STEREO BASED VERY HIGH RESOLUTION SATELLITE IMAGE CLASSIFICATION USING RPCS

Shabnam Jabari, PhD Candidate
Yun Zhang, Professor, P.Eng
University of New Brunswick
E3B 5A3, Fredericton, Canada
sh.jabari@unb.ca
yunzhang@unb.ca

ABSTRACT

Detection of urban objects in very high resolution (VHR) satellite imagery is challenging due to the similarities in the spectral and textural characteristics of urban land cover classes. Therefore, additional information such as elevation data is required for a proper classification. In this study, instead of LiDAR data, elevation information generated from satellite stereo images is used to assist the urban land cover classification in VHR imagery. RPCs are used to generate the elevation information. The classification process is performed using a fuzzy inference rule based system. This method is tested on GeoEye-1 and WorldView-2 satellite imagery. Preliminary results suggest that urban land cover classification is substantially improved by adopting elevation information from the stereo imagery, after it is transferred to the image domain.

Key words: Rational Polynomial Coefficients, Very High Resolution Satellite imagery, Object Based Classification, Stereo Imagery

INTRODUCTION

The main idea of this paper is to present an applicable methodology for very high resolution satellite images classification using stereo information. Considering the existence of details in high resolution satellite images and multiplex resemblance of the pixels to different urban classes, pixel based image classification methods are not of interest in the presented method. Instead, an object based classification methods is preferred since they have higher resemblance to human interpretation skills (Blaschke, 2003). Even using object based classification methods, differentiation of urban objects such as buildings from roads is challenging due to the similarities in their spectral and textural characteristics. Therefore, additional information is required to generate robust classification results (Bouziani M.,Goïta K., He D.C., 2010; Moskal, Styers, & Halabisky, 2011; Thomas, Hendrix, & Congalton, 2003,; Watanachaturaporn, Arora, & Varshney, 2008). This research also uses elevation information as ancillary data to differentiate buildings from roads.

However, because generally the coordinate system of elevation model is different than the image coordinates system, the registration between image and elevation information is a challenging task. In general, the coordinate system of elevation models is either a geodetic coordinate system (latitude, longitude, height) or a map projection system such as UTM (Universal Transverse Mercator), which are not consistent with the image coordinate system (Jabari, Zhang, & Suliman, 2014). Therefore, image to elevation model registration cannot be done accurately using a simple polynomial.

The rudimentary approach would be generating an ortho-photo, which has the same datum as DEM, then, registering the ortho-photo to the DEM. Though, ortho-photo generation destroys the geometry of an image (Moffitt & Mikhail, 1980) and a distorted image does not avail in classification. Another approach is to transfer the ground elevation points to the image and perform the image classification process in the image domain (Jabari et al., 2014). Therefore, the image geometry remains untouched and the elevation of different ground spots are transferred to the corresponding image points. Thus, in addition to the spectral bands, the elevation information is also used as a layer in the segmentation and classification process. In this research, the transformation is done through image RPCs and the output layer is called a height layer. In this study, the combination of the elevation information along with geometric and spectral features of urban objects is used in a fuzzy rule based classification system to differentiate image objects.

METHODOLOGY

In this project, the image is classified into 4 major urban classes: vegetation, road, building, and bare land. In order to detect each of the mentioned class objects, specific characteristics such as geometrical and spectral features are selected. The mentioned features along with the elevation information are used to detect segments belonging to each class. Figure 1 shows the flowchart of the presented method.

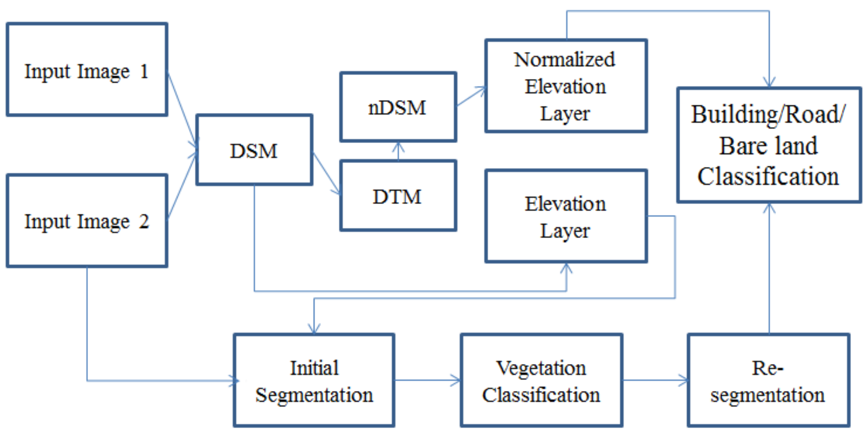

Figure 1: Flow chart of the presented work

Nowadays, elevation information extracted from stereo VHR satellite imagery, using RPCs, has high relative accuracy (Fraser & Hanley, 2005), which means that relative height if the objects is preserved in the elevation layer. Therefore, the extracted elevation model can be used for image classification (Jabari et al., 2014).

Alternatively, in order to classify the image segments, specific characteristic of each class is considered. E.g. buildings are elevated and mostly rectangular shaped features, while roads are elongated objects with lower elevations. Hence, the elevation above the ground can help in differentiating buildings from roads. Here, the normalized Digital Surface Model (nDSM) can be of assistance (Jabari et al., 2014).

Elevation Layer Generation

In this study, using stereo satellite imagery with their associated RPCs, the digital surface model of the area is generated. In this process, since the RPC model generates high relative accuracies in the ground coordinates (Fraser & Hanley, 2005), no ground control points are used.

Typically, using satellite or airborne imagery, the generated DEM is called Digital Surface Model (DSM). In DSM, the elevation model passes over the raised up objects such as trees or buildings. In contrast, the Digital Terrain Model specifies the elevation of the points relying on the terrain. Thus, instead of specifying the elevations of building rooftops or tree tops, the elevations of their footprints on the ground are stated in DTM (Li, Zhu, & Gold, 2010).

In order to generate a DTM out of a DSM, a function looking for the local minima in specific tile size sliding through the image is used from PCI Geomatica software. Using a tile size relevant to the size of the buildings can produce a reliable DTM. Tiles should be larger than the features being removed. After finding local area minimum/maximum values, a moving polynomial function using the local values is applied within the specified tile size to convert DSM to DTM (PCI help). The term nDSM (normalized digital surface model) is referred to the subtraction of DTM from DSM, which specifies the height above the ground (Brunn & Weidner, 1997).

$$nDSM = DSM\text{-}DTM \quad (1)$$

The result DTM, DSM and nDSM all have the same datum. Therefore, once the elevation of each point is transferred to the corresponding image coordinate system using RPCs, the associated normalized elevation (height) of the same spot can also be transferred to the same pixel in the image. Consequently, other than the gray level, each pixel will have an elevation and a height associated with. In this study after transferring the DSM and the nDSM to the image space, they are called elevation layer and height layer, respectively.

Segmentation

Since the elevation specification of the urban features plays an important role for the object differentiation, the elevation layer is also used with the spectral bands in the segmentation process.

In this research, the image segmentation is done in different scales for different image objects. After specific objects are extracted, the remaining unclassified objects are merged and a new segmentation with specific scale relevant to new object size is applied. For example for bare land extraction larger scale and for buildings smaller scales should be applied.

Classification

In this project, a fuzzy rule based system is used for image classification. Using fuzzy logic the borders are not crisp thresholds any more, but membership functions with which each parameter value will have a specific probability to be assigned to a specific class are used. Appending more number of parameters to this classification for example using NIR ratio and NDVI for vegetation classification, better results will be achieved. Using fuzzy logic, classification accuracy is less sensitive to the thresholds (Jabari & Zhang, 2013; Zhang, Zhao, Zhang, & Zhao, 2011).

Mathematically speaking μ_A is fuzzy membership function over domain X. $\mu_A(x)$ is called membership degree which ranges from 0 to 1 over domain X (Zhang et al., 2011).

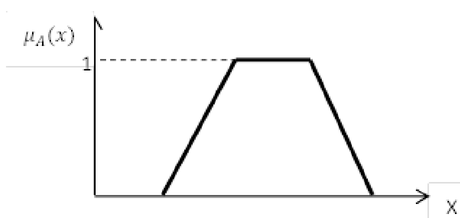

Figure 2: a typical trapezoidal membership function

$\mu_A(x)$ can be a Gaussian, Triangular, Trapezoidal or other standard functions depending on the application. In this research trapezoidal function are used.

Vegetation

Vegetation detection can be performed using NDVI and NIR ratio parameter in a Fuzzy rule based system (Jabari & Zhang, 2013).

$$NIR\ Ratio = \frac{NIR}{NIR + R + G + B} \tag{2}$$

The Fuzzy rule used for vegetation detection is (Jabari & Zhang, 2013):

If **NIR ratio** is ***High*** and **NDVI** is ***High*** then segment is ***Vegetation***.

Building

Typically, buildings are elevated rectangular shaped objects. In the height layer, buildings have higher gray values, but as Figure 3 depicts, the real borders of the buildings do not perfectly overlay that of height layer. Therefore, in this study building places are specified using the height layer and the exact borders are detected using spectral information.

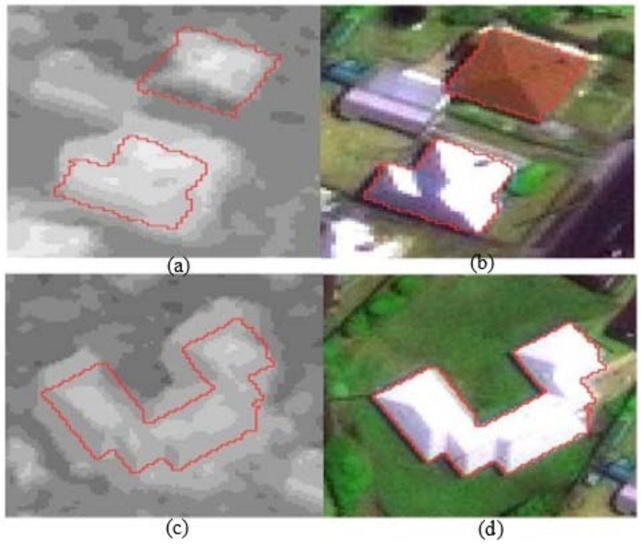

Figure 3: (b), (d): a typical building boundary; (a),(c):building boundary on top of the height layer.

Here, the segmentation optimization method presented in (Tong, Maxwell, Zhang, & Dey, 2012) is used to merge segments belonging to the same building. This process is based on the training segments.

Having assessed the different image segments, it is inferred that buildings have high similarities to rectangular as well as elliptical shapes. Therefore, tow parameters offered in eCognition software, rectangular fit and elliptic fit are used in building extraction.

In General, the produced fuzzy rule for building detection is (Jabari & Zhang, 2013):

If **Height** is ***high*** and **Rectangular fit** is ***high*** and **elliptic fit** is ***fairly high***, then segment is **Building**.

Road

Roads are elongated objects with smooth surface having low variations in the gray values. Therefore, parameters which show lengthened objects and gray value variance can be used for defining roads. I_{cm} and I_e, both ranging from 0 to 1, are used for defining elongated objects; equations 3 and 4 represent the related mathematical formulas (Bouziani, Goita, & Dong Chen, 2010).

$$I_{cm} = \frac{2\sqrt{\pi. Area(object)}}{perimeter(oject)} \quad (3)$$

$$I_e = \frac{Area(object)}{[length(object)]} \quad (4)$$

Standard deviation shows variation of pixel values in a specific segment. Since the surface of the roads are smooth with low variations in gray levels; Thus, the lower the standard deviation of a segments, the higher the possibility to be assigned to the road class.

Considering the above specifications for road, the produced fuzzy rule is (Jabari & Zhang, 2013):

If **Height** is ***low*** and **Icm** is ***low*** and **Ie** is ***low*** and **Std** is ***low*** then segment is **Road**.

Bare land

The remaining unclassified low elevated parts of the image are assigned to the bare land class.

Contextual check

As presented in (Jabari & Zhang, 2013), after image classification, contextual check is needed. For example if a small segment is surrounded by road segments, the mentioned segment probably is associated to a vehicle and must be assigned to the road class.

RESULTS AND CONCLUSION

Figure 4(a) shows a part of the original image and Figure 4(b) shows the classification result of the same area.

Figure 4: (a) original image; (b) Classification Results

In order to generate the ground truth data to check the accuracy of the presented results, after segmentation, the same parts of the images were classified manually into one the four aforementioned classes. Later on, the pixel labels were checked across the ground truth data. The confusion matrices and the associated accuracy assessment parameters are presented in Tables1 to 3.

Table 1:WorldView2 data Classification confusion matrix

	Building	Road	Vegetation	Bare Land	Sum	User's Accuracy
Building	5341	231	73	871	6516	**0.82**
Road	149	2846	39	347	3381	**0.84**
Vegetation	12	0	2018	33	2063	**0.98**
Bare Land	839	473	51	5124	6487	**0.79**
Sum	6341	3550	2181	6375	18447	
Producer's Accuracy	**0.84**	**0.80**	**0.93**	**0.80**		

Table 2: GeoEye data Classification confusion matrix

	Building	Road	Vegetation	Bare Land	Sum	User's Accuracy
Building	10786	341	13	935	12075	**0.89**
Road	541	6273	86	736	7636	**0.82**
Vegetation	26	56	4440	73	4595	**0.97**
Bare Land	1892	989	112	10523	13516	**0.78**
Sum	13245	7659	4651	12267	37822	
Producer's Accuracy	**0.81**	**0.82**	**0.95**	**0.86**		

Table 3: Kappa Coefficient and overall accuracy for both datasets

dataset	Kappa Coefficient	Overall Accuracy
WordView2	0.76	0.83
GeoEye	0.79	0.85

As can be seen from tables 1 to 3, the presented method produces high accuracies in the image classification. However, still around 20% of the pixels are miss-classified. Some of the buildings are missed from the building class and are miss-labelled as bare land generally because the associated height is not true. In other words, the DSM to DTM function in PCI Geomatica software has failed detecting the terrain in some places. Therefore, since the height above the ground for the associated segments is low, they are not labeled as buildings and are assigned to the bare land class. Using a more accurate height layer will be of great help in increasing the classification accuracy.

ACKNOWLEDGEMENT

This research has been funded by National Sciences and Engineering Research Council (NSERC) of Canada. We appreciate their support. We also acknowledge the Space Agency for providing GeoEye satellite images and also 2011 IEEE-DigitalGlobe Data Fusion Contest for providing the Brazil WorldView2 images.

REFERENCES

Blaschke, T. 2003. Object-based contextual image classification built on image segmentation. *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, Washington DC. pp. CD-ROM.

Bouziani M.,Goïta K., He D.C. 2010. Automatic change detection of buildings in urban environment from very high spatial resolution images using existing geodatabase and prior knowledge. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*, 143-153.

Bouziani, M., Goita, K., & Dong Chen, H. 2010. Rule-based classification of a very high resolution image in an urban environment using multispectral segmentation guided by cartographic data. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 48*(8), 3198-3211.

Brunn, A., & Weidner, U. 1997). Extracting buildings from digital surface models. *International Archives of Photogrammetry and Remote Sensing, 32*(3 SECT 4W2), 27-34.

Fraser, C. S., & Hanley, H. B. 2005. Bias-compensated RPCs for sensor orientation of high-resolution satellite imagery. *Photogrammetric Engineering and Remote Sensing, 71*(8), 909.

Jabari, S., Zhang, Y., & Suliman, A. 2014. STEREO-BASED BUILDING DETECTION IN VERY HIGH RESOLUTION SATELLITE IMAGERY USING IHS COLOR SYSTEM. *Submitted to:IGARSS2014,*

Jabari, S., & Zhang, Y. 2013. Very high resolution satellite image classification using Fuzzy rule-based systems. *Algorithms, 6*(4), 762-781.

Li, Z., Zhu, C., & Gold, C. 2010. *Digital terrain modeling: Principles and methodology* CRC press.

Moffitt, F., & Mikhail, E. 1980. Photogrammetry. *I-Lapper & Row Publisher, 648*

Moskal, M., Styers, D. M., & Halabisky, M. 2011. Monitoring urban tree cover using object-based image analysis and public domain remotely sensed data. *Remote Sens., 3*, 2243-2262.

Thomas, N., Hendrix, C., & Congalton, R. G. 2003,. A comparison of urban mapping methods using high-resolution digital imagery. *Photogramm. Eng. Remote Sensing, 69*, 963-972.

Tong, H., Maxwell, T., Zhang, Y., & Dey, V. 2012. A supervised and fuzzy-based approach to determine optimal multi-resolution image segmentation parameters. *PE&RS, 78*(10), 1029-1044.

Watanachaturaporn, P., Arora, M. K., & Varshney, P. K. 2008. Multisource classification using support vector machines: An empirical comparison with decision tree and neural network classifiers. *Photogramm. Eng. Remote Sensing, 74*, 239-246.

Zhang, C., Zhao, Y., Zhang, D., & Zhao, N. 2011. **Application and evaluation of object-oriented technology**
in high-resolution remote sensing image classification. *Control, Automation and Systems Engineering (CASE), 2011 International Conference On,* Singapore. pp. 1-4.