

COMBINATION OF OBJECT-BASED AND PIXEL-BASED IMAGE ANALYSIS FOR CLASSIFICATION OF VHR IMAGERY OVER URBAN AREAS

Bahram Salehi^a, PhD Candidate

Yun Zhang^a, Professor

Ming Zhong^b, Associates Professor

^aDepartment of Geodesy and Geomatics Engineering, ^b Department of Civil Engineering
15 Dineen Drive, Fredericton, New Brunswick, Canada, E3B 5A3

b.salehi@unb.ca, yunzhang@unb.ca, ming@unb.ca

ABSTRACT

Although spatial measures such as texture and shape extracted from very high resolution imagery (VHR) have been successfully employed in pixel-based classifications, the effectiveness of such measures in classification mainly depends on the optimal window size in which spatial measures are calculated. However, an optimal window size is usually subjective and varies for different image and different land cover types. Multiresolution segmentation of object-based image analysis, on the other hand, results objects with different size and shape, which are meaningful and better represent the real size and shape of land cover types. This paper introduces a new approach to land cover classification which benefits from both pixel-based and object-based image analyses. The VHR image is firstly segmented into different levels of segmentations. For each level, one set of spectral measures and two sets of spatial measures, texture and morphology, are extracted and then stacked to the original bands of VHR image forming a several-band image. To determine the contribution of each set of measures in separating urban land cover classes, the separability distance for all class pairs are calculated based on Bhattacharyya distance for each set of measures (i.e. spectral, texture and morphology). A pixel-based maximum likelihood classification is then applied to each set of bands. Results show that adding either texture or morphology to the original bands of VHR image has almost the same effect in increasing the overall classification accuracy. Furthermore, the classification accuracy of buildings and roads increases significantly by incorporation of spatial measures in classification procedure.

KEYWORDS: Feature extraction, Class separability, Pixel-based classification, Segmentation, Texture, Morphology

INTRODUCTION

Urban land cover classification using very high resolution (VHR) satellite imagery has been a challenging task in remote sensing communities since the launch of first VHR satellite in 1999. While the image of many VHR satellites such as IKONOS, QuickBird, and GeoEye contains only four multispectral bands in the visible and near infrared part of electromagnetic spectrum, these four bands are not sufficient for separating spectrally similar land cover types in an urban landscape (Ben-Dor, 2001; Herold et al., 2003; Warner and Nerry 2009). In an urban scene, when VHR imagery is utilized, the high within-class and low between-classes spectral variations lower classification accuracy if the classification is carried out solely based on the spectral information of the image. Generally, an urban landscape composed of vegetation and impervious surface including transportation areas such as roads/streets and parking lots and built-up classes such as buildings and houses. Some impervious surface classes are spectrally too similar to be separated using only spectral information of the VHR image. Tar roofs, roads and parking lots have very similar spectral signature, which even hyperspectral images have limitation in mapping them (Herold et al. 2003). Moreover, the spectral variation within the same land cover type is very high due to the small pixel size of VHR imagery. For instance, houses or buildings with different roof colors are hard to be classified without additional information to the spectral information of the image.

Despite the relative low spectral resolution, the high spatial resolution of VHR imagery makes them a potentially rich source of data for extracting spatial measures such as texture and morphology of land cover types. Several researchers have attempted to develop techniques in which spatial information of the image is incorporated into classification in order for mapping urban areas using VHR imagery. Carleer and Wolf (2006) employed several textural and morphological measures, extracted from a segmented QuickBird (QB) image, into a nearest neighbour object-based classifier to classify the area in a hierarchical scheme of classes including transportation areas and buildings as the last level of class hierarchy. They showed that morphological measures, rather than textural

measures, have higher effect on increasing the separability between buildings and transportation areas. Lu et al. (2010) utilized two texture measures, mean and dissimilarity of red band with a window size of 9x9 pixels, in land cover classification of Quickbird. They showed that adding texture to the spectral bands of Quickbird increases the overall classification accuracy by 11%. They concluded that adding texture to the spectral bands of VHR imagery not only reduces the intra-class spectral variations but also it increases inter-class spectral separability.

In this research, we are going to evaluate the effect of spectral and spatial measures extracted from VHR imagery on improving land cover classification of an urban environment by combination of object-based and pixel-based image analyses. For this, the image is firstly segmented into different levels of segmentations using multi-resolution segmentation algorithm. Then, for each level of segmentation different spectral and spatial measures (spectral, textural and morphological measures) are extracted. Having extracted the measures, they are stacked to the pansharped bands of VHR imagery forming a multi spectral-spatial band imagery. For each image, maximum likelihood classification is performed followed by the accuracy assessment.

STUDY AREA AND IMAGE DATA

The study area is the down town of city of Fredericton, Canada. The area is covered by five major land cover categories including vegetation, buildings, houses, streets, and parking lots. A part of QuickBird (QB) image collected in August 2002 was used in this study. The QB has a panchromatic band (Pan) with the spatial resolution of 0.7 m and four multispectral bands including blue (B), green (G), red(R) and near infra red (NIR) with the spatial resolution of 2.8 m at nadir. As the preprocessing step, four multispectral bands were fused with the panchromatic band using UNB Pansharpening algorithm (Zhang 2004) resulting four multispectral bands with the spatial resolution of 0.7 m (Pansharped QB image) (Figure 1). The classification of this image is very problematic because of the complex mixture of land cover types and the presence of trees' bushes. The classes of house and building are very diverse in terms of spectral signatures or simply colors. There are house and buildings with bright, dark and gray colors, which is hard to classify them as only two classes. Spectral diversity is also significant for streets and parking lots. While some parking lots contain cars, others are empty. Some streets and parking lots have also surface traffic marking, which in turn causes a diverse spectral signature for each class. Another major problem is the tree crown. The image is collected in late summer, when trees have maximum bushes. Consequently, part of streets and small houses are covered by tree crowns, which make the classification of them challenging. Despite such diverse spectral characteristics, we tried to classify the image into five aforementioned land cover types to see how adding spatial measure, extracted from the image, can increase the separability between class pairs and consequently improve classification result.

MULTIRESOLUTION SEGMENTATION

Traditionally spatial measures such as textural measures are extracted by applying a moving window of certain size to the panchromatic band of VHR image. However, it is not an easy task to define an optimal window size, because different land covers (e.g. small building, large building and roads) usually need different window sizes for extracting proper spatial information. On the other hand, multi resolution image segmentation of object-based methods gives meaningful regions which are close to the boundaries of real objects. In a typical urban environment there are small houses, large commercial buildings, narrow streets, wide highways, etc. In this situation a multi resolution segmentation in which object in different sizes and shapes are well represented in different levels of segmentation is desired. For this reason, in this work a multi resolution segmentation algorithm embedded in eCognition Developer (eCognition Developer, 2010) was applied to the pansharped bands of the QB image. Three different parameters should be set up by Analyst in multi resolution segmentation: scale, shape and compactness (eCognition Reference Book, 2010). Finding the optimal parameters for segmentation is a trial and error process which is very time consuming and directly depends on the analyst's experience (Zhang et al., 2010). Instead of trial and error method, we used the Fuzzy-based Segmentation Parameter optimizer (FbSP optimizer) developed at University of New Brunswick (Zhang and Maxwell, 2006, Zhang et al., 2010) to get proper parameters in three levels of segmentation. Figure 2 depicts three levels of segmentation on a small part of the image.



Figure 1: Pansharpened QB image of the study area.

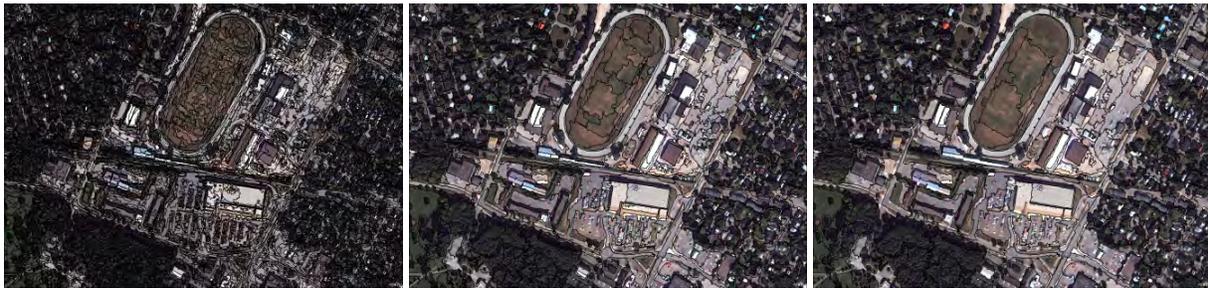


Figure 2: Three levels of segmentation (from left to right: L1, L2, and L3).

SPECTRAL AND SPATIAL MEASURE EXTRACTION

Having segmented the image, several spectral and spatial measures for each individual object can be extracted. Later, these measures are combined with the pansharpened bands of the image in classification process. One set of spectral measures and two sets of spatial measures including textural and morphological measures were extracted in this work. Table 1 shows the measures in each set. Each measure was extracted in three levels of segmentation. Therefore, spectral measures created 12 bands and spatial measures created 39 bands (18 textural bands and 21 morphological bands). These 51 extra bands were then stacked to the original four pansharpened bands creating a 55 spectral-spatial bands image.

Table 1. Extracted spectral and spatial measures from the segmented image

Spectral Measures	brightness, ratio B, ratio G, normalized difference vegetation index(NDVI)
Spatial Measures	Textural-GLCM angular 2 nd moment, contrast, correlation, homogeneity, entropy, standard deviation
	Morphological
	Shape compactness, roundness, density, shape index Extend length, width, length/width

Spectral Measures

Spectral information is the primary source of information in classification process. As mentioned, four spectral measures were extracted for each level of segmentation (see table 1). Brightness is the average of pixel (object) values in four pansharped bands. RatioB describes the amount that band blue of the image contributes to the total brightness for an object and is calculated as the ratio between objects' value in band blue over the summation of objects' value in four bands (eCognition Reference Book, 2010). Similarly RatioG is the ratio between band green and the summation of all four bands. Normalized difference vegetation index (NDVI) was also calculated for the three levels of segmentation. This widely used index is very helpful in separating vegetation from non vegetation areas. Figure 3 depicts brightness and NDVI for three levels of segmentation for small part of the study area. As can be seen, the brightness values of buildings roof are very diverse. This is partly because of diversity of materials used in building roofs which makes the classification of building very challenging. NDVI, on the other hand, has high values for vegetation and low values for non vegetation areas.

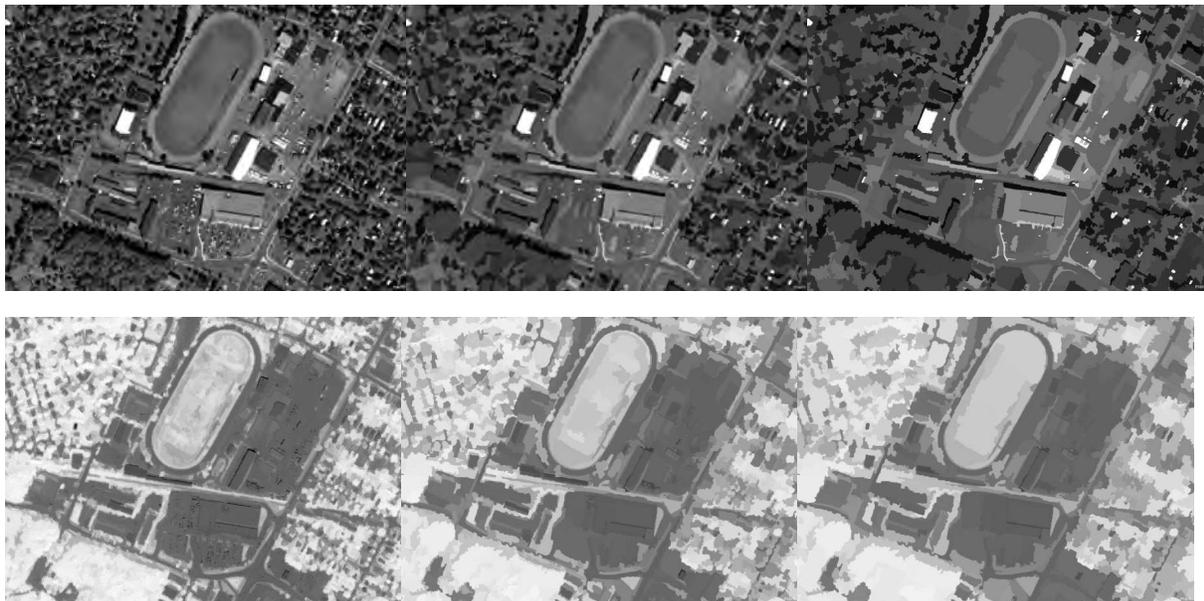


Figure 3: Brightness (top row) and NDVI (bottom row) measures of objects in three levels of segmentation (from left to right: L1, L2, and L3).

Textural Measures

In the past years, a considerable amount of literature has employed texture measures as additional spatial information in the land cover classification process of VHR images to overcome the lack of spectral information (Carleer and Wolff, 2006; Myint, 2007). There are several standard approaches for measuring the texture. Among them, Grey level co-occurrence matrix (GLCM) has been widely used in literature in order for classification of VHR imagery over urban areas, particularly when pixel-based approaches are employed. In this study, six GLCM texture measures (table1) were calculated for each level of segmentation forming 18 textural bands.

Morphological Measures

In a typical urban environment, member of same classes can have different spectral reflectance values, such as black roof and gray roof buildings; consequently, they may have different textural measures. However, such built-up areas have specific morphological characteristics such as shape and extent that can facilitate differentiating various impervious surface classes including buildings, parking lots, and streets. Although morphological measures can be extracted from the original image, which is usually done by applying a morphological operator to the image, segmenting the image allows extracting more meaningful morphological characteristics of objects in the image. In this research a list of seven morphological measures including the shape and extent (table 1) of objects in three levels of segmentations were extracted. The description of each morphological measure can be found in eCognition Reference Book (2010).

CLASSIFICATION

Several spectral and spatial measures in the previous section now must be combined together in conjunction with the original pansharped bands of the image in order to be utilized in classification procedure. We used the stacked vector or logical channel method (Jensen, 2005; Watanachaturaporn et al., 2008), which consider each measure as an extra channel and form a multi spectral-spatial band image. Pixel-based Maximum Likelihood (ML) classification algorithm was applied to the original pansharped and new multi spectral-spatial images to classify the area into five different classes. Table 2 represents land cover classes along with the number of training and testing pixels for each class.

Table 2. Number of training and test pixel used in classification and accuracy assessment, respectively

Class name	No of training pixels	No of test pixels
Vegetation	55906	1158796
Building	27893	145371
House	17272	213517
Road	29145	365649
Parking lot	23750	119946
Total	153966	2003279
% in the entire image	2.3%	30.2%

RESULTS AND DISCUSSION

In order to evaluate, the performance of spectral, textural, and morphological measures on classification, each set of measures (i.e. spectral, textural, and morphological) were considered as one group of bands and stacked to the original pansharped bands of the QB image. Therefore, in addition to the original pansharped image (PS), four new multi-band images were created. The first image has four pansharped bands in addition to 12 spectral measures forming a 16-band image (PS+SP). The second and third image are composed of the previous 16 bands in PS+SP image in addition to 18 textural and 21 morphological bands, respectively creating a 34-band(PS+SP+TXT) and 37-band images(PS+SP+MOR), correspondingly. Finally the last image contains all 51 spectral-spatial bands as well as 4 pansharped bands creating a 55-band image (PS+SP+TXT+MOR). ML classification was applied to each data set to evaluate the effectiveness of each set of measures on increasing the classification accuracy.

Classification Results

As described above, pixel-based ML classification algorithm was applied to the PS image as well as to the new four created multi-band images. For each data set, four accuracy measures including producer, user, and overall accuracies and kappa coefficient are reported in table3. A surprising result was achieved for class of vegetation. The highest PA for this class is resulted with only four PS bands of the QB image. In fact, adding spatial and even spectral measures declines the PA of vegetation area. Another interesting result is that the OA declines when all spectral and spatial measures are incorporated into classification (PS+SP+TXT+MOR) compared to the cases when either spectral-textural (PS+SP+TXT) or spectral-morphological (PS+SP+MOR) measures are incorporated into the classification procedure. Although higher number of bands potentially provides better class separability, the limited number of training pixels results inaccurate parameters estimation in ML classification, leading to a decline in classification accuracy (Salehi et al, 2008).

Spectral-textural and spectral-morphological measures have almost the same effect in increasing the overall classification accuracy. This increase is about 7% compared to when only four PS bands are utilized. The effect of spectral-textural and spectral- morphological measures, however, is more significant in increasing the PA of building and road, respectively. This can be explained by the fact that roads/streets are elongated features and are better represented by their morphological characteristics such as length and rather than their textural characteristics. On the other hand, buildings are well-textured features in VHR imagery and are better represented by their textural

characteristics. Although the PA of class of house is high in all new generated images, especially in the fifth image, this is not a reliable accuracy since the UA of house is significantly low representing an overestimation of this class by the classifier. This is because of the diverse spectral and spatial characteristics of the class of house in the image confirming that classifying houses in VHR imagery is a very challenging task and needs additional data such as ancillary data in addition to the VHR imagery. The class of parking lot has the lowest PA amongst the five classes in all five images. This is also because of the significant spectral and spatial diversity of this class, which even adding additional spectral and spatial measures does not improve the classification result of it.

Table 3. Classification accuracies for different data sets

Class name	ML Classification									
	PS		PS+SP		PS+SP+TXT		PS+SP+MOR		PS+SP+TXT+MOR	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Vegetation	99.0	99.2	95.6	99.7	92.7	99.8	92.4	99.8	87.4	99.9
Building	23.2	59.1	41.1	79.8	59.0	80.4	53.1	79.9	45.2	82.5
House	23.4	37.6	72.7	42.6	86.2	34.0	91.0	31.7	94.8	23.5
Road	70.7	34.7	69.0	39.3	66.7	60.0	76.4	70.3	69.8	87.8
Parking lot	33.0	22.0	37.6	55.6	37.0	83.0	31.0	95.6	25.9	97.0
OA	72.7		77.7		80.0		79.7		74.7	
KC	0.56		0.65		0.68		0.68		0.61	

PA: producer's accuracy, UA: user's accuracy, OA: overall accuracy, KC: kappa coefficient

CONCLUSION

Because of the high inter- class spectral variation and low intra-class separability of impervious land cover types, classification of an urban area using VHR imagery becomes inaccurate if it is performed solely based on the spectral bands of the image. On the other hand, utilizing spatial information, inherent in the image, in the classification leads to an increase of class separability and consequently improves the classification accuracy. In this research, several spectral, textural, and morphological measures were extracted from objects after segmenting the image, and then employed in pixel-based ML classification procedure. Results show that such measures increases the overall classification accuracy and the PA of impervious land cover types, but decreases the classification accuracy of vegetation areas. Specifically, the classification accuracy of roads and buildings were improved significantly by incorporating the spatial measures into the classification procedure. This is because, such impervious classes are better represented by their spatial characteristics such as texture and shape in VHR imagery than large homogeneous classes such as vegetation areas. The result of this research also revealed that classifying small family houses and parking lots in an urban area needs additional data than using exclusively VHR imagery. This is can be ancillary data such as DSM and/or existing GIS layers, which is our next topic of research.

ACKNOWLEDGEMENT

The authors would like to thank City of Fredericton for providing the QuickBird image used in this work.

REFERENCES

- Ben-Dor, E., 2001. Imaging Spectroscopy for Urban Applications in Imaging Spectroscopy (F. van Der Meer. And S.M. de Jong eds), Kluwer Academic Press, pp. 243-281.
- Carleer, A.P., E., Wolff, 2006. Urban land cover multi-level region-based classification of VHR data by selecting relevant features. *International Journal of Remote Sensing* 27 (6), 1035-1051.

eCognition Developer 8.0.1 *Reference Book*, 2010. www.definiens.com

**ASPRS 2011 Annual Conference
Milwaukee, Wisconsin ♦ May 1-5, 2011**

- Herold, M., M.E. Gardner, D.A., Roberts, 2003. Spectral resolution requirements for mapping urban areas. *IEEE Transactions on Geoscience and Remote Sensing* 41 (9 PART I), 1907-1919.
- Jensen, J.R., 2005. *Introductory digital image processing : a remote sensing perspective*, 3rd ed. Prentice Hall, Upper Saddle River, N.J.
- Lu, D., S., Hetrick , E. Moran, 2010. Land cover classification in a complex urban-rural landscape with Quickbird imagery. *Photogrammetric Engineering and Remote Sensing* 76(10), 1159-1168.
- Myint, S.W., 2007. Urban mapping with geospatial algorithms. In: Weng, Q, Quattrochi,D,A. *Urban remote sensing*.CRC press, Taylor & Francis group; Boca Raton, Florida ,pp. 109-135.
- Salehi, B., M.J. Valadan Zouj ,M. Varshosaz , 2008. Projection Pursuit and Lowpass Filtering for Preprocessing of Hyperspectral Images. *World Applied Sciences Journal*, 3(5): 785-796.
- Warner, T.A., F. Nerry, 2009. Does single broadband or multispectral thermal data add information for classification of visible, near and shortwave infrared imagery of urban areas? *International Journal of Remote Sensing* 74(7), 2155-2171.
- Watanachaturaporn, P., M.K. Arora, P.K. Varshney, 2008. Multisource classification using support vector machines: An empirical comparison with decision tree and neural network classifiers. *Photogrammetric Engineering and Remote Sensing* 74 (2), 239-246.
- Zhang, Y., T. Maxwell, H. Tong,V. Dey, 2010. Development of supervised software tool for automated determination of optimal segmentation parameters for ecognition. *ISPRS TC VII symposium-100 Years ISPRS, Vienna, Austria, July 5-7, 2010.*
- Zhang, Y., T. Maxwell, 2006. A trained segmentation technique for optimization of object-oriented classification. *ISPRS commission VII Mid –term Symposium-Remote Sensing from Pixels to Processes, Enschede, the Netherland, 8-11 May 2006.*
- Zhang, Y. 2004. *The Front Cover: Standard and Colour Enhanced Pansharpening of Quickbird Imagery - Sydney, Australia. Photogrammetric Engineering & Remote Sensing, Vol. 70, No. 6.*