Development of an Automatic Land Use Extraction System in Urban Areas using VHR Aerial Imagery and GIS Vector Data

Seyed Ahad Beykaei[1], Ming Zhong[2], Sajad Shiravi[3], Yun Zhang[4]

Received: 13.01.2013 *Accepted: 20.06.2013*

Abstract:

Lack of detailed land use (LU) information and efficient data collection methods have made the modeling of urban systems difficult. This study aims to develop a novel hierarchical rule-based LU extraction framework using geographic vector and remotely sensed (RS) data, in order to extract detailed subzonal LU information, residential LU in this study. The LU extraction system is developed to extract residential LU at a fine spatial level̃ parcel through morphological analysis. First, a novel hybrid pixel- and object-based land cover (LC) classification system, coupled with a sophisticated GIS post-classification correction process, is developed to extract land cover, including vegetation, parking lot, and bare soil, required for LU classification. The land cover classification system developed results in an overall accuracy of 96.4%. Residential LUs are then extracted by examining the morphological properties of individual parcels (which are derived from RS and geographic vector data) using a binary logistic model, which results in an overall accuracy of 97.5%. The above results show that the LU classification expert system developed can classify and then divide large zones with mixed LUs into single-LU subzones with a high accuracy. Therefore, it has a significant value to address several persistent issues caused by using large zones in urban modeling, such as intrazonal travel and mixed-LU zones.

Keywords: Land use classification, land cover classification, remote sensing, morphological analysis

Corresponding author E- mail: beykaei@gmail.com
1. Ph.D Candidate, Department of Civil Engineering, University of New Brunswick, Canada
2. Associate Professor, Department of Civil Engineering, University of New Brunswick, Canada
3. MSc Student, Department of Civil Engineering, University of New Brunswick, Canada
4. Professor, Department of Geomatics and Geodesy, University of New Brunswick, Canada

1. Introduction

Land use (LU) information is important for urban and transportation planners to understand spatial orientation of existing activities and to forecast urban changes or trends of development over the space-time continuum. Lack of detailed land use information has forced city planners and modellers to use large aggregated zones in their models, and consequently accept undesirable approximations and errors in their analyses and planning workflow. Since traditional acquisition of LU data takes several months or years to complete, frequent land surveys are thus impractical. There is also confidentiality issue when requesting detailed LU information from an organization or a government agency. However, urban planning and especially integrated LU and transportation modeling process require up-to-date and detailed land and floorspace inventory data. Unfortunately, there is no traditional way to collect such data quickly and efficiently. Therefore, it is highly desirable to develop a tool to assist planners to collect such data in a quick, efficient, reliable and non-intrusive way.

Remote sensing (RS) provides fast and efficient ways for monitoring urban LUs and it is growing as a promising tool to extract/update detailed LU information. It offers a great potential for enhancing current urban planning processes. Preliminary investigations from this study indicate that land cover information extracted from RS imagery is vital for extracting detailed urban LUs. Most of previous researches extract land cover information from images (such as bare soil, green area, building, and traffic area) or at best man-built features at very coarse level [Zhang, 1999; Zhang, 2001; Mundia, 2005; Deng et al, 2005]. Although many studies have been dedicated to extracting land cover information using RS imagery, many issues/challenges remain [Gomarasca, 1993; Hill, 1996; McGibbon and Eyton, 1996]. Further, man-built features in the urban area (such as parking, street, and building) are built by similar materials (e.g., concrete) and therefore, they show similar spectral signatures, which make land cover classification difficult. Moreover, it is found that there is a need to incorporate very high resolution (VHR) imageries and spatial information for LU classification purpose in urban areas (Beykaei et. al., 2010; Liu and Prinet 2005; Mohaptra and Wu 2008; Pacifici et. al 2007; Xiao et al. 2004). Chan et al. (2009) mentioned that "… it seems the choice for extracting land uses in urban areas is VHR satellite images".

It is found that extracting LU information is the most challenging part of classification process and it needs other auxiliary data and significant post-classification analysis. Mesev (2005) mentioned that LU classification inevitably relies on auxiliary GIS data (non-spectral data) as well as images (spectral data) to improve the results, due to its abstract nature. This is because that LU is defined based on its function, rather than its physical or chemical properties. The function of a LU cannot be directly observed, but need to be generalized through its various components, such as building size, the quantity of parking and vegetation area. In this regard, a far less number of studies have been devoted for extracting urban LU information via remotely sensed imageries.

Barnsely et al. (2003) used LiDAR (Light Detection And Ranging) and multispectral image data to determine urban land uses through an analysis of the spatial composition of buildings. Three morphological properties of building, namely roof area, compactness and height, were examined in their study. Yoshida and Omae (2005) studied the form and structure of urban features in Tokyo based on their urban landscape model (ULM) using RS data. They used city blocks (spaces between street grids) as study unit. The interrelationships as well as geographical distribution of blocks were interpreted based on their morphological properties. In this regard, they concluded that the correlation and interrelation of the parameters (morphological properties) are important to extract LU information from images. For extracting residential and industrial areas from IKONOS and LiDAR data, they mentioned by quoting from Pesaresi and Bianchin (2001) that morphological properties and structural relations must be employed as extra bands besides spectral wavebands in per-pixel image classification. The key issues they raised are (i) what combination of properties and relations provides the best separation of the two land use categories, and (ii) are they sufficient to identify areas of each land use unambiguously?

Despite the fact that the above studies attempted to improve LU classification in different fashions, only limited enhancements have been achieved. Further, no one has really answered the two above questions raised by Pesaresi and Bianchin (2001) ten years ago. No solid and effective method has been materialized.

Overall, the literature demonstrates that the application of remote sensing technology in the urban and transportation planning/modeling is promising, yet still immature, especially for extracting detailed LU information from complex urban areas. Moreover, it also suggests that auxiliary geographic data is required for an accurate LU classification process. This literature review shows that most of previous studies in this area are about land cover classification. However, from the view point of planners, land use information is more important than land cover in their modeling and forecasting workflow, to which few studies have been dedicated. This paper proposes a rule-based LU extraction system using geographic vector data and remotely sensed imagery for extracting subzonal residential LUs.

2. Study Area and Data

This research uses the City of Fredericton, Canada as the testbed. Due to the abstract nature of land uses, several types of data, including geographic vector and remote sensing (RS), are used to discover correlation/association rules related to urban LU classes. Based on the results of previous research on land cover/use classification [Beykaei et al., 2010], it is found that the similarity of spectral characteristics of urban land covers/uses requires incorporating geometrical/spatial information within the classification process. In this regard, ancillary data such as geographic vector data (hereafter referred as "geographic data") plays an important role to understand LU patterns and their association rules. Due to significant processing time required with each image, a bunch of sample dissemination blocks (DBs) from the testbed are selected and considered. Please note that DB is the smallest zone unit from the Canadian census. The testing area includes 43 DBs with a variety of LUs, ranging from highly complex downtown area at the center of the city, different residential LUs, institutional and industrial area, to exclusive commercial area. The following geographic data (which are GIS shapefiles) are used in this study: digital Property Map (DPM), building footprint, and street network. DPM dataset contains New Brunswick tax assessment records, which divides the entire province into parcels with known LU purposes. Reference data on the size, shape and spatial location of building objects in the study area are also available in the format of geographic vector layer provided by the City of Fredericton. However, the building footprint data has no information about activity uses. In order to assign LU information to the building footprint data, they are spatially joined with the tax assessment records from the DPM data. Street network is another geographic data used in this study which is a poly-line shape file and has the information about each road segment.

Besides geographic data, digital aerial images collected by the City of Fredericton in 2008, which have been corrected and ready for image processing (also called orthophotos), are used to extract land covers. The orthophotos are VHR images with 0.15 m resolution, 8 bit depth, multispectral with three bands of Red, Green, and Blue (RGB). Each image has a size of 6667×6667 pixels. In order to cover the entire testing area, 34 orthophotos are used for segmentation and classification purposes. Figure 1 shows the orthophotos used, sample DBs, and built areas to be analyzed.

3. Method of Land Use Extraction System

Due to the abstract nature of land uses, it is difficult and sometimes impossible to extract LUs directly from RS imagery. A morphological analysis with RS and auxiliary geographic data are required to develop association/correlation rules with respect to residential LU. Land covers are first extracted using RS imagery. First, a hybrid pixel- and object- based land cover classification method is developed to extract urban land covers. Extracted land covers, including parking lots, vegetation areas, and bare soils, are essential for the morphological analysis of urban features. In this regard, geographic data and land covers extracted are used to examine and recognize different land use patterns at the parcel-level. A rule-based LU extraction system is developed to differentiate residential versus non-residential LU.

3.1 Method of Land Cover Classification

Deriving association/correlation rules for LU classification utilizing urban physical objects necessitates accu-

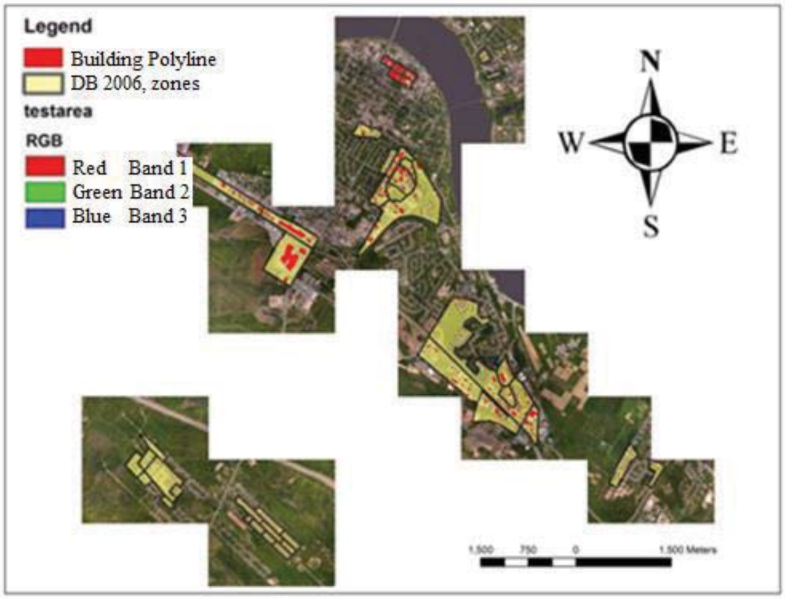

Figure 1. Orthophotos, sample DB zones, and testing area

rate and detailed land cover information at a fine spatial scale. As mentioned in the literature review, low and medium spatial resolution of images are not useful for this purpose due to their low spatial resolution. In addition, both pixel- and object-based land cover classification approaches, have their limitations to extract accurate land cover information. Therefore, in this study, a novel hybrid approach, which benefits from both object- and pixel-based methods, is developed for land cover classification using VHR aerial imagery from City of Fredericton the study area.

Preliminary analysis indicates that land covers, which are important for LU classification, include buildings, streets, parking lots, green areas, and bare soil. On the other hand, several geographic data, such as street network and building footprint, are collected for further analysis. Street network data is usually provided by municipalities or transportation organizations in many communities. However, building footprint data may not be available in many of them. In this regard, some experiments were conducted for extracting buildings using various remote sensing data, such as QuickBird imagery and elevation (e.g., digital elevation model (DEM)) data. It is found, however, that building extraction accuracy is low and itself is still an immature research area (Beykaei et. al., 2010). Furthermore, urban land cover classification indicates that there is a problem of differentiating traffic areas (i.e. streets and parking lots) with buildings. Spectral information alone cannot differentiate these man-built features constructed with similar materials (e.g., concrete or asphalt). Therefore, in this research we assume that buildings are known or well classified to better derive association rules concerning LU pattern recognition.

Based on the above discussion, vegetation area, parking lot, and bare soil are the three urban features required for LU classification purpose in this research. The details of proposed hybrid pixel- and object- based method, which is a hierarchical rule-based system, to extract three above land covers are summarized in Figure 2.

First, the pixel-based classification (see the top center of Figure 2) is carried out with the classification software, ENVI 4.7, in order to perform an initial extraction of all land covers from the study area including vegetation areas, parking lots, bare soils, buildings, and streets using maximum likelihood (ML) classifier. Because we assume that buildings and streets are known or well classified, using the geographic data (building footprint and street network), the classified areas occupied by the two classes are deleted from the results of the pixel-based classification. The final results of the rest of land covers extracted (including parking lot, bare soil, and vegetation area) are then converted to the geographic vector layers for further object-based analysis, so as to be consistent with the geographic data and the results of object-based classification.

In the object-based classification process, first, land cover segmentation is carried out with the classification software, ENVI 4.7, by which the testing areas were divided into different unknown features/objects based on their spectral information. The results of pixel- and object-based classification approaches are then fused into a hybrid rule-based classification process to classify parking lots and bare soils, which will be explained later in detail as the following.

Extracting vegetation areas

Vegetation areas are first extracted due to the fact that they cover the majority of the study area and can be extracted with a higher accuracy than the rest. The flowchart on the far left of Figure 2 shows the process of extracting vegetation areas. In order to do so, the GreenRatio values for all segmented features (derived from the land cover segmentation process) are calculated in terms of their mean values of RGB bands as below:

$$GreenRatio=G_b/(G_b+R_b+B_b) \tag{1}$$

Where G_b is the mean value of green band, R_b is the mean value of red band, and B_b is the mean value of blue band for the segmented features. Based on trial and errors and accuracy assessment, the threshold value of 0.35 for the GreenRatio proves to be most effective in terms of distinguishing vegetation from non-vegetation areas. In this regards, all the features with the GreenRatio equal to and greater than 0.35 are classified as vegetation areas.

Despite the result of vegetation extraction from the above method is very accurate, visual interpretation and accuracy assessment of the testing areas (10% of the total area of the testbed) show that there are some misclassification between the areas of "dead" trees and parking lots/buildings and also a few bare soil areas are classified as vegetation. Two more steps presented in the following sections are added onto the above results to fix the misclassifications and to refine the vegetation

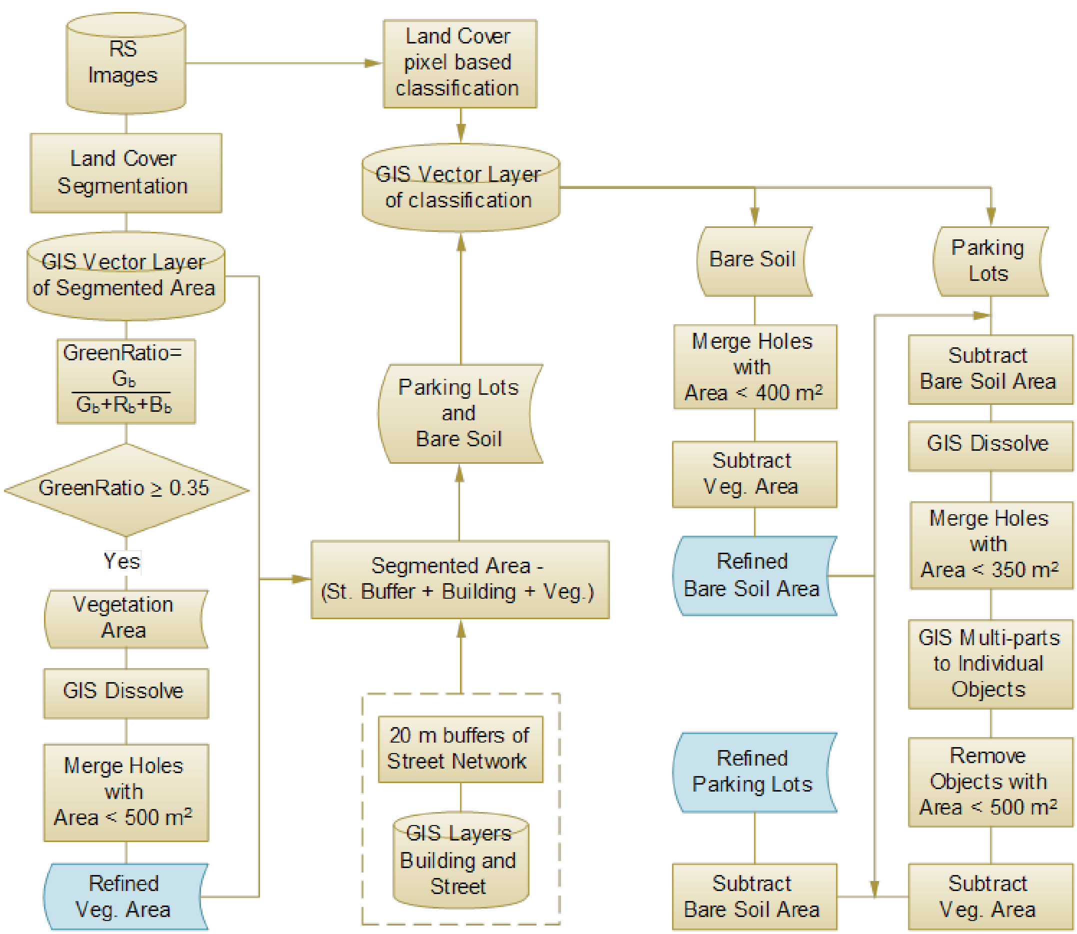

Figure 2. Hierarchical pixel- and object-based land cover classification system

areas, as shown in Figure 2. First, all the vegetation areas are dissolved together to create one GIS layer. Second, all the holes enclosed by the vegetation with an area of 500 m^2 or less are extracted and merged with the vegetation layer. This area threshold is selected based on the trial and error method to minimize misclassifications.

Extracting bare soils and parking lots

Pixel-based classification is used to perform an initial extraction of bare soils and parking lots using maximum likelihood method. It is found that it is relatively easy for extracting bare soils because of its homogeneous characteristics and similar spectral information. However, parking lots are found to be spectrally diverse, as they can contain many different objects, such as paved areas, cars, flowerbeds, strip markings, containers, trailers, and sheds. For this reason, extracting parking lots is challenging and requires careful considerations.

As mentioned above, this study uses a hybrid pixel- and object-based classification method to tackle this issue. For this purpose, auxiliary geographic data such as street buffer and building footprint are used, so that we assume that buildings and street areas are known based on the geographic data provided by City of Fredericton. The street areas are determined based on a buffer of 20 meters over the street centerline (this buffer threshold is selected based on a trial and error method). And finally, street, vegetation, and building areas (based on their footprints) are subtracted from the results of segmentation. The remaining areas are considered to be either parking lots or bare soils.

The procedure for extracting bare soils and parking lots are presented on the right side of Figure 2. In order to mitigate the misclassifications, the holes within the bare soil and parking lot classes with the area less than or equal to the defined thresholds are “dissolved” and then merged into the corresponding classes. Again, the area thresholds are based on trails and errors and accuracy assessment results. Based on our experience and previous studies, it is understood that vegetation areas can be extracted most accurately than all other land covers due to its homogeneous nature and specific spectral information (e.g., near infrared). Therefore, at the end, the vegetation layer is subtracted from the bare soil and parking lot layers to reduce the misclassification between such layers and vegetation areas. Finally, the resulting layer from parking lot classification is subtracted from the bare soil class to remove the misclassification from the strip markings, which were previously misclassified as bare soils. More details about this process were presented in an earlier study of this research [Beykaei et. al., 2011].

The results of extracting vegetation areas, parking lots, and bare soils, through the above proposed hybrid method, are used in the following morphological analysis for the LU classification process.

3.2 Method of Land Use Extraction

In order to define the signatures of the residential LUs, several morphological properties derived from geographic and RS data are examined in the classification system. The following parcel morphological properties are considered as potential LU classification indicators, including:

1. Morphological properties derived from geographic data: sum of building area (sB_Area), sum of building perimeter (sB_Per), average building compactness (AvB_Com), average building height (AvB_Height), built ratio (Built_Ratio, which is the ratio of total building area and parcel area), floorspace area ratio (FAR), parcel area (P_Area), distance to CBD (Dis_CBD), distance to nearest local road (Dis_Local), distance to nearest arterial (Dis_Arterial), distance to nearest collector (Dis_Collector), and distance to nearest freeway (Dis_Freeway),
2. Morphological properties derived from RS data: vegetation ratio (Veg_Ratio), parking ratio (Park_Ratio), and bare soil ratio (BareS_Ratio).

The sample area in this study contains 824 parcels which are divided into two subsets; 75% of these parcels are used for training and the other 25% are used for testing the classifiers/models.

Urban LUs are categorical responses which consist of a set of categories, in this case, residential vs. non-residential. Different properties derived from RS and geographic data are tested through logistic regression models (LRM). LRM is a form of generalized linear model (GLM) for categorical and discrete response data. This model has a linear form for the logit of the probability which is calculated as (Agresti, 1996):

$$logit[\pi(x)]=log((\pi(x))/(1+\pi(x)))=\alpha+\beta x \qquad (2)$$

An alternative formula for this regression to calculate the probability of being in one class (e.g., Residential vs. non-Residential) is in the form (Agresti, 1996):

$$P(Res)=exp(\alpha+\beta x)/(1+exp(\alpha+\beta x)) \quad (3)$$

where $\pi(x)$ denotes the "success" probability, α is the constant parameter, β is the vector of variable's coefficients to be estimated based on maximum likelihood method in the regression analysis, and x is the vector of independent variables.

In order to identify best indicators for differentiating residential vs. non-residential parcels, a binary logistic model is tested with the training dataset. The step-wise regression is used to select effective indicators (independent variables) to be used in the binary logistic model. In other words, such a model is used to select a subset of parcel properties, which will be considered as the indicators in the LU classification model, in order to provide the best extraction results for the residential LU.

All the tests are carried out at the 95% confidence interval for Exp(β), where β is the coefficient of independent variable to be estimated by the model. In this regard, the probability of a parcel being a residential or a non-residential class is calculated using the binary logistic model. The cut-off probability for a parcel to be assigned into either LU class is set as 0.5. For example, if the probability of a parcel belonging to the residential class vs. non-residential is calculated as 0.42 (P(Res)=0.42), it is less than the cut-off value 0.5, which means the parcel should be classified as non-residential rather than residential.

4. Experimental Results

4.1 Land cover classification results

First, land cover segmentation is carried out with the classification software, ENVI 4.7, by which the testing area were divided into different unknown features based on their spectral information. In order to obtain the best result for extracting all three target land covers, a trial and error method was used to select right segmentation parameters. Finally, the segment scale level parameter was set as 30 and the merge level parameter was set as 97.7 (see results in Figure 3(a)). The results of segmentation are then exported to geographic vector layers with several spatial and spectral attributes produced by the software.

Pixel-based land cover classification performs an initial extraction of vegetation areas bare soils and parking lots using maximum likelihood classifier. This study proposed a hybrid pixel- and object-based land cover classification system to better extract urban land covers. The results of the two methods are summarized and compared in Table 1.

The classification results show that in the pixel-based approach, vegetation areas are extracted with a user accuracy of 92.34%, which is highest for all three land cover categories. Figure 3(b) shows the results of vegetation extraction including woods and grasses in the testing areas displayed in green. Visual interpretation of the results and the accuracy assessment for the en-

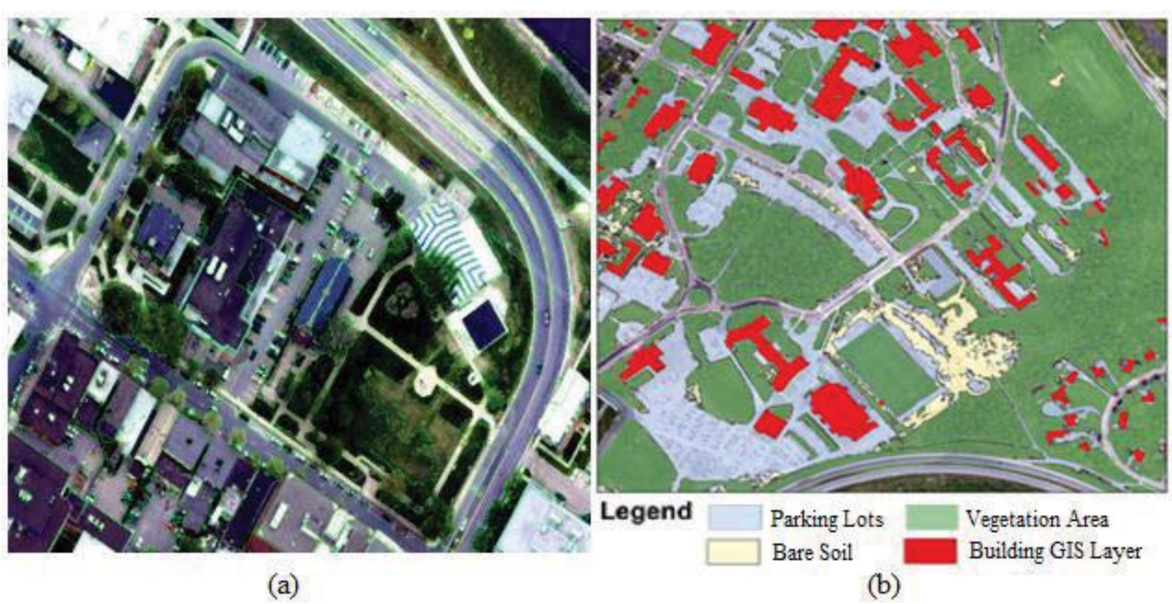

Figure 3. Land cover segmentation and classification results

tire testing area indicates that the user accuracy of vegetation extraction is improved to 97.21% through the proposed hybrid land cover classification method (see Table 1).

The pixel-based method results in a user accuracy of 91.75% for extracting bare soils. This accuracy is improved to 95.27% through the proposed hybrid method (see Table 1). The accuracy of parking lot extraction (with the user accuracy of 84.27%) using the pixel-based method is much lower than those for the other two land cover categories, which is most likely due to its heterogeneity and various spectral signatures. The proposed method shows that it mitigates the misclassifications arising from the above mentioned issues during the process of classifying the parking lots. The user accuracy of extracting parking lot is improved to 93.23% through the presented hybrid method (see Table 1).

Figure 3(b) illustrates the final results of parking lots, bare soil, and vegetation area for the testbed. Table 1 illustrated that the proposed hybrid land cover classification method results in consistent higher user accuracies for all three land cover categories than the pixel-based (Maximum Likelihood) method. Visual examination and accuracy assessment indicate that the extraction results are highly desirable for most planning exercises.

4.2 Land use extraction results

In order to develop a "best" binary logistic model, five steps from the stepwise regression are used to search the significant variables that should be included in the classification model. Considering the percentage correct and the significance value of each indicator (which should be less than the significance level specified (0.05)), it is found that the model using all the variables selected from the Step 5 offers the best overall accuracy of 97.8% (see Table 2). In this case, the least number of independent variables are included, which is consistent with the principle of "parsimony". Further, all of the following independent variables are significant: AvB_Height, Built_Ratio, FAR, P_Area, Dis_Local, Dis_Arterial, Dis_Collector, Dis_Freeway, and Park_Ratio and they are selected and used in the residential binary logistic LU classification model.

Then the probability of a parcel belonging to the residential class is calculated based on the binary logistic model identified from the Step 5 as the following:

$$Pr(Res) = 1 - \frac{Exp(-6.829+\sum_{i=1}^{14} \beta_i v_i)}{1+Exp(-6.829+\sum_{i=1}^{14} \beta_i v_i)} \qquad (4)$$

where β_i are the coefficients of independent variables estimated by the model and v_i are independent variables/indicators selected through the Step 5 of the above test, which are listed in Table 2.

The residential classification model is then tested with the testing dataset to check its prediction power. The accuracy assessments against the training and the testing datasets are then summarized in Table 3. The results show that the classification model developed can accurately classify parcels into residential and non-residential LUs. Residential parcels are classified with a high accuracy of 98.8% for the training dataset and 97.6% for the testing dataset. The accuracy of extracting non-Res parcels is slightly lower than the Res extraction, which results in 93.5% for the training and 91.7% for the testing of the model. Such results are expected as the non-Res parcel data is found to have more heterogeneity. The overall accuracy for training and testing the model is 97.8% and 96.6% respectively.

Visual interpretation to the residential LU classification results at the parcel level indicates that most of misclas-

Table 1. Accuracy assessment of land cover classification

Land Cover	Vegetation (m^2)		Parking Lot (m^2)		Bare Soil (m^2)		Accuracy (%)	
	1*	2**	1*	2**	1*	2**	1*	2**
Vegetation	75,950	79,961	5,851	2,120	450	170	92.34	97.21
Parking Lot	0	0	15,231	16,849	2,842	1,224	84.27	93.23
Bare Soil	0	0	672	385	7,470	7,757	91.75	95.27

* Maximum Likelihood Classifier, ** Hybrid Pixel- and Object-Based Classification Method

sified residential parcels are those for the medium and high density residential uses. This is because the signatures of medium and high density residential parcels are very similar to those for commercial or institutional purposes.

5. Conclusion and Discussion

Lack of elaborate land use information has forced city planners and modelers to use large LU zones in their models, and consequently accept undesirable approximations and errors in their analyses and planning workflow. Arbitrary/coarse zone-based system traditionally used in urban planning incurs the following two kinds of problems. First, intra-zone travel and spatial distributions of economic activities within zones are missing. Second, intra-zone variation has to be tolerated even economic activities are highly mixed in many cases. Extracting detailed subzonal LU information using remotely sensed imageries and geographic data is a choice to overcome the above issues. However, LUs, particularly in urban areas, exhibits unique natures, such as heterogeneity, complexity and abstract, which

Table 2. Statistical tests for residential LU classification using stepwise binary logistic model

	Step 1		Step 2		Step 3		Step 4		Step 5	
Variables	*β*	Sig.	*β*	Sig.	*β*	Sig.	*β*	Sig.	*β*	Sig.
Geographic Data										
sB_Area	-.002	.304	-	-	-	-	-	-	-	-
sB_Per	.014	.286	-	-	-	-	-	-	-	-
AvB_Com	-3.73	.234	-	-	-	-	-	-	-	-
AvB_Height	-.634	.083	-.67	.063	-.69	.042	-.61	.089	-.73	.029
Built_Ratio	-21.6	.055	-2.6	.053	-21.5	.025	-22.2	.03	-20.24	.025
FAR	12.31	.014	12.13	.016	13.49	.004	13.21	.006	14.05	.002
P_Area	.001	.000	.001	.000	.001	.000	.001	.000	.001	.000
Dis_CBD	.000	.69	.000	.447	-	-	-	-	-	-
Dis_Local	.007	.004	.007	.005	.007	.004	.134	.006	.007	.003
Dis_Arterial	-.003	.064	-.002	.104	-.003	.069	-	-	-.003	.036
Dis_Collector	.001	.028	.001	.032	.001	.038	.001	.037	.001	.031
Dis_Freeway	.001	.042	.001	.047	.001	.028	.001	.029	.001	.027
RS Data										
Veg_Ratio	-.042	.311	-.035	.387	-.023	.403	-.038	.134	-	-
Park_Ratio	.093	.022	.1	.013	.112	.000	.103	.000	.129	.000
BareS_Ratio	-.051	.395	-.03	.621	-	-	-	-	-	-
Constant	-1.519		-3.704		-5.01		-4.946		-6.829	
Percentage Correct										
Res	98.7		98.7		98.7		98.8		98.8	
Non-Res	93.1		93.1		93.1		92.4		93.5	
Overall	97.7		97.7		97.7		97.7		97.8	

Table 3. Accuracy assessments for residential LU classification at parcel level

Sample	Observed	Predicted		
		Res	non-Res	Percent Correct
Training	Res	504	6	98.8
	non-Res	7	101	93.5
	Overall Percent			97.8
Testing	Res	166	4	97.6
	non-Res	3	33	91.7
	Overall Percent			96.6
Overall Training and Testing				97.5

Dependent Variable: Residential Class

make LU classification a very challenging process. Previous studies show that extracting LU information is the most challenging part of the entire image classification process, and it requires auxiliary data and additional spatial analysis.

Extracting detailed subzonal LU information using remotely sensed imagery and geographic data has a good potential to solve the above issues. The morphological properties of different parcels derived from a set of geographic (DPM, Building footprint, and street network) and RS data function as potential indicators in the classification procedure. In order to obtain informative morphological properties from RS imageries, a hybrid image analysis method, which used a combined object- and pixel-based classification approach, is utilized to extract vegetation area, bare soil, and parking lot. The results of image processing demonstrate that it is more reliable and accurate, when compared to either single approach. The potential indicators for classifying different LUs are selected and used in land use classification through an initial investigation and statistical analysis.

In this study, a rule-based LU extraction system is developed to extract residential LU through the morphological analysis at the parcel-level. A Binary logistic model is regressed with the geographic and RS data to extract residential LUs, which results in an overall accuracy of 97.5%.

The LU extraction framework developed and reported in this paper shows a high accuracy in extracting detailed urban residential LUs at the fine spatial level̃ parcel. Future research is going to expand this to all types of LUs. At this fine spatial level, zones with mixed LUs can be further divided into subzones with single LU. Such tool would eventually solve the intra-zonal travel and mixed activity distribution issues frequently encountered in the traditional urban modeling process, which has been discussed above.

6. Acknowledgements

This study was primarily funded by a Natural Science and Engineering Research Council (NSERC) Strategic Project grant. The authors would like to thank NSERC, Canada for the financial aid and City of Fredericton for the data used in this study.

7. References

- Agresti, A. (1996) “An introduction to categorical data analysis”, New York: Wiley.

- Barnsley, M. J., Steel, A. M. and Barr, S. L. (2003) “Determining urban land use through an analysis of the spatial composition of buildings identified in lidar and multispectral image data”, In V. Mesev (Ed.), Remotely sensed cities (pp. 47–82), London: Taylor & Francis.

- Beykaei, S. A., Zhong, M. and Zhang, Y. (2010) “Evaluating the accuracy of extracting urban land cover/use from remotely sensed imagery and its potential application in urban planning”, 8th International Transportation Specialty Conference, CSCE, Winnipeg, Manitoba.

- Chan, J. C., Bellens, R., Canters, F. and Gautama, S. (2009) “An assessment of geometric activity features for per-pixel classification of urban man-made object using very high resolution satellite imagery”, Photogrammetric Engineering & Remote Sensing, 75(4), pp.397-411.

- Deng, J.S., Wang, K., Li, J., Feng, X.L., and Huang, J.Q. (2005) “Integration of SPOT-5 and ETM+ images to detect land cover change in urban environment”, Proceedings of Geoscience and Remote Sensing Symposium, IGARSS '05, IEEE International, Vol 4, pp.2284- 2286.

- Gomarasca, M. A., Brivio, P. A., Pagnoni, F. and Galli, A. (1993) “One century of land use changes in the metropolitan area of Milan (Italy)”, International Journal of Remote Sensing, Vol. 14, pp.211– 223.

- Hill, J. and Hostert, P. (1996) “Monitoring the growth of a Mediterranean Metropolis based on the analysis of spectral mixtures – A case study on Athens (Greece)”, in Progress in Environmental Remote Sensing Research and Applications, edited by Parlow (Rotterdam: Balkema), pp.21–30.

- Liu, W. and Prinet, Y. (2005). “Building detection from high-resolution satellite image using Probability Model.” Proceedings of Geoscience and Remote Sensing Symposium. IGARSS '05. 2005 IEEE International, Vol. 6, pp.3888- 3891.

- McGibbon, G. S. and Eyton, J. R. (1996) "Frequency-based contextual landform classification." The Canadian Surveyor, GEOMATICA, 50(3), pp. 287-299.

- Mesev, V. (2005) "Identification and characterization of urban building patterns using IKONOS imagery and point-based postal data." Comput., Environ.Urban Syst., 29(5), pp.541-557.

- Mohaptra, R. P. and Wu, C. (2008) "Subpixel imperviousness: Estimation with IKONOS imagery: An artificial neural network approach." Remote Sensing of Impervious Surfaces, pp.21-37.

- Mundia, C. N., and Aniya, M. (2005) "Analysis of land use/ cover changes and urban expansion of Nairobi City using remote sensing and GIS." International Journal of Remote Sensing, 26 (13), pp.2831-2849.

- Pacifici, F., Fabio, D. F., Solimini, D. and Burini, A. (2007) "Urban land cover classification: potential of high and very-high resolution SAR imagery." IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 1982-1985.

- Pesaresi, M. and Bianchin, A. (2001) "Recognizing Settlement Structure Using Mathematical Morphology and Image Texture." In: Donnay, J.-P., Barnsley, M. and Longley, P., Editors, 2001. Remote sensing and urban analysis, Taylor & Francis, London, 55–67.

- Xiao, Y., Lim, S. K., Tan, S.T., and Tay, S. C. (2004) "Feature extraction using very high resolution satellite imagery", IEEE International Geoscience and Remote Sensing Symposium, Anchorage, Alaska, pp.2004-2007.

- Yoshida, H., and Omae, M (2005). "An approach for analysis of urban morphology: Methods to derive morphological properties of city blocks by using an urban landscape model and their interpretations", Comput., Environ.Urban Syst., 29(2), pp. 223-247.

- Zhang, Y. (1999) "Optimization of building detection in satellite data by combining multispectral classification and texture filtering." ISPRS Journal of Photogrammetry & Remote Sensing, 54(1), pp.50-60.

- Zhang, Y. (2001) "Detection of urban housing development by fusing multisensor satellite data and performing spatial feature post-classification", International Journal of Remote Sensing, 22(17), pp.3339-3355.