban object extraction. The data is captured in urban area over the Vaihingen/Enz, Germany. The dataset delivered in 33 patches, each contains true orthophoto in near infrared, red and green spectral bands. DSM is produced through the dense image matching and used to build true orthophoto of the images. The spatial resolution of both true orthophoto and DSM is 9 cm. Patch number 17 is used in our experiments. Labelled ground truth is provided for this patches in 5 land cover classes including: impervious surface, building, low vegetation, tree and car. A snapshot of the dataset is provided in Figure 1.

2.2 Segmentation scale space

Hierarchical theory originally was developed to analyse the effect of scale on the performance of complex systems. In other words, any complex system includes a number of components that interacts in the system (Hay, 2014). Due to the interaction of different elements, urban areas could be considered as complex systems. From a remote sensing perspective, urban complex system includes the objects with different shape, size and spectral signature. In object-based image classification the main challenge is to create the image objects by subdividing the image pixels using the segmentation algorithms. The difference in shape and size of real world objects, in addition to their spectral signature and brightness variations make their detection, classification and identification, using of a single scale parameter, very difficult.

Using hierarchical principal is one of the solutions to deal with this issue in object creation. Multi-scale image segmentation is proposed based on the hierarchical theory, in which multiple scales are used to create the image objects (Benz et al., 2004). Segmentation scales for building the segmentation hierarchy can be selected manually e.g. in (Gao et al., 2011; Johnson and Xie, 2013; Johnson, 2013) or increased gradually in order to reach the best scale, respect to a pre-defined criteria (Drăguț et al., 2014; Drăguț et al., 2010). None of the proposed methods considers any logical approach to select the scales in the hierarchical analysis. In this paper, we propose a method to choose the number of levels in hierarchy and to select the scale parameter based on the previous knowledge obtained from pixel-based maps. The flowchart of the proposed method is presented in Figure 2.

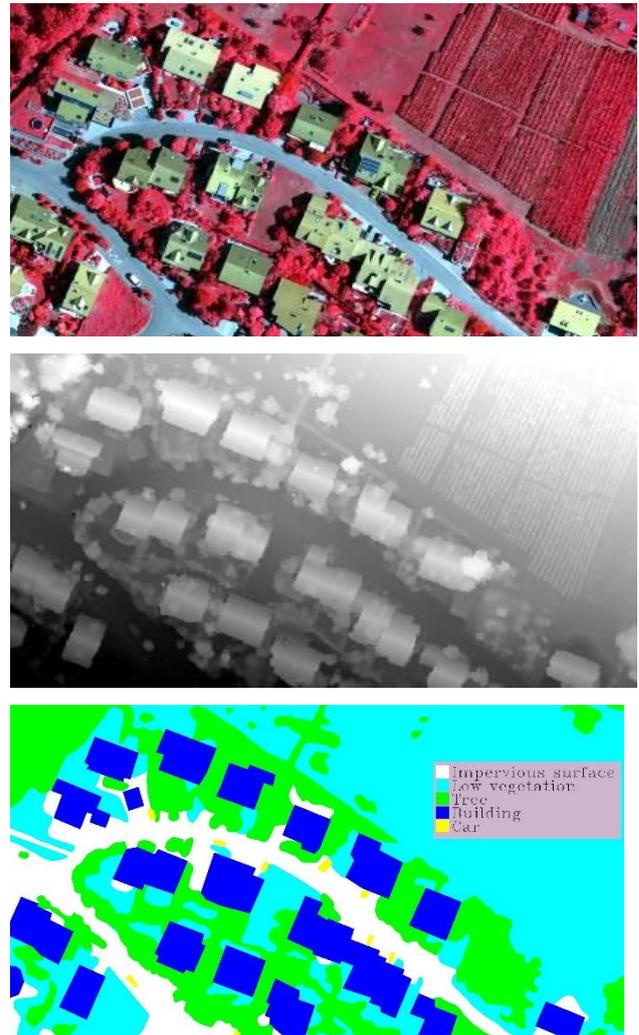


Figure 1. True orthophoto (top), DSM (middle) and labelled ground truth (bottom) used in the experiments

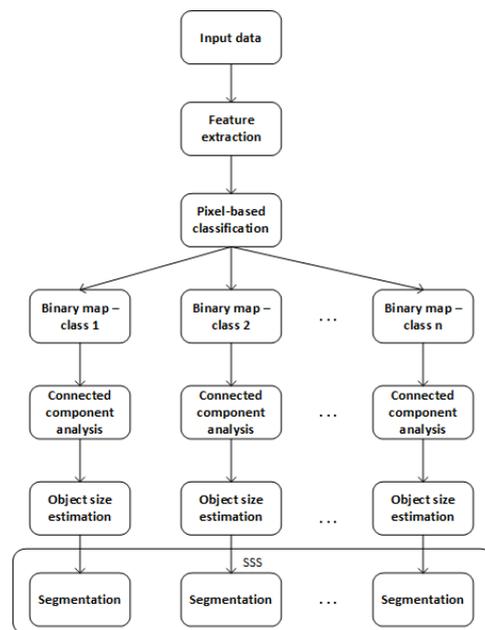


Figure 2. Process of generating segmentation scale space

In our proposed method, at first step, several features are extracted from input data. Then, a pixel based classifier runs on the features and individual land cover maps are converted into the binary maps. Connected component analysis is, then, used in order to integrate the neighbouring pixels of each class. Due to the limited accuracy of pixel-based maps, one expects that there are some misclassified pixels appear as small patches in each class. To reduce this effect small patches, based on the area of smallest considerable object are omitted from the result. Small misclassified patches usually yields to noisy land cover map and can bias the size estimation process towards smaller values. In the next step, the mean size of all patches for each class are computed and considered as the size index for the class. Then, using the obtained size indices, the segmentation scale space are formed. Segmentation parameters tuned in a way that the mean size of obtained image objects equals the computed size index. The number of segmentations in SSS equals to the number of land cover classes.

2.3 Optimizing SSS

After the creation of segmentation hierarchy in SSS, one important issue is “how to use the hierarchy in further processing?” Selecting the optimized level among the hierarchy by optimizing a cost function (Ikoku and Smit, 2013) or analysing the coincidence of a pixel-based map with objects in different levels (Zerrouki and Bouchaffra, 2014) is a basic method to use the hierarchy. In these methods the relationship among different levels of hierarchy and the ability of some levels in better delineation of objects with specific shape and size is neglected. A good solution is to find a way to decide on using the best level in local areas.

Based on our knowledge, there is lack of methods considering the local change and variation of land cover class in scale analysis. In addition, we need new methods to find optimal scale of analysis for multi-source data. In this paper, we proposed a method in order to find the best scale in local areas, by optimizing SSS produced in previous step. The flowchart of the proposed method is presented in Figure 3.

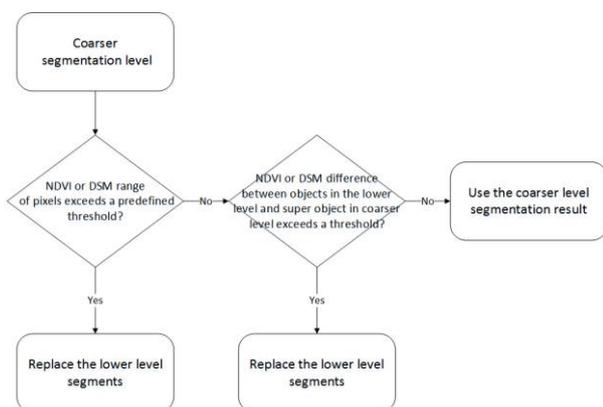


Figure 3. Process of segmentation scale space optimization

The optimization of SSS from the previous step, is based on the analysis of NDVI and DSM values in the objects in different scales. The process starts from the coarsest scale in SSS which contains the biggest image objects. As depicted in the Figure 3, for the image objects with higher scale, the range of NDVI and DSM values for pixels comprising the image object is calculated. High variation interval in NDVI warns about the

mixing vegetated and non-vegetated pixels, and for the DSM it represents the mixture of elevated and non-elevated pixels in an image object. If this condition is satisfied, then the image objects of the lower scale will be replaced in the optimized segmentation result. Otherwise, if this condition isn't satisfied, the objects in lower scale will be compared individually to the super objects in higher scale level. The mean difference of NDVI and DSM in super object and each object in lower level is then considered. To replace each segments in lower level with those of the super object, defining the threshold of NDVI and DSM difference plays an important role. To find the proper thresholds, a grid search is employed to optimize the inter-segment heterogeneity and intra-segment homogeneity. A combination of Moran's index and weighted variance, called global score introduced in (Johnson and Xie, 2011) is used in order to find the best threshold and consequently the best sub objects.

3. EXPERIMENTS AND DISCUSSION

Random forest (RF) algorithm (Breiman, 2001), as a well-known state of the art classifier, is used in the experiments. RF algorithm fuses the result of several CART classifier to get the final classification map. The number of CART classifiers coincides the number of trees in RF classifier and should be set as an input parameter. One hundred trees is used in the experiments in this paper. The features used include three spectral bands, nine different vegetation indices, two water indices, four soil indices, normalized DSM and three topographic properties (gradient, roughness and curvature). As presented in Figure 4, for each land cover class, a binary map is produced and connected patches are detected using connected component analysis

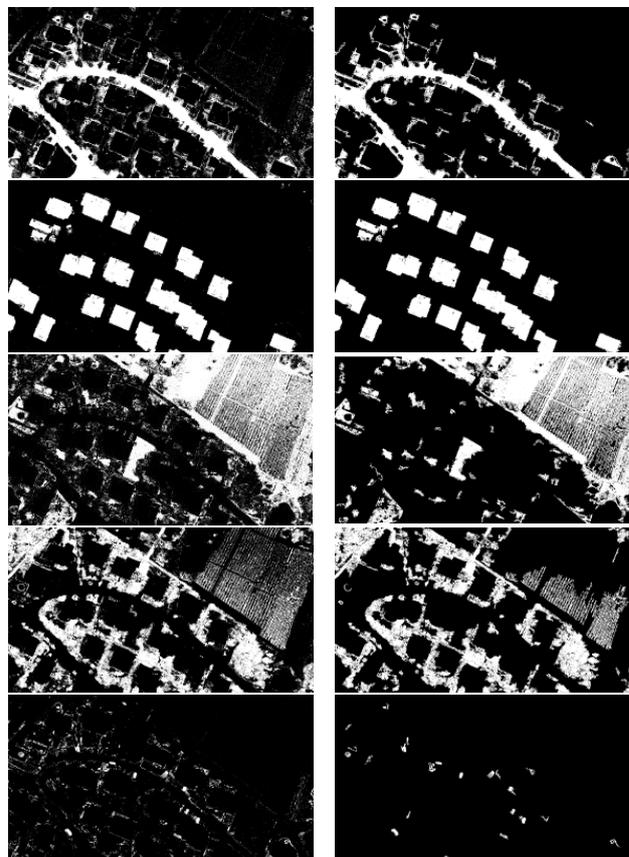


Figure 4. Binary map for each class (Left column) and enhanced binary map by filtering small patches (Right column)

.Small objects in the binary map are excluded by filtering the objects with the area less than four square meters. This will reduce the effect of small patches which mostly contain misclassified pixels in estimating mean object size in each class.

Elimination of salt and pepper noises from the binary maps is evident in the right column of Figure 4. In the next step, the mean size of patches are used to segment the image, for each land cover class. Fractal Net Evolution Approach (FNEA), implemented in eCognition software (Benz et al., 2004) is used to segment image and build the SSS. FNEA uses the region growing methodology which starts the process by selecting some seed points and merge other pixels to the initial seed points until the increase in homogeneity reaches a predefined threshold which is called scale parameter. To build SSS the scale parameter should yield to the objects with mean size according to the value obtained from the pixel-based binary map.

Segmentation process starts from the highest scale value and super objects are built in this highest level. For lower scale values, the segmentation in a lower level is build respect to the boarder of segments in higher level. This process continues until to reach the lowest level and the SSS is built.

Table 1 contains the mean object size and the scale parameters for each level. It is clear that the mean object size obtained from the binary maps is highly affected by noise and small patches. The high number of single pixels and the small patches biases the mean object size towards the smaller values. Consequently, the mean object sizes will be obtained smaller than real objects. In addition, the mean object size values will come close together and reduce the separability of different levels of SSS in terms of scale and size.

Table 1. Mean object size and estimated scale parameter to build SSS

	Class 1	Class 2	Class 3	Class 4	Class 5
Mean object size for binary maps	294	851	313	283	40
Mean object size for enhanced binary maps	9721	22031	14112	11177	950
Estimated scale parameter	170	265	215	185	43

Visual evaluation of SSS for different parts of image, as depicted in Figure 5, demonstrates the coincidence of image objects and the real objects in the different scales of SSS. Objects in land cover classes with bigger size, such as buildings and roads, are well detected in higher scales of SSS. However, for others such as trees and low vegetation and especially for those areas affected by shadows, the lower levels of SSS are necessary.

The next step is to optimize the SSS to reach a unique segmentation, contains the objects from different scales to create the final land cover map. As mentioned earlier, the global score, a combination of Moran's index and weighted variance, is employed in order to find the optimal segments through the SSS in each local area. These values are also prepared for each scale in SSS and for the optimized segmentation results. In addition, for comparison purpose, the results of our proposed method is compared with those of ESP tools prepared to use in eCognition software (Drăguț et al., 2010). Weighted variance is a measure of intra-segment quality measure which reaches

lower values for the homogeneous objects. Contrarily, Moran's index is an inter-segment quality measure which its lower values is more desirable in segmentation and object creation process.

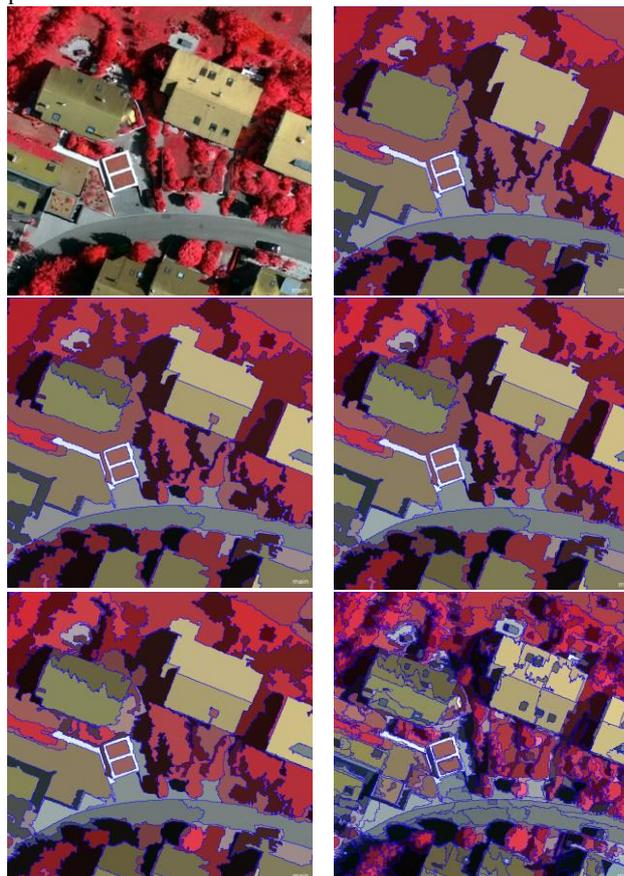


Figure 5. A small part of original image and its SSS

In Table 2 the values of these measures are listed for different spectral bands. It is evident that weighted variance values are increased and Moran's index values are decreased when the scale increases in SSS. For getting the best segmentation results, one should find a balanced state of both these measures. The comparison of image segments obtained by ESP tools and the optimized SSS shows that the proposed method have higher weighted variance and also higher Moran's index. As a result, it would be better to test the ability of both segmentations in classification for the purpose of land cover mapping as well.

Table 2. Weighted variance and Moran's index for segmentations in SSS, ESP tools and proposed method

Segmentation	Near IR		Red		Green	
	WV*	MI*	WV	MI	WV	MI
Scale 1	296	0.29	105	0.44	100	0.34
Scale 2	852	-0.09	281	0.11	263	0.08
Scale 3	897	-0.13	291	0.09	273	0.07
Scale 4	1007	-0.14	323	0.05	299	0.04
Scale 5	1125	-0.14	385	0.00	355	0.02
ESP result	426	0.14	141	0.30	136	0.22
Proposed method	730	0.24	235	0.31	224	0.24

*WV: Weighted Variance, MI: Moran's Index

As we aimed to produce the land cover map using object-based image analysis process, the obtained objects are classified using RF classifier. Then, three different land cover maps including

pixel-based, object-based of ESP segments and object-based map from optimized SSS as presented in

Figure 6 are compared. Visual assessment demonstrates the superiority object-based maps over pixel-based one, and also the result of proposed method over ESP tools method is evident. The results of these land cover maps are also evaluated using well-known criterions including overall accuracy, kappa coefficient and F1-score for each class and are summarized in Table 3. Results show the efficiency of object-based method over the traditional pixel-based. In addition, the overall accuracy and kappa coefficients are improved from 79 to 80 percent and from 0.69 to 0.71 respectively. Furthermore, F1-score shows improvements in impervious surface, tree, and car class and worsen a little for building class.

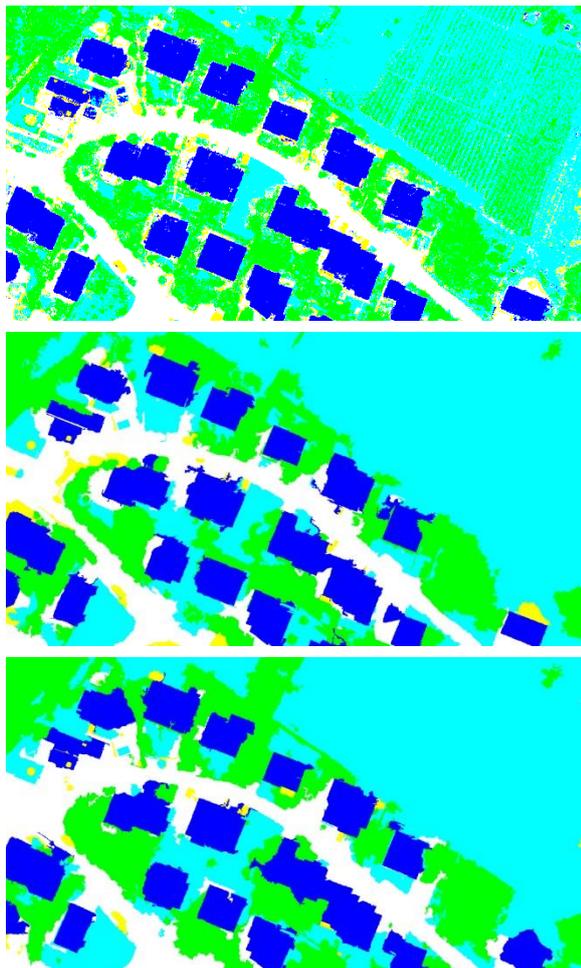


Figure 6. Pixel-based map (top), ESP object-based map (middle) and object-based map of proposed method (bottom)

Table 3. Evaluation on land cover maps obtained from pixel-based classification, object-based classification on ESP tool objects and proposed method objects

	Pixel-based	ESP tool object-based	Proposed method object-based
Overall accuracy	74	79	80
Kappa coefficient	0.64	0.69	0.71
F1-score impervious surface	0.71	0.74	0.77
F1-score building	0.94	0.91	0.90
F1-score low vegetation	0.75	0.82	0.82
F1-score tree	0.68	0.69	0.71

F1-score car	0.14	0.12	0.17
--------------	------	------	------

4. CONCLUSION

Object creation through the segmentation algorithm is a main processing step in object-based image analysis process. It highly depends on the segmentation scale parameter. In this paper, a new framework is proposed for estimating segmentation scale parameter. This method uses the primary land cover maps obtained by classical pixel-based classifier in order to estimate the proper scale for each land cover class and generate the SSS. SSS is then optimized using the NDVI and the DSM data in each object in different scales. Finally, RF classifier is employed in order to produce the final land cover map. The evaluations demonstrate that SSS optimization process produces image objects comparable with those produced by ESP tools. Moreover the potential of proposed method is demonstrated in extracting land cover information. As expected, setting proper scale parameter for segmentation of small objects, such as cars, is more effective in final classification. Here, we see our solution causes to significantly improvement of F1-score for impervious surface, car and tree land cover classes.

REFERENCES

- Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing* 58, 239-258.
- Berger, C., Voltersen, M., Hese, S., Walde, I., Schullius, C., 2013. Robust extraction of urban land cover information from HSR multi-spectral and LiDAR data. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* 6, 2196-2211.
- Breiman, L., 2001. Random forests. *Machine learning* 45, 5-32.
- Drăguț, L., Csillik, O., Eisank, C., Tiede, D., 2014. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing* 88, 119-127.
- Drăguț, L., Tiede, D., Levick, S.R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science* 24, 859-871.
- Gao, Y., Mas, J.F., Kerle, N., Navarrete Pacheco, J.A., 2011. Optimal region growing segmentation and its effect on classification accuracy. *International Journal of Remote Sensing* 32, 3747-3763.
- Hay, G.J., 2014. Visualizing Scale-Domain Manifolds: A Multiscale Geo-Object-Based Approach. *Scale Issues in Remote Sensing*, 139-169.
- Hirschmugl, M., Ofner, M., Raggam, J., Schardt, M., 2007. Single tree detection in very high resolution remote sensing data. *Remote Sensing of Environment* 110, 533-544.
- Ikokou, G.B., Smit, J., 2013. A Technique for Optimal Selection of Segmentation Scale Parameters for Object-oriented Classification of Urban Scenes. *South African Journal of Geomatics* 2, 358-369.
- Johnson, B., Xie, Z., 2011. Unsupervised image segmentation evaluation and refinement using a multi-scale approach. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, 473-483.
- Johnson, B., Xie, Z., 2013. Classifying a high resolution image of an urban area using super-object information. *ISPRS Journal of Photogrammetry and Remote Sensing* 83, 40-49.

- Johnson, B.A., 2013. High-resolution urban land-cover classification using a competitive multi-scale object-based approach. *Remote Sensing Letters* 4, 131-140.
- Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. *International journal of Remote sensing* 28, 823-870.
- Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J.G., Worthy, L.D., 2006. Land-cover change detection using multi-temporal MODIS NDVI data. *Remote sensing of environment* 105, 142-154.
- Mokhtarzade, M., Zoj, M.V., 2007. Road detection from high-resolution satellite images using artificial neural networks. *International journal of applied earth observation and geoinformation* 9, 32-40.
- Pal, N.R., Pal, S.K., 1993. A review on image segmentation techniques. *Pattern recognition* 26, 1277-1294.
- Samadzadegan, F., Azizi, A., Hahn, M., Lucas, C., 2005. Automatic 3D object recognition and reconstruction based on neuro-fuzzy modelling. *ISPRS journal of photogrammetry and remote sensing* 59, 255-277.
- Voltersen, M., Berger, C., Hese, S., Schmullius, C., 2014. Object-based land cover mapping and comprehensive feature calculation for an automated derivation of urban structure types at block level. *Remote Sensing of Environment* 154, 192-201.
- Wu, H., Li, Z.-L., 2009. Scale issues in remote sensing: A review on analysis, processing and modeling. *Sensors* 9, 1768-1793.
- Zerrouki, N., Bouchaffra, D., 2014. Pixel-based or Object-based: Which approach is more appropriate for remote sensing image classification?, *Systems, Man and Cybernetics (SMC)*, 2014 IEEE International Conference on. IEEE, pp. 864-869.