Remote Sensing of Environment 115 (2011) 1145–1161

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery

Soe W. Myint [a,*], Patricia Gober [a], Anthony Brazel [a], Susanne Grossman-Clarke [b,d], Qihao Weng [c]

[a] School of Geographical Sciences and Urban Planning, Arizona State University, P.O. Box 875302, Tempe, AZ 85287-5302, United States
[b] Global Institute of Sustainability, Arizona State University, PO Box 875402, Tempe, AZ 85287, United States
[c] Department of Geography, Geology, and Anthropology, Indiana State University, Terre Haute, IN 47809, United States
[d] Potsdam-Institute for Climate Impact Research, P.O. Box 60 12 03, 14412 Potsdam, Germany

ARTICLE INFO

Article history:
Received 29 September 2009
Received in revised form 8 December 2010
Accepted 29 December 2010

Keywords:
Urban
High resolution
Object-based classifier
Membership function
Nearest neighbor

ABSTRACT

In using traditional digital classification algorithms, a researcher typically encounters serious issues in identifying urban land cover classes employing high resolution data. A normal approach is to use spectral information alone and ignore spatial information and a group of pixels that need to be considered together as an object. We used QuickBird image data over a central region in the city of Phoenix, Arizona to examine if an object-based classifier can accurately identify urban classes. To demonstrate if spectral information alone is practical in urban classification, we used spectra of the selected classes from randomly selected points to examine if they can be effectively discriminated. The overall accuracy based on spectral information alone reached only about 63.33%. We employed five different classification procedures with the object-based paradigm that separates spatially and spectrally similar pixels at different scales. The classifiers to assign land covers to segmented objects used in the study include membership functions and the nearest neighbor classifier. The object-based classifier achieved a high overall accuracy (90.40%), whereas the most commonly used decision rule, namely maximum likelihood classifier, produced a lower overall accuracy (67.60%). This study demonstrates that the object-based classifier is a significantly better approach than the classical per-pixel classifiers. Further, this study reviews application of different parameters for segmentation and classification, combined use of composite and original bands, selection of different scale levels, and choice of classifiers. Strengths and weaknesses of the object-based prototype are presented and we provide suggestions to avoid or minimize uncertainties and limitations associated with the approach.

1. Introduction

When extracting urban land cover information from remote sensor data, analysts tend to consider spatial resolution to be more important than spectral resolution. In other words, it is more useful to have finer spatial resolution (i.e., smaller pixel size) than higher spectral resolution (i.e., greater number of spectral bands or narrower interval of wavelengths). This is the major reason why aerial photography has traditionally been the key source for urban planning and management. There has recently been a shift from air photo to satellite based imagery for urban applications, because of the availability of a new generation of very high spatial resolution multispectral sensor data (e.g., QuickBird and IKONOS). The objective of launching and deploying the above commercial remote sensing satellite was to increase the visibility of terrestrial features, especially urban objects, by reducing per-pixel spectral heterogeneity and thereby improving land cover identification.

These finer resolution or larger scale image data exhibit higher levels of detailed features than those from preceding sensors (e.g., Landsat Thematic Mapper and SPOT), but this greater level of detail may lead to complicated urban features in the spectral domain. (Myint et al., 2006). This is because many small objects are concentrated in a small area when dealing with an urban space, and they become more and more visible as the spatial resolution gets finer and finer. This situation potentially leads to lower accuracy in urban image classification. This may not be the case for other environments, especially when dealing with other natural land covers and land uses (rangeland, evergreen forests, broad leaved forests, pine forests, mangroves, wetland, desert landscape, and agriculture).

Despite the above limitation, many urban spatial analysts and modelers must take approaches for urban decision-making that increasingly require urban land-use and land-cover maps generated from very high resolution data. For example, a remote sensing application to estimate population based on the number of dwellings of different housing types in an urban environment (single-family, multi-family), usually requires a pixel size ranging from about 0.25 to

* Corresponding author.
E-mail addresses: soe.myint@asu.edu (S.W. Myint), gober@asu.edu (P. Gober), abrazel@asu.edu (A. Brazel), sg.clarke@asu.edu (S. Grossman-Clarke), qweng@indstate.edu (Q. Weng).

0034-4257/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2010.12.017

5 m in order for there to be an identification of the type of individual structures (Jensen & Cowen, 1999). In general, any visible band or infrared spectral bands at this range of spatial scale should provide different spectral signatures between the object of interest (e.g., single family house) and its surrounding environment (e.g., roads, driveways, sidewalks, trees, shrubs, grass, and swimming pools). Most remote sensing scientists would agree that higher radiometric resolution or number of bits (e.g., 8 bit vs. 16 bit) would not noticeably increase information about small objects and surrounding features in high resolution data.

Cowen et al. (1995) report that there needs to be a minimum of four spatial units (e.g., 4 pixels) within an urban object area to effectively be identified using a remotely sensed image. In other words, the sensor spatial resolution needs to be at least one-half the diameter of the smallest object of interest. For example, if we need to identify a mobile home (an urban object) that is 4 m wide, the minimum spatial resolution of high quality imagery without haze or other atmospheric problems would be 2.0 by 2.0 m (Fig. 1). This implies that the required spatial resolution of remotely sensed data to prepare an urban land use and land cover map needs to be at least one half the size of the smallest object to be identified in the image. This operational concept or real world situation is not precisely in line with the theoretical definition of the spatial resolution as the smallest linear separation between objects that can be separated in an image (Jensen, 2005; Lillesand et al., 2008). In a real world situation, a 4 m wide object to be identified in an image most likely does not locate perfectly over 4 pixels in a 2 m resolution image as illustrated in Fig. 1. Hence, we need a pixel size that is remarkably smaller than an object to identify that object in an image. Since urban objects are notably smaller in comparison to natural features, it is apparent that we need a significantly small pixel size for urban applications.

The geometric elements of image interpretation (e.g., pattern, shape, size, and orientation) are important when using high-resolution image data for urban applications. However, the question of whether we should evaluate the usefulness of a given type of imagery (e.g., Landsat Thematic Mapper and IKONOS) for extracting specific types of information (e.g., swimming pools) based solely on its spatial characteristics alone is hard to answer. The accuracy of urban-suburban image interpretation from a panchromatic satellite imagery or an aerial photo may be improved by adding additional spectral bands of the same resolution. The above-mentioned new generation of high spatial resolution satellite data with multispectral capability has recently been available (IKONOS initiated in 1999, and QuickBird in 2001) on the market to provide detailed information (smaller objects) in their spatial as well as spectral domain.

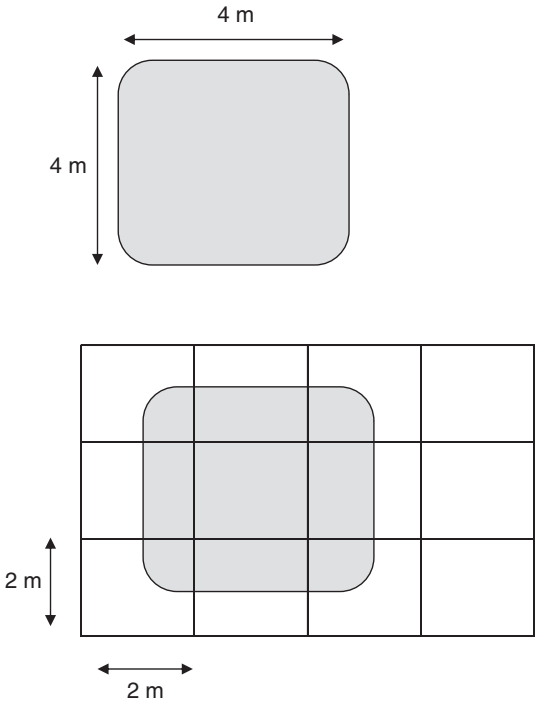

Fig. 1. A comparison of a 4-m wide object and 2- by 2-m resolution image data. There is only 1 pixel that contains a pure spectral response from the object.

As the spatial resolution gets smaller, the spectral response from these different small objects in an urban environment exhibits complex patterns in fine-resolution images. Even though human interpreters can naturally recognize such complex patterns (e.g., commercial) and individual land covers (e.g., individual houses, swimming pools, cement roads, and asphalt roads), traditional digital classification algorithms generally encounter serious problems in identifying urban classes in such scenes (Campbell, 2007), because they use spectral information (pixel values) alone as a basis to analyze and classify remote sensing images, and ignore spatial information and a group of pixels that need to be considered together as an object (Bentz et al., 2004; Walter, 2004). One of the major limitations in urban mapping is that many different urban land covers may share the same or similar spectral responses (e.g., cement roads, cement sidewalks, cement parking lots, cement rooftops, and other bright surface features). If our objective is to identify buildings and roads separately in a high resolution image, it could be anticipated that the classical algorithms (sometimes referred to as per-pixel classifiers — e.g., maximum likelihood, minimum distance to the mean, and Mahalanobis distance) would not be very effective to perform the job, since many rooftops and roads are made of same materials (e.g., cement and asphalt). In other words, some urban classes share the same or similar spectral responses that could lead to interpretive confusion when using traditional approaches. For example, asphalt roads and asphalt parking lots share the same reflectance as asphalt rooftops. A similar situation exists for many other land covers: cement roads and cement parking lots vs. cement rooftops, asphalt roads, asphalt parking lots, and asphalt rooftops vs. deep clear water, bright desert soil vs. bright manmade features, and grass vs. shrubs. Hence, accurately classifying urban land categories from high-resolution image data remains a challenge despite significant advances in geographic information science and technology. In this study, we employ an object-based classification approach that separates spatially and spectrally similar pixels at different scale levels as segmented objects which we expect can effectively identify urban land cover classes. We would like to emphasize that we consider contiguous pixels of similar properties that represent urban objects.

2. Background

The object-centered classification prototype generally starts with the generation of segmented objects at multiple level of scales as fundamental units for image analysis, instead of considering a per-pixel basis at a single scale for classification (Desclée et al., 2006; Im et al., 2008; Myint et al., 2008; Navulur, 2007; Stow et al., 2007). The fundamental concept of an object-based paradigm also differs from sub-pixel classifiers in that it does not consider spectra of different land covers that would quantify percent distribution of these land covers. The sub-pixel approach may not be appropriate for urban mapping with high resolution image data, since it is originally designed to identify percent distribution of different land covers in a coarse resolution imagery (e.g., Landsat TM and MODIS) (Asner & Heidebrecht, 2002). The sub-pixel processor is based on the concept that the spectral reflectance of the majority of the pixels in remotely sensed imagery is assumed to be a spatial average of spectral signatures from two or more surface categories or endmembers (Schowengerdt, 1995). As discussed earlier, high resolution remotely sensed data were designed to capture smaller objects with only a few

multiple bands (generally 3–4 bands), ranging from the visible to near infrared portion of the electromagnetic spectrum. The sub-pixel tool is intended to quantify materials that are smaller than image spatial resolution (Weng & Hu, 2008). On one hand, it may not be necessary to identify percent distribution of land covers in a small pixel (e.g., QuickBird multispectral at 2.4 m spatial resolution, QuickBird panchromatic at 60 cm spatial resolution). On the other hand, it may not be appropriate to model spectral responses from ground features in fewer bands (both IKONOS and QuickBird contain 4 bands) to effectively quantify percent distribution of many different land cover classes, since sub-pixel approaches use spectra of all possible land covers in all available bands. For this reason, hyperspectral remote sensing has been used effectively to quantify fractions of image endmembers at a sub-pixel level (Okin et al., 2001; Roberts et al., 1998; Roberts et al., 2003). Moreover, the linear spectral unmixing classifier, the most widely used sub-pixel approach, does not permit number of representative materials or endmembers greater than the number of spectral bands (Lu & Weng, 2004).

There have been several geospatial techniques that have emerged as an alternative to spectral-based traditional classifiers to improve the classification accuracy: the image spatial co-occurrence matrix (Franklin et al., 2000); local variance (Ferro & Warner, 2002); the variogram (De Jong & Burrough, 1995); fractal analysis (Lam & Quattrochi, 1992), Getis index (Myint et al., 2007), spatial autocorrelation (Purkis et al., 2006), lacunarity (Myint & Lam, 2005), and wavelet transforms (Myint, 2006). These approaches have demonstrated significant improvements in the classification accuracy of land-covers and subsequent inference of urban land-use from land-cover classes. The above approaches compute geospatial indices within a local moving window that represent spatial arrangements and patterns of urban objects (land covers) to identify land use classes. Most of these approaches are designed to identify the composite (e.g., residential) of many features rather than in making an inventory of many small objects (swimming pool, tree, shrubs, driveway, and sidewalk) that may be of little or no interest (Campbell, 2007). Hence, geospatial approaches using different window sizes to characterize land use that can be inferred from detailed land covers is not part of the study as well.

3. Data and study area

A QuickBird image data over a central region in the city of Phoenix acquired on 29 May 2007 is used. The study area is about 178 km^2 covering a little more than 51 census tracts (upper left longitude 112° 7′ 45″ and latitude 33° 33′ 15″, lower right longitude 112° 0′ 50″ and latitude 33° 26′ 2″) (Fig. 2). The dataset has 2.4 m spatial resolution with 4 channels: blue — B1 (0.45–0.52 μm), green — B2 (0.52–0.60 μm), red — B3 (0.63–0.69 μm), and near infrared — B4 (0.76–0.90 μm). The radiometric resolution of the dataset is 16 bit. Even though the study area is only part of Phoenix, the image data is sizeable (5339 rows × 5570 columns) due to its fine spatial resolution. The area includes urban segments (commercial, industrial, and residential) and undeveloped regions (grassland, unmanaged soil, desert landscape, and open water), giving a diversity of urban land use and land cover classes. The selected land-cover classes that we identified for the study include buildings, other impervious surfaces (e.g., roads and parking lots), unmanaged soil, trees/shrubs, grass, swimming pools, and lakes/ponds. These particular land-cover classes are important to ongoing analysis of the urban energy budget using a model that requires these land cover classes (e.g., Gober et al., 2010, Grimmond & Oke, 2002). In addition to the original bands, principal component analysis (PCA) bands stretched to 16 bit and normalized difference vegetation index (NDVI) were used in the analysis.

In addition to the above QuickBird image, we selected another QuickBird image acquired on 11 July 2005 over Tempe, Arizona (hereafter referred to as test image) to evaluate the effectiveness of the multiple classification strategies employed for urban object extraction. The selected test image is a relatively small image where spatial arrangements of objects and type of land covers are significantly different from most urban features in the main image. We carefully and intentionally selected this image to examine if the same approach was consistently effective in identifying similar classes in urban areas with different environmental settings. The subset has 541 columns and 851 rows (upper left longitude 111° 56′ 24″ and latitude 33° 25′ 58″, lower right longitude 111° 55′ 33″ and latitude 33° 24′ 52″).

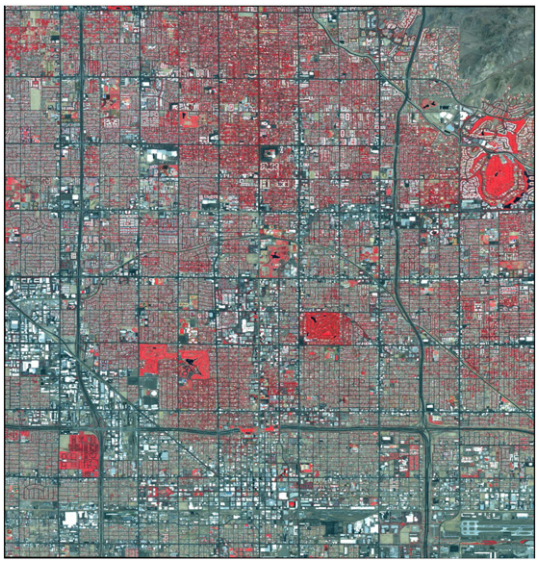

Fig. 2. Study area located in Phoenix, Arizona.

4. Methods

4.1. Spectral information of urban land-cover classes

To demonstrate if spectral information based on a single pixel alone is effective in urban classification, we used spectra of the selected classes from thirty randomly selected points (brightness values of the selected classes in all four bands) to examine if they can be accurately discriminated. The target materials of the classes that we selected were the same as the original land-cover categories used in the object-based classification: buildings, other impervious surfaces (e.g., roads and parking lots), unmanaged soil, trees/shrubs, grass, swimming pools, and lakes/ponds. Fig. 3 shows image brightness values of the target materials in the study area. We observed some shadows around high-rise buildings in the central business district in the city of Phoenix. However, we do not consider this a serious issue since there were not many high-rise buildings in the study area. Moreover, shadows around these high-rise buildings were about two to 3 pixels away from where the actual buildings were located. We also noticed that there were no shadows around some high-rise buildings depending on where they are located. This could have been due to the fact that the sun angle of the time the image was acquired was near nadir. Hence, we manually edited shadows in the study area instead of adding shadows as a class. To evaluate the classical per-pixel classification technique, a linear discriminant analysis approach was used. The spectra of the target materials at randomly selected points were subjected to discriminant analysis. The procedure generates a discriminant function (or, for more than two groups, a set of discriminant functions) based on linear combinations of the predictor variables, which provide the best discrimination between the groups. It can be anticipated from the statistics (Table 1) of the selected classes that the data range and standard deviations were high

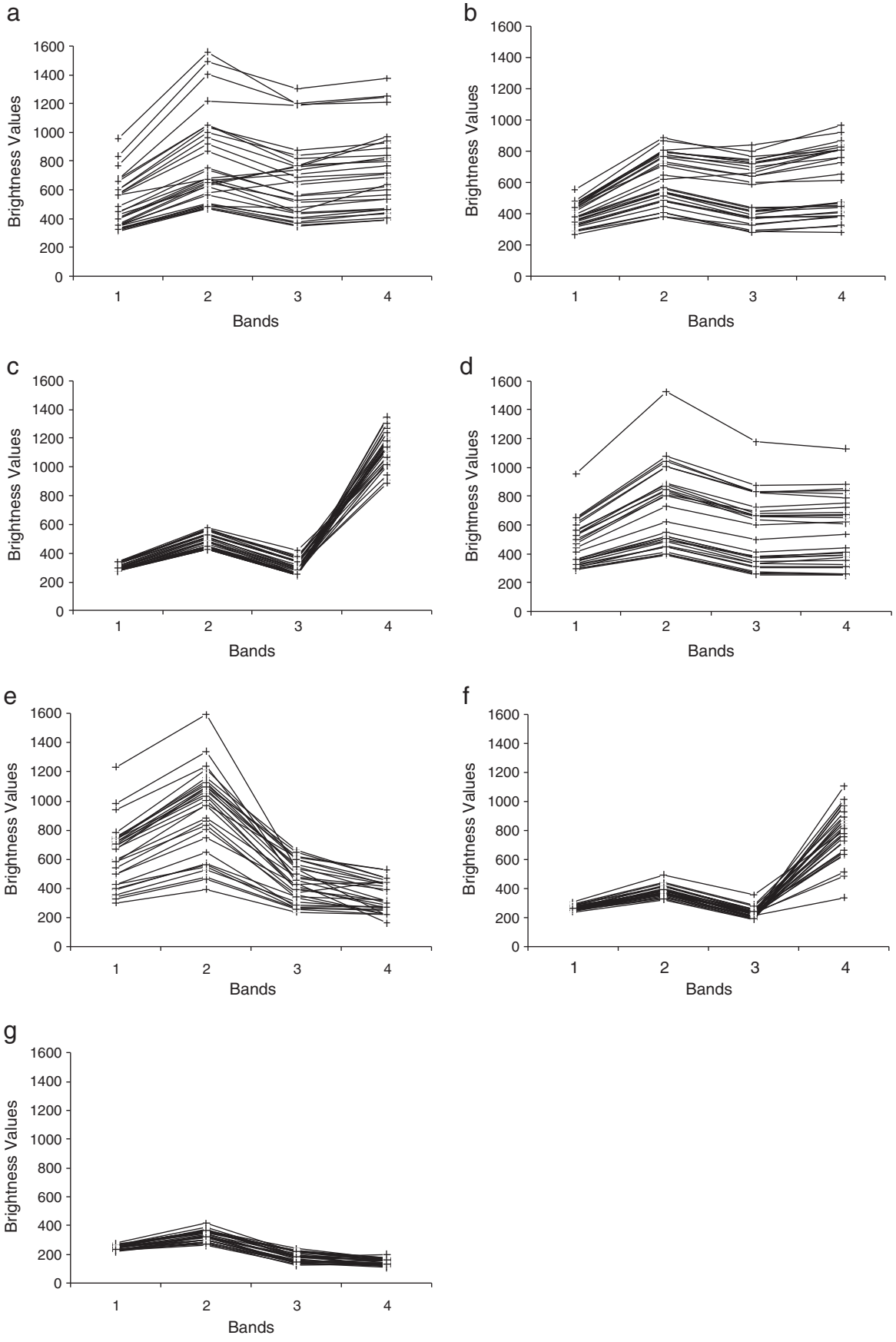

Fig. 3. Image brightness values from urban land cover classes: (a) buildings, (b) unmanaged soil, (c) grass, (d) other impervious surfaces, (e) swimming pools, (f) trees/shrubs, and (g) lakes/ponds.

Table 1
Statistics of the selected land-cover classes.

	Band 1	Band 2	Band 3	Band 4
Buildings				
Minimum	314	469	345	388
Maximum	957	1559	1306	1376
Mean	489.66	798.59	678.72	733.45
Standard deviation	173.92	317.25	274.16	282.67
Unmanaged soil				
Minimum	268	386	286	286
Maximum	559	889	840	966
Mean	392.77	626.53	538.93	602.73
Standard deviation	71.28	158.62	181.34	214.29
Grass				
Minimum	274	425	245	889
Maximum	340	576	419	1350
Mean	301.77	492.10	302.57	1132.70
Standard deviation	21.07	49.82	48.55	117.28
Other impervious				
Minimum	288	393	253	250
Maximum	954	1524	1176	1130
Mean	460.84	710.77	544.94	557.71
Standard deviation	151.58	272.09	235.69	230.15
Pools				
Minimum	301	393	236	161
Maximum	1234	1593	662	530
Mean	627.50	920.40	439.27	339.27
Standard deviation	212.16	294.49	131.50	105.61
Trees/shrubs				
Minimum	235	324	184	338
Maximum	307	494	357	1105
Mean	263.00	382.70	230.10	792.23
Standard deviation	16.64	36.06	35.59	170.77
Lakes and ponds				
Minimum	218	262	124	108
Maximum	279	415	239	202
Mean	247.50	332.07	180.60	147.00
Standard deviation	16.23	37.39	33.34	22.48

for some classes (e.g., buildings, other impervious surfaces, and swimming pools). Consequently, there are some spectra overlaps among the selected classes. Table 2 shows producer's accuracy, user's accuracy, overall accuracy, and the kappa coefficient. The overall accuracy and kappa coefficient reached only about 63.33% and 0.59 respectively. It can be anticipated that using the spectra information alone, as in the classical approaches, may not be effective for urban land cover classification.

4.2. Object-based analysis

4.2.1. Image segmentation

Image segmentation is a principal function that splits an image into separated regions or objects depending on parameters specified (Im et al., 2008; Lee & Warner, 2006; Myint et al., 2008; Stow et al., 2008). A group of pixels having similar spectral and spatial properties is considered an object in the object-based classification prototype. We used Definiens Developer 7.0 (formerly known as eCognition software — Definiens, 2008) to perform an object-based classification approach (Baatz & Schape, 1999). A number of studies have demonstrated methods of assessing segmentation accuracy (Lucieer, 2004) and comparing sample segment objects against corresponding reference objects (Weidner, 2008; Winter, 2000). Möller et al. (2007) developed an approach to identify a segmentation scale that is close to optimal using trial-and-error tests in combination with an index called the "comparison index". They generated objects at different scales using a segmentation procedure called fractal net evolution approach (Baatz & Schape, 2000) that is implemented in the eCognition software. Munoz et al. (2003) pointed out advantages and disadvantages of various segmentation approaches that integrate region and boundary information, and reported that there is no perfect segmentation algorithm, which is crucial for the advancement of computer vision and its applications. Mueller et al. (2004) employed an object-based segmentation with special focus on shape information to extract large man-made objects, mainly agriculture fields in high resolution panchromatic data. It is important to note that there is no standardized or widely accepted method available to determine the optimal scale for all types of applications, areas with different environmental and biophysical conditions, and different kinds of remotely sensed images.

We used a segmentation algorithm available in Definiens known as the multiresolution segmentation which is based on the Fractal Net Evolution Approach (FNEA) (Baatz & Schape, 2000). The first step in the object-based paradigm with Definiens software is that we need to assign appropriate values to three key parameters, namely shape (S_{sh}), compactness (S_{cm}), and scale (S_{sc}) to segment objects or pixels having similar spectral and spatial signatures in an image. Users can apply weights ranging from 0 to 1 for the shape and compactness factors to determine objects at different level of scales. These two parameters control the homogeneity of objects. The shape factor adjusts spectral homogeneity vs. shape of objects, whereas the compactness factor, balancing compactness and smoothness, determines the object shape between smooth boundaries and compact edges. There is also a parameter called "smoothness" that is directly linked to compactness as the sum of smoothness and compactness is equal to one. The compactness or smoothness is effective only when the shape factor is larger than zero. The scale parameter that controls the object size that matches the user's required level of detail can be

Table 2
Overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient produced by the discriminant analysis using brightness values of the urban land cover classes.

Classified	Reference								Producer's accuracy (%)	User's accuracy (%)
	Buildings	Unmanaged soil	Grass	Other impervious	Pools	Trees/shrubs	Lakes/ponds	Total		
Buildings	15	11	4	0	0	0	0	30	40.54	50.00
Unmanaged soil	10	11	2	0	0	0	7	30	34.38	36.67
Grass	12	10	7	0	0	0	1	30	50.00	23.33
Other impervious	0	0	0	26	4	0	0	30	78.79	86.67
Pools	0	0	1	7	21	0	1	30	84.00	70.00
Trees/shrubs	0	0	0	0	0	23	7	30	100.00	76.67
Lakes/ponds	0	0	0	0	0	0	30	30	65.22	100.00
Total	37	32	14	33	25	23	46	210		

Overall accuracy = 63.33%.
Overall kappa statistics = 0.59.

considered the most crucial parameter of image segmentation. Different levels of object sizes can be determined by applying different numbers in the scale function. The higher number of scale (e.g., 100) generates larger homogeneous objects (similar to a smaller cartographic or mapping scale), whereas the smaller number of scale (e.g., 10) will lead to smaller objects (larger scale). We would like to emphasize that this is a spatially aggregated scale (more similar pixels or bigger objects vs. less similar pixels or smaller objects). A larger number used in the scale parameter is considered lower level in the segmentation procedure. The decision on the level of scale depends on the size of object required to achieve the goal. The software also allows users to assign different level of weights to different bands in the selected image during image segmentation.

The shape parameter (S_{sh}) was set to 0.1 to give less weight on shape and give more attention on spectrally more homogeneous pixels for image segmentation. The compactness parameter (S_{cm}) and smoothness were set to 0.5 to balance compactness and smoothness of objects equally. After testing different scale levels and parameter values and evaluating qualitatively, we considered scale levels from 10 to 100 to be appropriate for the study. To evaluate if these scale levels are appropriate for the classification, we selected 10 different objects per class at scales 10, 25, 50 and 100 to perform a discriminant analysis. We visually checked segmented objects at the selected scale levels and selected significantly different objects of the same classes. The target objects of the classes that we selected were the same as the original land-cover categories. We used mean values of objects in each spectral band to generate a discriminant function (or, for more than two groups, a set of discriminant functions) based on linear combinations of the predictor variables to evaluate if they can be separated effectively. Table 3 shows producer's accuracy, user's accuracy, overall accuracy, and the kappa coefficient generated at the selected scale levels. As expected the lowest scale level or scale level 1 (i.e., scale 10) produced the highest accuracy (100%) whereas the highest scale level or scale level 4 produced the lowest accuracy (75.71%), since a higher scale level generates larger objects that can potentially miss smaller objects of land covers (e.g., residential buildings and swimming pools). The overall accuracies produced by the discriminant analysis at scale 10, 25, 50, and 100 were 100%, 97.14%, 97.14%, and 75.71% respectively. Producer's accuracies and user's accuracies were 100% for all classes. This demonstrates that small objects within each category can be identified accurately using mean values of the original bands. However, this does not necessarily mean that this is the exact level of scale that generates optimal object sizes for all classes. Buildings, grass, and trees/shrubs produced the highest producer's and user's accuracies (100%) at levels 2 and 3. It was found that the best producer's and user's accuracies at the highest scale level (producer's accuracy = 100%, user's accuracy = 90%) were produced by other impervious surfaces. Even though this analysis provides classification accuracies at different levels, it may not explicitly imply classification accuracies for a particular class, since some land cover classes contain different size of objects. For example, area coverages between residential buildings and commercial buildings are significantly different. Results from this analysis can be taken as a guide to make decisions on object extraction for different land covers at different scale levels. However, to determine an optimal scale level to effectively extract particular objects it is necessary to qualitatively analyze them on the display screen and determine accuracies of different land cover classes as a trial and error approach. This is because the selection depends on several factors such as the nature of the study area, classification specificity, type of land cover classes, spectral and spatial properties of objects within classes, and variation of object sizes within classes. Results from the discriminant analysis suggest that all 4 levels tested in the segmentation are suitable for different objects in different classes. Hence, we employed 4 different scale levels (S_{sc}) to segment objects: 10, 25, 50, and 100 in the study. Fig. 4 shows segmented images of a subset at scale_level 1, scale_level 2, scale_level 3, and scale_level 4 (scale parameters 10, 25, 50, and 100) using shape (S_{sh}) 0.1 and compactness (S_{cm}) 0.5 (smoothness of 0.5).

4.2.2. Classification methods

There are two options to assign classes to segmented objects: membership function and the nearest neighbor classifier.

4.2.2.1. Membership function classifier. By using the user's expert knowledge, we can define rules and constraints in the membership function to control the classification procedure. The membership function describes intervals of feature characteristics that determine whether the objects belong to a particular class or not. The membership function is a non-parametric rule and is therefore independent of the assumption that data values follow a normal distribution.

4.2.2.2. Nearest neighbor classifier. To classify image objects using the nearest neighbor classifier, we need to define the feature space (e.g., original bands, transformed bands, and indices), define training samples (objects), classify, review the outputs, and optimize the classification. The nearest neighbor classification procedure uses a set of samples that represent different classes in order to assign class values to segmented objects. The procedure therefore consists of two steps: (a) teach the system by giving it certain image objects as samples, and (b) classify image objects due to their nearest sample neighbors in their feature spaces (Definiens, 2008). The nearest neighbor option is also a non-parametric rule and is therefore independent of a normal distribution. The nearest neighbor approach allows unlimited applicability of the classification system to other areas, requiring only the additional selection or modification of new objects (training samples) until a satisfactory result is obtained (Ivits & Koch, 2002). Application of the nearest neighbor method is also advantageous when using spectrally similar classes that are not well separated using a few features or just one feature (Definiens, 2008). The nearest neighbor approach in eCognition can be applied to any classes at a scale level using original, composite, transformed, or

Table 3
Overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient produced at four different scale levels (i.e., 10, 25, 50, and 100) by the discriminant analysis.

Land cover	Scale 10		Scale 25		Scale 50		Scale 100	
	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy
Buildings	100%	100%	100%	100%	100%	100%	85.71%	60%
Unmanaged soil	100%	100%	90.91%	100%	90.91%	100%	70%	70%
Grass	100%	100%	100%	100%	100%	100%	43.75%	70%
Other impervious	100%	100%	100%	90%	100%	90.00%	100%	90%
Pools	100%	100%	90%	90%	90.91%	100%	71.43%	100%
Trees/shrubs	100%	100%	100%	100%	100%	100%	100%	60%
Lakes/ponds	100%	100%	100%	100%	100%	90.00%	100%	80%
Overall accuracy	100		97.14		97.14		75.71	
Kappa	1		0.9674		0.9674		0.7233	

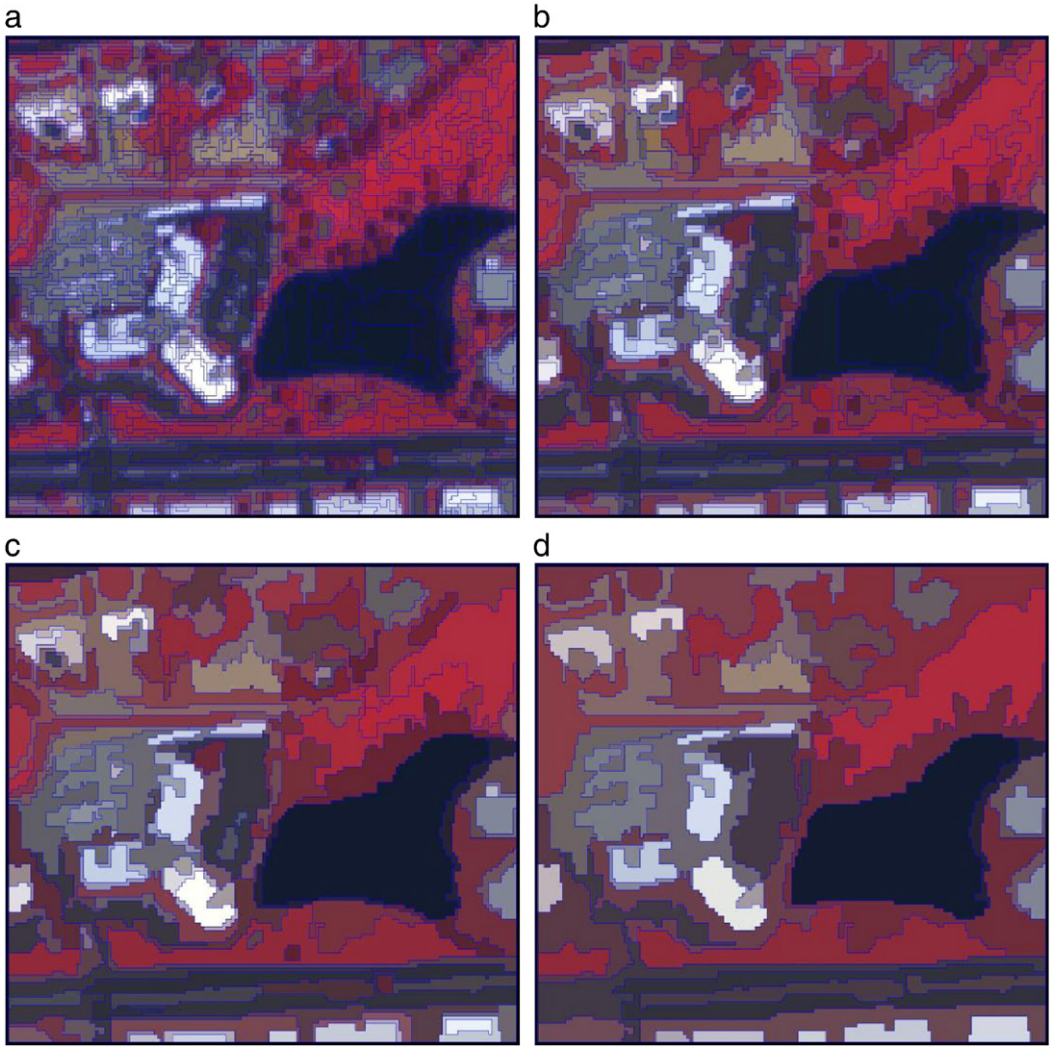

Fig. 4. Image segmentation at four scale levels: (a) scale parameter 10; (b) scale parameter 25; (c) scale parameter 50; and (d) scale parameter 100.

customized bands. There are two options available with the nearest neighbor function in Definiens or eCognition software, namely (1) Standard Nearest Neighbor, and (2) Nearest Neighbor. The Standard Nearest Neighbor option automatically selects mean values of objects for all the original bands in the selected image whereas the second option requires users to identify variables (e.g., shape, texture, and hierarchy) under object features, class-related features, or global features. The steps employed in this study to map different urban land-use and land-cover classes at each scale level are described below. We classified different group of classes or individual classes separately using different sets of parameters, different feature space (different bands, indices, or composite bands), different level of scales, and different classification rules. After completion of identifying classes separately, we combined them using a GIS overlay function.

4.2.3. Classification of grass, trees, and others

We employed the nearest neighbor classifier to identify grass, trees/shrubs, and other classes (e.g., impervious surfaces, swimming pools, and other water features). The objective of this component was to effectively discriminate grass from trees. After visually examining all scale levels, we decided to use scale level_1 (scale parameter = 10). The selected bands for the feature space include mean of the original bands 2, 3, 4, ratio of PCA band 1, and NDVI image. The ratio band in Definiens or eCognition is defined as a particular band divided by the summation of all other bands and contains digital numbers (DN) between 0 and 1. We identified 4 to 5 training samples (training objects) for each choice of grass and trees, and several samples of other classes. Fig. 5 shows a flow chart to identify grass, trees, and others.

4.2.4. Classification of buildings

We employed our expert knowledge in the membership function classifier to identify buildings in the study area. We found out that scale_level 2 (scale parameter = 25) was the optimal scale for the classification. After testing many different band combinations and composite bands with various parameters in the membership function, the selected bands for the feature space include the mean of the original band 1, ratio of PCA band 3, and NDVI image. We found that digital values of the ratio of the PCA band 3 between 0.45 and 0.58 which intersect with a DN value of the mean of the original band 1 higher than 390, determines buildings and vegetated areas, especially grassy features. To exclude vegetated areas, we again intersect the above output with pixels having NDVI values less than 0.1. The following is the expert system rule that was employed to extract buildings in the study area.

(1) Ratio PCA band 3 DN values < 0.45 (and)
(2) Ratio PCA band 3 DN values > 0.58 (and)
(3) Mean band 1 DN values > 390 (and)
(4) NDVI values < 0.1

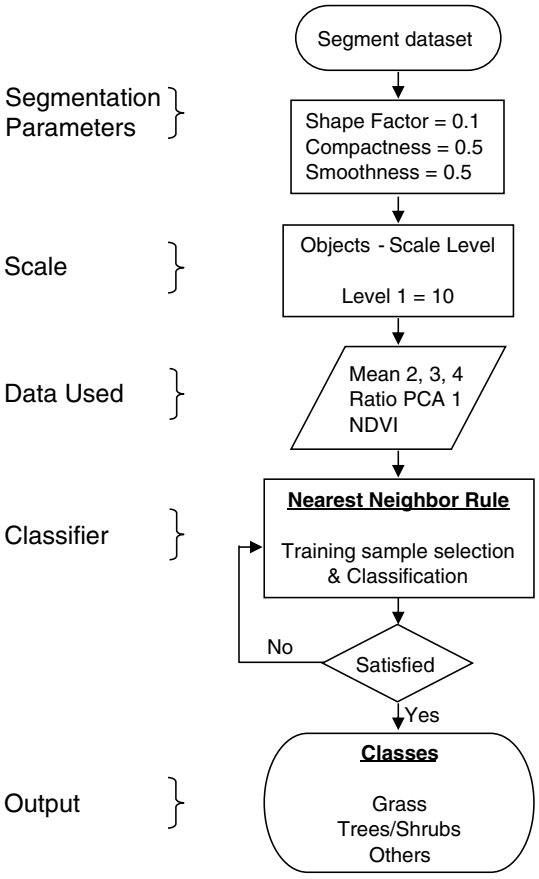

Fig. 5. Procedure to extract grass and trees/shrubs.

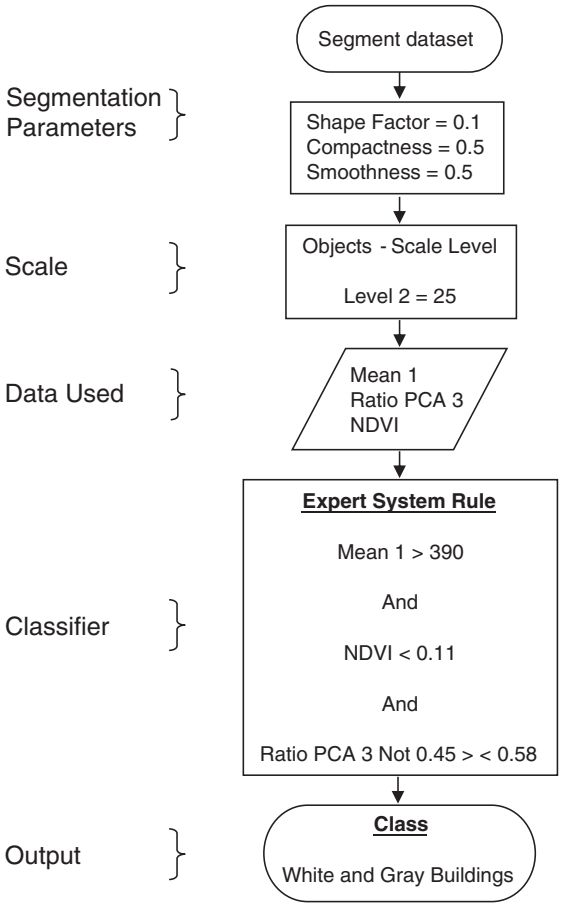

Fig. 6. A flow chart demonstrating how to extract buildings.

Fig. 6 demonstrates a flow chart to extract buildings in the study area.

4.2.5. Classification of other impervious surfaces

As stated earlier, the category of other impervious surfaces includes roads, driveways, sidewalks, and parking lots. The expert system rule that was employed for the extraction of roads is the reverse of the approach employed above for identifying buildings. Hence, the selected bands for the feature space include mean of the original band 1, ratio of PCA band 3, and NDVI image. We also used the same scale-level that we used for the building (scale parameter = 25). The expert system rule employed can be described as

(1) Ratio PCA band 3 DN values between 0.45 and 0.58 (and)
(2) Mean band 1 DN values < 390 (and)
(3) NDVI values > 0.1

Since the output is the opposite of buildings, we do not provide the impervious area map of the same subset above. Fig. 7 presents a diagram that explains how to extract other impervious surfaces in the study area.

4.2.6. Classification of swimming pools

We used an expert system rule at the scale level of 1 (scale parameter = 10) to extract swimming pools in the study area. The selected bands for the feature space include a mean of PCA band 2 and mean of PCA band 3. We found that DN values of the mean of PCA band 2 that are lower than 15,000 and DN values of mean of PCA band 3 that are lower than 24,000 can effectively identify swimming pools. We decided to use both options to make sure there is no missing area over pools with slightly different spectral response. The expert system rule employed to determine swimming pools can be described as

(1) Mean PCA band 2 DN values < 15,000 (or)
(2) Mean PCA band 3 DN values < 24,000

Fig. 8 demonstrates a classification procedure to extract the swimming pool class in the study area.

4.2.7. Classification of lakes and ponds

This class covers all natural and man-made water bodies (e.g., ponds and lakes) except swimming pools. For this class, we considered the scale level of 3 (scale parameter = 50) to be the optimal scale to perform the classification. We used an expert system rule using the mean of PCA band 1 and ratio of PCA band 3. It was found that both bands were equally effective. To make sure the classification as effectively as possible we used both bands to identify the class. The following describes the expert system rule that was used to capture lakes and ponds.

(1) Mean PCA band 1 DN values < 4000 (or)
(2) Ratio PCA band 3 DN values > 0.62

Fig. 9 demonstrates a flow chart to extract lakes and ponds in the study area.

4.2.8. Classification of unmanaged soil and other land covers

This component was to identify unmanaged soil effectively in the study area. We observed that there was some significant signature confusion between unmanaged soil and some building rooftops. For example, unmanaged soil in false color composite displaying near infrared, red visible, and green visible bands in red, green, and blue

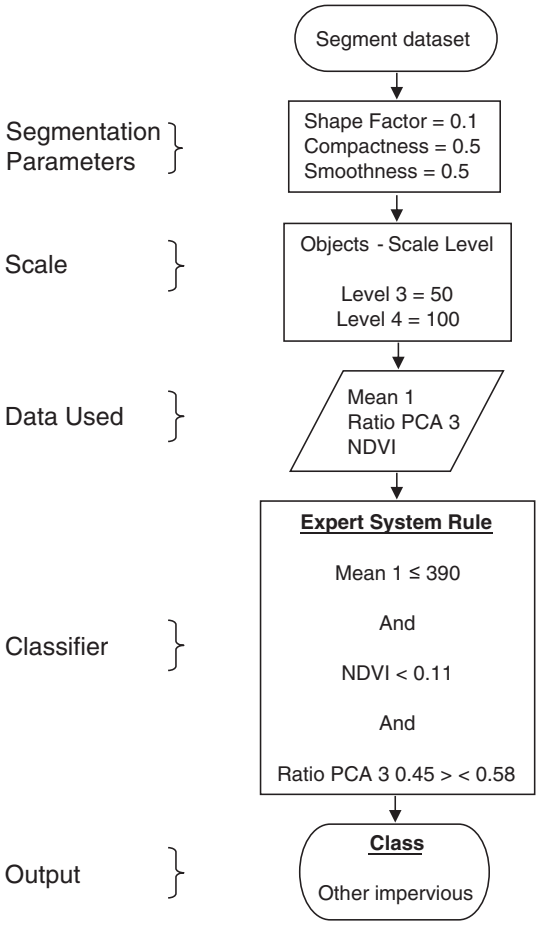

Fig. 7. A flow chart demonstrating how to extract other impervious surfaces.

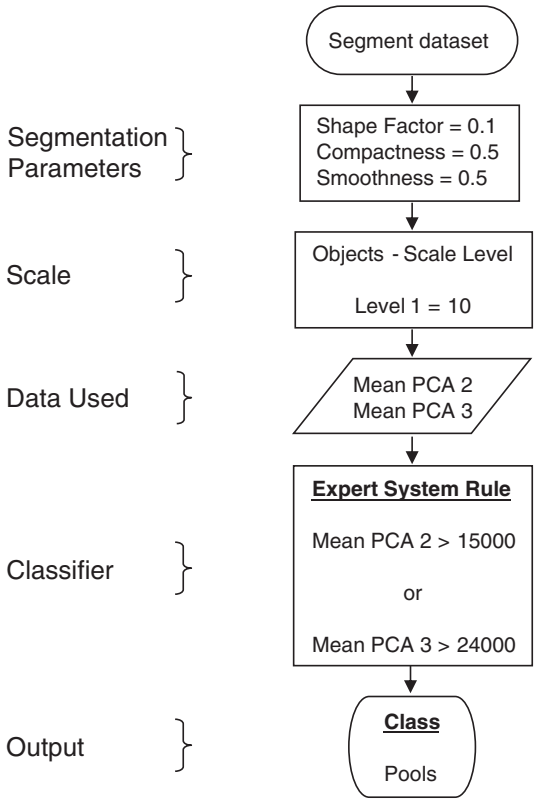

Fig. 8. A flow chart demonstrating how to extract swimming pools.

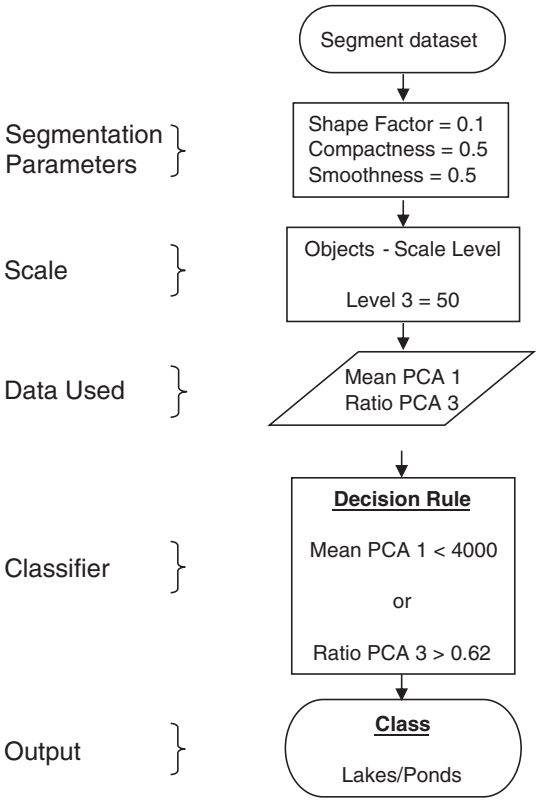

Fig. 9. A flow chart demonstrating how to extract lakes/ponds.

appeared as yellow, orange, or brown colors. There are some rooftops that appeared as yellow, orange, or brown colors in the same false color display. This type of signature confusion was also evident in the spectra statistics (Table 1) and the discriminant analysis (Table 2). We attempted many different classification options with the nearest neighbor analysis as well as the membership function, and the nearest neighbor classifier was a better option to effectively discriminate these two classes. We attempted many different training samples per class as a trial and error approach. The selected bands for the feature space include the mean of the original band 2, mean of the original band 3, and mean of the original band 4. Training samples selected include unmanaged soil similar to buildings, yellow or brown buildings, white buildings, gray buildings, asphalt roads/parking lots, lakes/ponds, grass, and trees/shrubs. We did not attempt to identify swimming pools as a separate class, since there is some signature confusion between swimming pools and buildings. On the other hand, we were satisfied with the output map of swimming pools generated by the expert system rule that we employed. Even though the objective was to identify unmanaged soil accurately, this classification output was also treated as a base map to avoid any missing pixels or unclassified pixels after integrating the previously identified individual classes. Fig. 10 demonstrates a classification procedure to extract the unmanaged soil and other land-cover classes in the study area.

4.2.9. GIS overlay function to produce final output

We overlaid individual layers at different levels to produce a final output map of urban land-cover classes. The first GIS overlay function started with the last two layers. This was mainly to add lakes/ponds generated by the membership function to the last output map. Hence, a priority is given to water bodies identified with the membership function when intersecting the water output with the last output generated by the nearest neighbor classifier. In other words, all water pixels identified in the first layer that intersect with any other classes

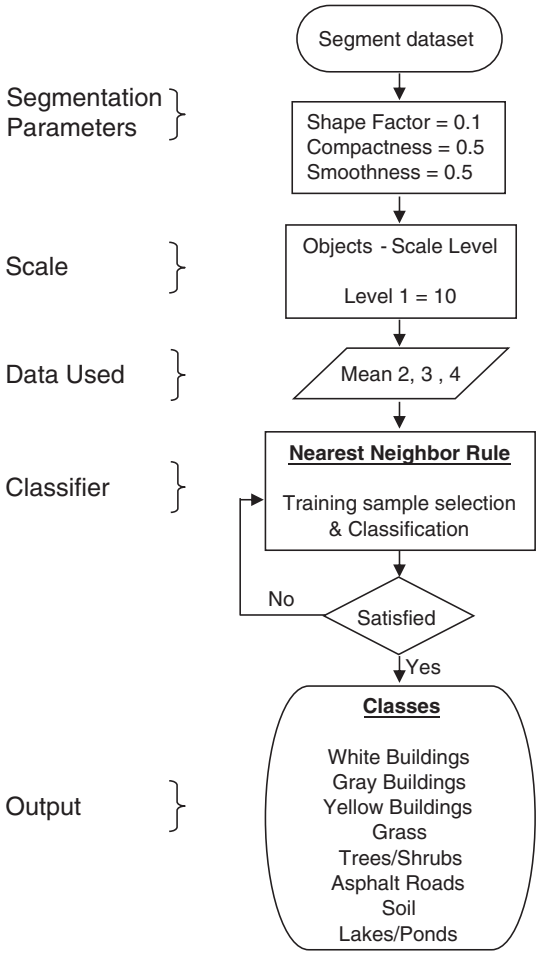

Fig. 10. A flow chart demonstrating how to extract unmanaged soil and other land-cover classes.

in the second layer were identified as lakes/ponds. This output map was overlaid with the map produced for the impervious surface to add that class. We later obtained buildings, trees, and grass from the first two layers generated at scale levels 2 and 1. The swimming pool output map was overlaid last to minimize a possible signature confusion with swimming pools and some building rooftops. We later merged different types or colors of building rooftops as building class. A flow chart that demonstrates a step-by-step procedure to conduct this research study is presented in Fig. 11. A subset of the original QuickBird image and the output map of the object-based approach is shown in Figs. 12 and 13 respectively.

4.3. Traditional classifier

We also employed the most commonly used supervised decision rule, namely maximum likelihood classifier to thoroughly evaluate the effectiveness of the object-based approach. The maximum likelihood decision rule is based on the probability that a pixel belongs to a particular class. The basic equation of the decision rule assumes that these probabilities are equal for all classes (Jensen, 2005; Lillesand et al., 2008). Traditional per-pixel classifiers use a combined spectral response from all training set pixels for a given class. Hence, the resulting signature comprises responses from a group of different land covers in the training samples, and the classification system simply ignores the impact of mixed pixels (Lu & Weng, 2004).

The maximum likelihood classifier is a type of parametric decision rule that is based on the assumption that data values follow a normal distribution, and that the statistical parameters (e.g. mean, variance, covariance matrix) of the training samples are representative. However, the assumption of a normal spectral distribution could potentially lead to some errors if the data is not normally distributed. We selected 5 to 7 training samples per class that are spectrally different for the classification. We also attempted several different sets of training samples and qualitatively evaluated the outputs. We merged those classes generated by different training samples under the same land-use and land-cover category. The resulting land-use and land-cover categories were the same as those identified with the object-based approach. The output map produced by the traditional classifier is presented in Fig. 14.

4.4. Additional classification with test image

As mentioned earlier we selected another QuickBird image that covers part of Tempe, Arizona where nature of the study area, especially spatial arrangements of objects and type of land covers, are significantly different from most urban features in the main image to treat as a test image. We employed the same multiple classification strategies developed for urban object extraction in the main image. This is to demonstrate whether the same methodology was consistently effective in identifying the selected classes in urban areas with different environmental settings. To better evaluate the approach and for comparison purposes, we also used the maximum likelihood classifier to identify the selected classes. The original test image and two output maps generated by the object-based and maximum likelihood approaches are presented in Fig. 15a,b, and c respectively.

5. Accuracy assessment

5.1. Pixel-based accuracy assessment

An adequate accuracy assessment of a remotely sensed image is to compare the land-use land-cover classification at every pixel in an image with a reference source or a ground truth information. While this approach is ideal, gathering reference data for an entire study area is expensive (i.e., costly, labor intensive, and time consuming) and ruins the main purpose of performing a remote-sensing classification (Lillesand et al., 2008). Selection of a certain number of sample pixels that are assumed to represent the whole image has been used to avoid the above issue (Campbell, 2007; Jensen, 2005). Since accuracy assessment assumes that the sample points selected are the true representation of the map being evaluated, an improperly gathered sample will produce meaningless information on the map accuracy (Congalton & Green, 1999; Jensen, 2005; Lillesand et al., 2008). We produced error matrices in order to analyze and evaluate each method. These error matrices show the contingency of the class to which each pixel truly belongs (columns) on the map unit to which it is allocated by the selected analysis (rows). From the error matrix, overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient were generated. It has been suggested that a minimum of 50 sample points for each land-use land-cover category in the error matrix be collected for the accuracy assessment of any image classification (Congalton, 1991). We selected 500 samples points that led to approximately 70 points per class (7 total classes) for the accuracy assessment. A minimum of 50 points per class was set for generating 500 points using a stratified random sampling approach. To be consistent and for precise comparison purposes, we applied the same sample points generated for the output generated by the object-based classifier as the output produced by the traditional classification technique (i.e., maximum likelihood). The same accuracy assessment using the same sampling procedure was also performed for the test image.

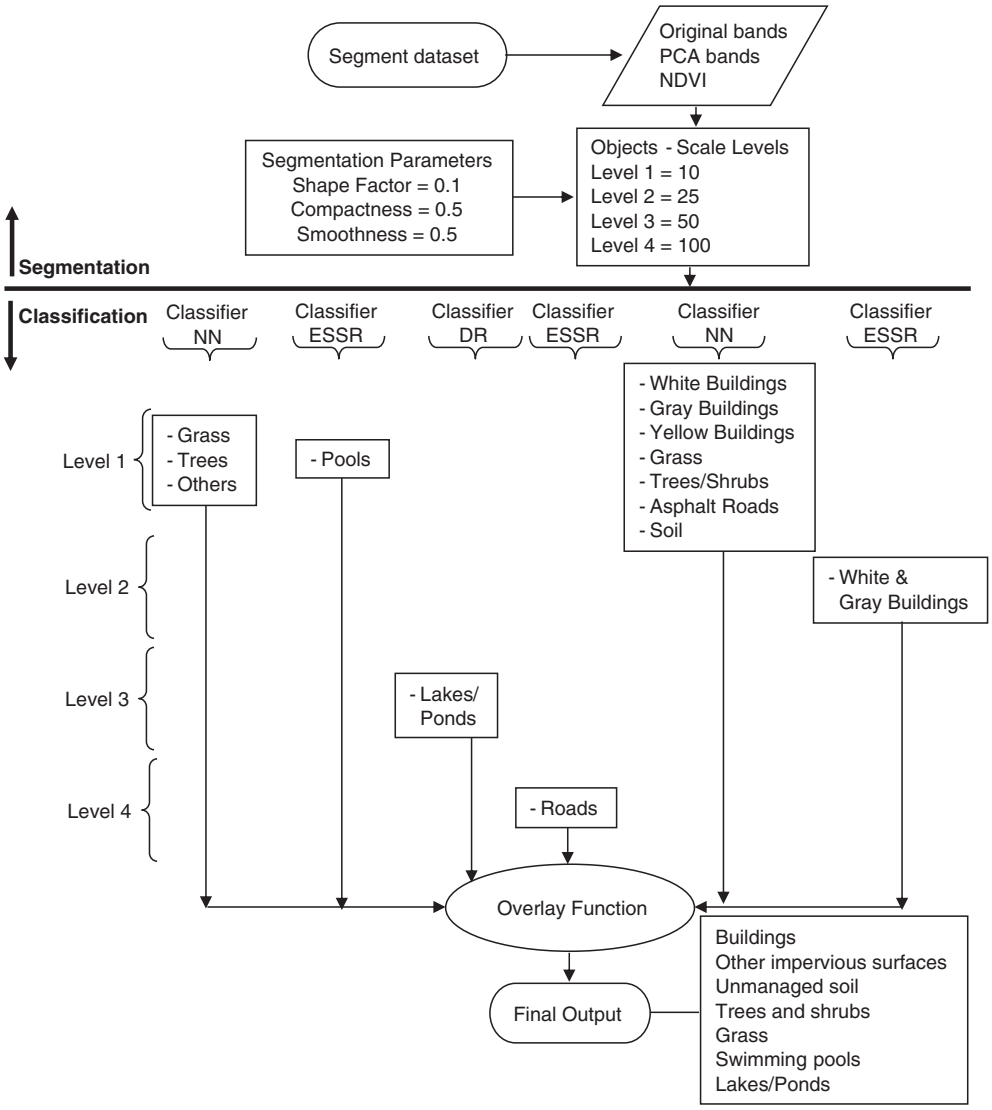

Fig. 11. A flow chart demonstrating the overall procedure to generate a final output.

5.2. Object-based accuracy assessment

We also performed an accuracy assessment at the object level. This was to demonstrate if the segmentation or size of objects at different

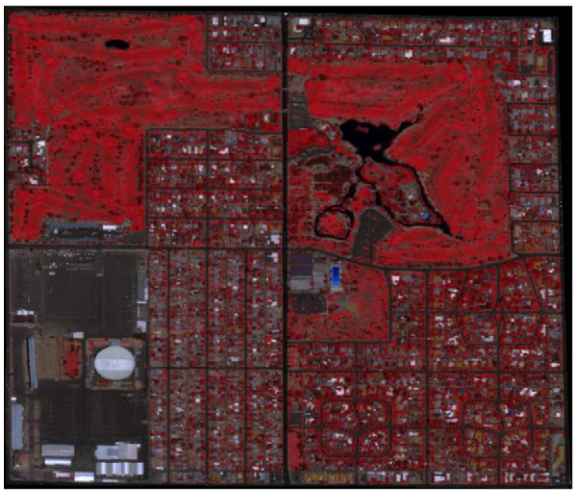

Fig. 12. A subset of the original QuickBird image.

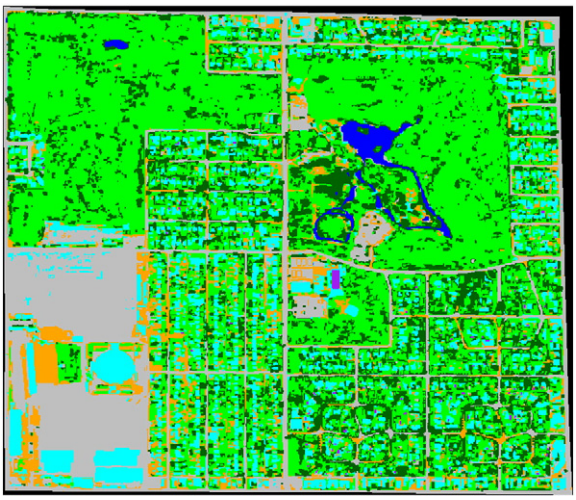

Fig. 13. Output map produced by the object-based approach. Note: Cyan = buildings; orange = unmanaged soil; light green = grass; gray = other impervious surfaces, purple = swimming pools; dark green = trees and shrubs; and blue = lakes and ponds.

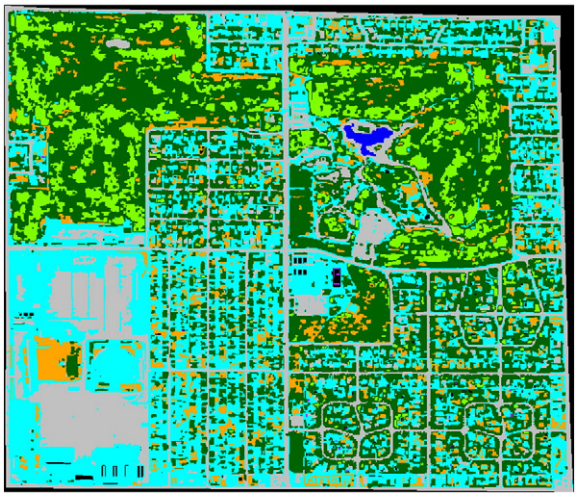

Fig. 14. Output map produced by the classical per-pixel classifier. Note: Cyan = buildings; orange = unmanaged soil; light green = grass; gray = other impervious surfaces, purple = swimming pools; dark green = trees and shrubs; and blue = lakes and ponds.

scale levels were appropriate for the classes selected. The object-based classification accuracy is determined by dividing the total number of correctly classified objects by the total number of objects. However, it should be noted that this type of assessment does not reflect the actual accuracy of an output whether each image pixel is accurately identified according to the reference data.

To perform the object-based accuracy assessment or segmentation assessment a total of 210 land-cover polygons or objects were randomly selected from the study site (the original QuickBird image). The polygons or objects were manually or visually interpreted and compared with the classification results derived from the object-based approach. We randomly selected 30 objects per class at each scale level and checked one class at a time with the final classification output since we identified different classes at different scale levels and integrated them at the final stage using a GIS overlay function. Moreover, some classes were identified at multiple scales, and some classes contained spectrally different members of the same class (e.g., yellow rooftops, white rooftops, and gray rooftops) that were extracted using different classification specificity (e.g., nearest neighbor, decision rule, scale, band combinations, and features). We checked grass, trees/shrubs, swimming pools, and soil at scale level 1, buildings at scale level 2, lakes/ponds at scale level 3, and roads at scale level 4. Since we followed the same segmentation procedure for both images, we performed the object-based accuracy assessment for the original image alone to demonstrate if the segmentation procedure was appropriate.

6. Results and discussion

6.1. Classification accuracies (object-based) — original image

At scale level 1, grass, trees/shrubs, pools, and unmanaged soil objects achieved classification accuracies of 93.33%, 80.00%, 96.67%, and 80.00% respectively (Table 4). Percentage of total areas correctly identified and total area of objects correctly identified for the same classes were 92.67% (641 acres), 81.68% (288 acres), 99.03% (91 acres), and 86.36% (544 acres) respectively. Buildings at scale level 2, lakes/ponds at scale level 3, and other impervious surfaces at scale level 4 produced object-based accuracies of 83.33%, 90.00%, and 93.33% respectively. Percentage of total areas correctly identified and total area of objects correctly identified for the same classes were 92.57% (792 acres), 98.53% (1815 acres), 97.06% (22,540 acres) respectively. It should be noted that object overall accuracies in general are a little lower than percent of total areas correctly identified. This could have been due to the fact that average size of correctly identified objects was bigger than objects of incorrectly identified objects in almost all classes. Two class objects that received the lowest accuracies were trees and soil. The second lowest object-based accuracy was produced by building category. As demonstrated earlier there were some signature confusion between the above three classes and other categories (e.g., trees vs. grass, soil vs. impervious, and buildings vs. other impervious surfaces). An average overall accuracy and an average percent of total areas correctly identified for all classes were 88.10% and 92.56% respectively. We believe object-based classification accuracies show that segmented objects at different scale levels were relevant. As mentioned earlier, the object-based classification accuracy does not reflect the actual classification accuracy of whether pixels in the image are accurately classified. This was just to demonstrate if the segmented objects at different level of scales were reasonable to perform object-based classification. It is important to note that image segmentation depends on many different factors such as classification system, area coverage of the study area, type of sensor, nature of the selected classes, radiometric resolution, spectral resolution, variation of object sizes within each class, and minimum size of object in a class that normally contains smaller objects (e.g., buildings). The limitation is that we cannot generate an object size bigger than minimum object size of any class since bigger objects will include other land cover classes around each object in the entire image. In general, smaller objects would work better for most classes and can be expected to obtain pure land cover categories. The best option would be to test potentially effective scale levels and select appropriate ones. Hence the image segmentation is somewhat subjective. But a rigorous visual inspection typically plays an important role. Testing class separability based on segmented objects may be taken as a rough assessment for scale selection but results from this type of test may have some uncertainty as the actual situation of all segmented objects in the image and variation of object sizes depend on many factors. On the other hand, the decision on whether a particular object boundary is correct is subjective. For the accuracy assessment of a final map generated by an object-based approach we would like to suggest that pixel-based accuracy assessment is the most appropriate approach since the smallest unit of an output is a pixel.

6.2. Classification accuracies (pixel-based) — original image

As mentioned earlier, we anticipated that the classification of urban land-cover classes with the traditional per-pixel approach (i.e., maximum likelihood) may not be very effective, since many urban land-cover classes are spectrally similar and classical approaches do not consider spatial arrangements of pixels. This was also demonstrated earlier in the discriminant analysis of the selected classes from thirty randomly selected points (Table 2). The overall classification accuracy produced only about 63.33% in the discriminant analysis. The spectra statistics (Table 1) of the selected land-cover classes indicated earlier that the traditional approaches would not be effective. By qualitative evaluation (visual examination on screen) of the output maps, we noticed that the output map generated by the traditional per-pixel approach contains many mistakenly identified pixels of classes (Fig. 14), whereas the output map generated by the object-based classifier approach looks more accurate (Fig. 13). We used output maps of a small portion of the entire study area to qualitatively demonstrate which output map seems more accurate.

From Table 5, it can be observed that the per-pixel classifier produced low overall accuracy (67.60%) and kappa coefficient (0.62). This accuracy was slightly higher than the discriminant analysis (63.33%) with the use of spectra of the selected classes. The lowest producer's accuracy (50%) was given by the building class. The

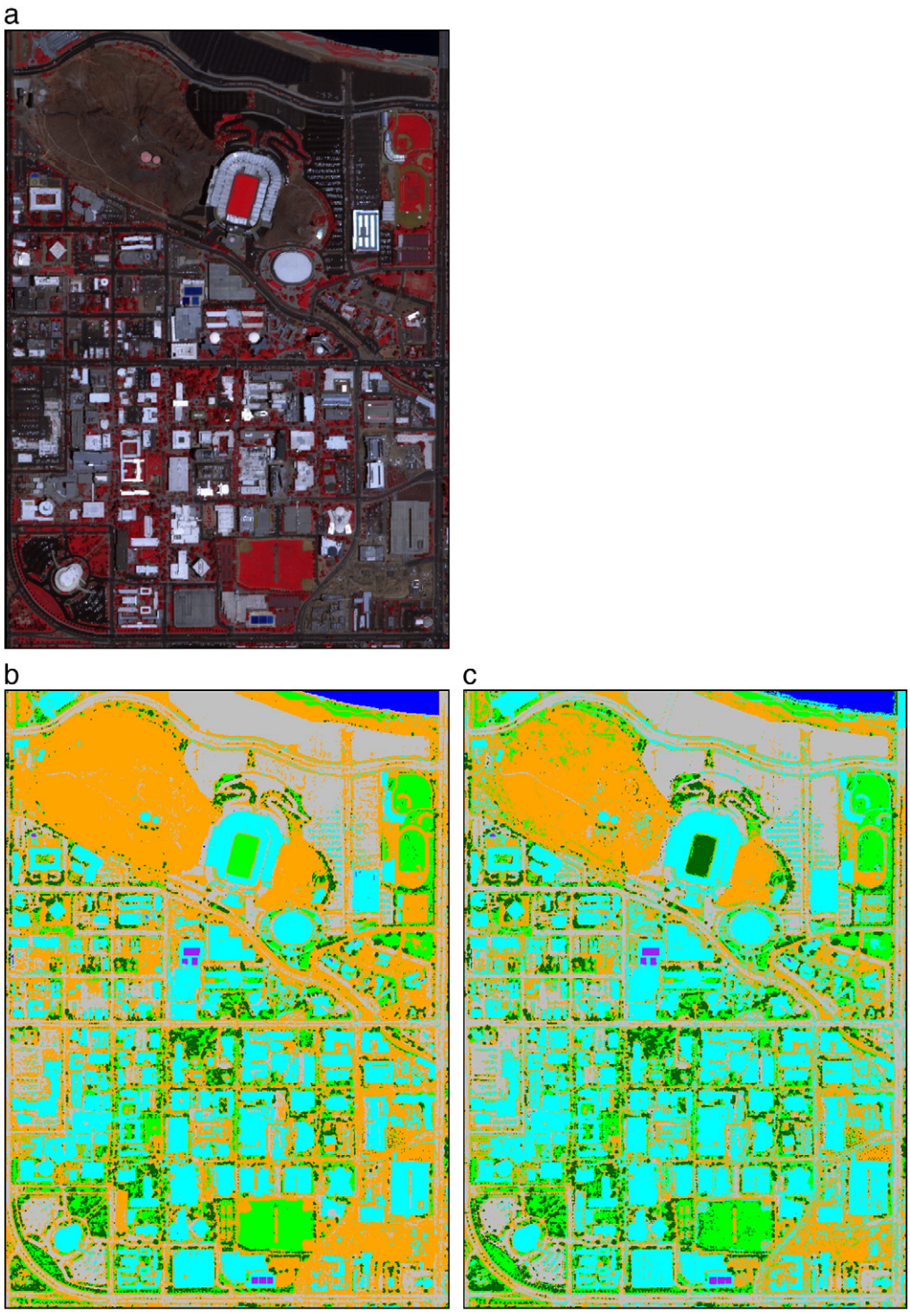

Fig. 15. (a) Test image; (b) output map produced by the object-based approach; (c) output map produced by the classical per-pixel classifier. Note: Cyan = buildings; orange = unmanaged soil; light green = grass; gray = other impervious surfaces, purple = swimming pools; dark green = trees and shrubs; and blue = lakes and ponds.

classification can reasonably claim that only 50% of the time an area that was identified as buildings was identified as such in the output map. As expected earlier, there was significant signature confusion among buildings, other impervious surfaces, swimming pools, and unmanaged soil. Trees and shrubs category produced the second lowest producer's accuracy (62.50%), since there was some significant signature confusion between trees/shrubs and grass. Almost one third of the total sample points identified as grass were found to be trees and shrubs on the ground (23 out of 72). The other low producer's accuracies were produced by grass (64.41%) and impervious (62.50%) categories. It was also found that about one third of the sample points identified as unmanaged soil were found to be the impervious category. Significant number of sample points that were identified as unmanaged soil, swimming pools, and trees/shrubs actually belonged to grass. It was found that the lowest user's accuracy (50.00%) was produced by the unmanaged soil category, and a majority of the mistakenly identified sample points under this category belong to buildings and other impervious classes. As we discussed earlier,

Table 4
Object-based classification accuracy — original image.

Class	Scale	Correctly identified objects	Incorrectly identified objects	Correctly identified area (acre)	Incorrectly identified area (acre)	Percent correctly identified objects	Percent correctly identified area
Buildings	2	25	5	792.17	63.60	83.33	92.57
Unmanaged soil	1	24	6	543.53	85.84	80.00	86.36
Grass	1	28	2	640.72	50.71	93.33	92.67
Other impervious	4	28	2	22540.38	683.20	93.33	97.06
Pools	1	29	1	91.18	0.89	96.67	99.03
Trees/shrubs	1	24	6	287.56	64.49	80.00	81.68
Lakes/ponds	3	27	3	1814.96	27.13	90.00	98.53

spectral responses from bright soil are similar to bright impervious surfaces and some building rooftops. The user's accuracy of the swimming pool category was also exceptionally low (50.77%). There were many different sample points identified as swimming pools actually belonging to all other classes. We did not anticipate this signature confusion, since spectral responses from some categories do not seem similar to swimming pools. Another low user's accuracy was produced by the grass category. There were many sample points that were identified as grass which were trees and shrubs. From the preceding discussion, it can be noted that the traditional per-pixel approach is not effective in urban mapping, since the approach is unable to handle spectrally similar classes. The accuracy of the classical approach is low because there are many spectrally similar land-cover classes within an urban environment.

In contrast to the classical approach, the object-based classifier produced a significantly higher overall accuracy (90.40%) and kappa coefficient (0.89) (Table 6). Buildings and other impervious categories produced relatively low producer's accuracies (83.91%, 83.65%). The spectral response from buildings has some signature confusion with unmanaged soil and grass. It is not very unusual to see signature confusion between buildings and impervious surfaces, as both categories share many land cover materials (e.g., cement and asphalt). However, signature confusion between buildings and grass seemed a little unusual. This could have been due to the fact that some grassy areas in and on the edges of residential and commercial areas might have been partly included as building objects, especially when considering lower level scales (larger scale parameters) that consider bigger objects having contiguous pixels with similar spectral and spatial properties. Even though we focus on segmented objects of the object-based paradigm, it is important to note that spectral information still plays an important role in determining segmented objects. There are only 4 bands in the image data and three of them were in the visible portion of the electromagnetic spectrum. There is also a signature confusion between tress/shrubs category. This could have been due to the same reason given above. The lowest user's accuracy was given by the grass category (79.07%). This user's accuracy was significantly lower than all other categories, since they all reached at least 85.00%. This accuracy was low due to some signature confusion with buildings and trees/shrubs classes. Lakes/ponds category achieved the highest accuracy (100%) for both producer's and user's accuracies. The second highest producer's and user's accuracy was produced by swimming pool category (97.96%, 96.00%, respectively).

To evaluate the output in a different way, we performed a regression analysis between swimming pool area (sq m), identified in each census tract by the object-based technique, and number of pools within the census tract compiled from, and prepared by the Maricopa

Table 5
Overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient produced by the per-pixel classifier — original image.

Classified	Reference								Producer's accuracy (%)	User's accuracy (%)
	Buildings	Unmanagedsoil	Grass	Other impervious	Pools	Trees/shrubs	Lakes/ponds	Total		
Buildings	39	4	0	2	3	0	0	48	50.00	81.25
Unmanagedsoil	12	61	8	39	0	1	1	122	81.33	50.00
Grass	0	2	38	0	0	23	0	63	64.41	60.32
Other impervious	16	5	0	82	1	0	8	112	63.57	73.21
Pools	11	3	8	6	33	3	1	65	89.19	50.77
Trees/shrubs	0	0	5	0	0	45	0	50	62.50	90.00
Lakes/ponds	0	0	0	0	0	0	40	40	80.00	100.00
Total	78	75	59	129	37	72	50	500		

Overall accuracy = 67.60%.
Overall kappa statistics = 0.62.

Table 6
Overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient produced by the object-oriented classifier — original image.

Classified	Reference								Producer's accuracy (%)	User's accuracy (%)
	Buildings	Unmanaged soil	Grass	Other impervious	Pools	Trees/shrubs	Lakes/ponds	Total		
Buildings	73	2	1	3	0	1	0	80	83.91	91.25
Unmanaged soil	6	70	0	3	1	0	0	80	94.59	87.50
Grass	6	2	68	2	0	8	0	86	95.77	79.07
Other impervious	1	0	0	87	0	0	0	88	83.65	98.86
Pools	1	0	0	1	48	0	0	50	97.96	96.00
Trees/shrubs	0	0	2	8	0	56	0	66	86.15	84.85
Lakes/ponds	0	0	0	0	0	0	50	50	100.00	100.00
Total	87	74	71	104	49	65	50	500		

Overall accuracy = 90.40%.
Overall kappa statistics = 0.89.

Table 7
Overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient produced by the per-pixel classifier — test image.

Classified	Reference								Producer's accuracy (%)	User's accuracy (%)
	Buildings	Unmanaged soil	Grass	Other impervious	Pools	Trees/shrubs	Lakes/ponds	Total		
Buildings	84	4	1	5	0	2	0	96	95.45	87.50
Unmanaged soil	1	77	2	16	0	2	0	98	89.53	78.57
Grass	1	3	49	2	0	11	0	66	89.09	74.24
Other impervious	1	1	1	73	0	1	0	77	74.49	94.81
Pools	0	0	0	0	50	0	0	50	100.00	100.00
Trees/shrubs	1	1	2	2	0	55	1	62	77.46	88.71
Lakes/ponds	0	0	0	0	0	0	51	51	98.08	100.00
Total	88	86	55	98	50	71	52	500		

Overall accuracy = 87.80%.
Overall kappa statistics = 0.86.

County Assessor Office — 2003. The entire study area covers 51 census tracts, and we used all census tracts to conduct the analysis. The relationship was found to be exceptionally strong (0.98). This implies that our membership function using an expert system rule to identify pools was effective and the output map for this category was accurate. The regression to predict swimming pool area in sq m per census tract from the number of swimming pools per census tract in city of Phoenix can be described as

$$y = 36.888x + 845.22 \quad (1)$$

where y = swimming pool area per census tract; x = number of swimming pools per census tract. The analysis to examine the effectiveness of identifying pool area per census tract can be used to predict swimming pool area in any census tracts.

6.3. Classification accuracies (pixel-based) — test image

Since the area coverage of the test image is smaller and spatial arrangements of objects and land cover classes are less complex than the main image, the overall classification accuracies for both the object-based approach and the maximum likelihood classifier were higher than those generated by the main image. The Object-based approach reached an overall accuracy as high as 95.2% (Table 8), whereas the maximum likelihood approach yielded 87.80% (Table 7). In both cases, the pool category produced the highest producer's and user's accuracy (100%). The second highest producer's accuracy (98%) and user's accuracy (100%) was produced by lakes/ponds class. The reason is that there were not many pools and lakes/ponds in the study area, since the study area does not contain any recreational, residential and commercial areas. This is one of the reasons why the overall accuracies generated by both approaches were high. The two lowest producer's accuracies produced by the maximum likelihood were other impervious (74.49%) and trees/shrubs (77.64%). The two lowest users' accuracies produced by the same traditional approach were other impervious (74.24%) and grass categories (78.57%). However, the object-based approach produced high producer's and user's accuracies. The only slightly lower users' accuracy was produced by unmanaged soil (83.48%). The overall accuracy produced by the object-based approach is about 8% higher than the maximum likelihood. The gap between the overall accuracies between the two approaches with the test image was smaller than those with the original image. This is mainly because the test image covers a smaller area with less complex urban objects and categories. For example, some land cover categories may contain significantly different sizes of objects and noticeably different spectral responses or surface materials within class. Finer resolution images that cover larger urban areas tend to cover more complex or irregular objects and more surface materials in each land cover category that can potentially lead to classification errors especially when dealing with decision rule approach. By visual inspection, it can be observed from Fig. 14 that there are noticeable errors in the output produced by the maximum likelihood approach. However, this is not the case for the object-based method.

7. Conclusion

We have shown that the traditional per-pixel approaches were not very effective in identifying urban land-cover classes. This was proven by the classification of the entire QuickBird image using the most widely used classifier namely maximum likelihood rule and spectra of the selected land cover classes generated from the QuickBird using discriminant analysis. The discriminant analysis of spectra information received an overall accuracy of 63.33%. The object-based classifier produced a significantly higher overall accuracy (90.40%), whereas the maximum likelihood classifier produced 67.60%. Segmentation procedures and scale levels employed to identify objects of different classes were found to be relevant. The same classification procedures and classification accuracy assessment employed to classify the same classes using the test image confirms that the object-based approach (95.2%) outperforms the traditional classifier (87.8%). This study reveals that it is more difficult to achieve higher accuracies for larger

Table 8
Overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient produced by the object-oriented classifier — test image.

Classified	Reference								Producer's accuracy (%)	User's accuracy (%)
	Buildings	Unmanaged soil	Grass	Other impervious	Pools	Trees/shrubs	Lakes/ponds	Total		
Buildings	82	1	0	0	0	0	0	83	94.25	98.80
Unmanaged soil	5	96	7	3	0	4	0	115	97.96	83.48
Grass	0	0	63	0	0	1	0	64	88.73	98.44
Other impervious	0	0	1	76	0	1	0	78	96.20	97.44
Pools	0	0	0	0	50	0	0	50	100.00	100.00
Trees/shrubs	0	1	0	0	0	58	0	59	90.63	98.31
Lakes/ponds	0	0	0	0	0	0	51	51	100.00	100.00
Total	87	98	71	79	50	64	51	500		

Overall accuracy = 95.20%.
Overall kappa statistics = 0.94.

images when dealing with a detailed urban mapping. There is no universally accepted method to determine an optimal scale level to segment objects. Moreover, a scale level may not be suitable for all classes in an image classification. The best approach to select which bands to be considered for the membership function and which scale level to employ for a particular class would be to identify the class with different options and qualitatively analyze them on the display screen as a generate and test approach. The output map needs to be checked carefully throughout the image. We note that some membership functions seemed to work very well for a particular class in one part of the study area, but it did not perform well for the same class in other parts. The analyst needs to check a zoom-in version of the output at all possible spots before observing other classification options at different level of scales. It would be a good idea to treat individual classes separately to explore which method and scale level with which feature space is potentially good to extract a class. However, in some cases, classification of a few classes together in multispectral bands with the use of the nearest neighbor option could be better than individual classes separately using a membership function with a set of expert system rules. Even though there were seven classes in our classification, we kept several different classes or many different training samples of the same classes to perform the classification. The threshold values given in this study may not be applicable to other urban mapping using the same satellite data (QuickBird), even though the classification system employs the same classes in a similar urban environment. However, similar or considerably different threshold values with a slight modification of the parameters can be expected to be effective for urban mapping in different environmental settings.

Since both classifiers available in the object-based approach are non-parametric rules, they are independent of the assumption that data values need to be normally distributed. This is advantageous, because most data are not normally distributed in many real world situations. One of the other advantages of the object-based approach is that it allows additional selection or modification of new objects (training samples) each time, after performing a nearest neighbor classification quickly until the satisfactory result is obtained. There are many possible combinations of different functions, parameters, features, and variables available with the object-based approach. The successful use of the object-based paradigm largely relies on repeatedly modifying training objects, performing the classification, observing the output, and/or testing different combinations of functions as a trial-and-error process.

Our experience was that Definiens or eCognition software was not able to perform many features or bands at many different scale levels for image segmentation and classification. This was simply because the computer memory needs to be used extensively to segment tremendous numbers of objects from many different bands, especially when requiring smaller scale parameters (larger scale segmentation). We used different computer hardware and experienced numerous computer breakdowns and freezes during the segmentation, even though our study area is a small part of the whole Phoenix metropolitan area. This should be considered a limitation especially when dealing with a large dataset (finer resolution data for a relatively large area). Nonetheless, the object-based classification system is a better approach than the traditional per-pixel classifiers in urban mapping using high-resolution imagery.

References

Asner, G. P., & Heidebrecht, K. B. (2002). Spectral unmixing of vegetation, soil and dry carbon cover in arid regions: Comparing multispectral and hyperspectral observations. *International Journal of Remote Sensing*, *19*, 3939−3958.

Baatz, M., & Schape, A. (1999). Object-oriented and multi-scale image analysis in semantic networks. *Proceedings of the 2nd international symposium on operationalization of remote sensing, 16–20 August 1999*. Enschede: ITC.

Baatz, M., & Schape, A. (2000). Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. *Proceedings of the Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT Symposium. Salzburg, Austria.*

Campbell, J. (2007). *Introduction to remote sensing* (4th ed.). New York: The Guilford Press.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*, 35−46.

Congalton, R. G., & Green, K. (1999). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Lewis Publishers, Boca Raton, Florida, 137 p.

Cowen, D. J., Jensen, J. R., & Bresnahan, P. J. (1995). The design and implementation of an integrated geographic information system for environmental applications. *Photogrammetric Engineering and Remote Sensing*, *61*, 1393−1404.

De Jong, S. M., & Burrough, P. A. (1995). A fractal approach to the classification of Mediterranean vegetation types in remotely sensed images. *Photogrammetric Engineering and Remote Sensing*, *61*, 1041−1053.

Definiens (2008). *Definiens Developer 7.0, user guide* (pp. 506).

Desclée, B., Bogaert, P., & Defourny, P. (2006). Forest change detection by statistical object-based method. *Remote Sensing of Environment*, *102*, 1−11.

Ferro, C. J. S., & Warner, T. A. (2002). Scale and texture in digital image classification. *Photogrammetric Engineering and Remote Sensing*, *68*, 51−63.

Franklin, S. E., Hall, R. J., Moskal, L. M., Maudie, A. J., & Lavigne, M. B. (2000). Incorporating texture into classification of forest species composition from airborne multispectral images. *International Journal of Remote Sensing*, *21*, 61−79.

Gober, P., Brazel, A. J., Myint, S., Quay, R., Miller, A., Rossi, S., & Grossman-Clarke, S. (2010). Using watered landscapes to manipulate urban heat island effects: How much water will it take to cool Phoenix? *Journal of the American Planning Association*, *76*, 109−121.

Grimmond, C. S. B., & Oke, T. R. (2002). Turbulent heat fluxes in urban areas: Observations and local-scale urban meteorological parameterization scheme (LUMPS). *Journal of Applied Meteorology*, *41*, 792−810.

Im, J., Jensen, J. R., & Hodgson, M. E. (2008). Object-based land cover classification using high posting density lidar data. *GIScience and Remote Sensing*, *45*, 209−228.

Im, J., Jensen, J. R., & Tullis, J. A. (2008). Object-based change detection using correlation image analysis and image segmentation techniques. *International Journal of Remote Sensing*, *29*, 399−423.

Ivits, E., & Koch, B. (2002). Object-oriented remote sensing tools for biodiversity assessment: A European approach. *Proceedings of the 22nd EARSeL symposium, Prague, Czech Republic, 4–6 June 2002*. Rotterdam, Netherlands: Millpress Science Publishers.

Jensen, J. R. (2005). Introductory digital image processing: A remote sensing perspective (3rd ed.). Upper Saddle River: Prentice-Hall 526 pp.

Jensen, J. R., & Cowen, D. C. (1999). Remote sensing of urban/suburban infrastructure and socio-economic attributes. *Photogrammetric Engineering and Remote Sensing*, *65*, 611−622.

Lam, N. S. N., & Quattrochi, D. A. (1992). On the issues of scale, resolution, and fractal analysis in the mapping sciences. *Professional Geographer*, *44*, 88−97.

Lee, J. Y., & Warner, T. A. (2006). Segment based image classification. *International Journal of Remote Sensing*, *27*, 3403−3412.

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2008). *Remote sensing and image interpretation.* New York: John Wiley and Sons.

Lu, D., & Weng, Q. (2004). Spectral mixture analysis of the urban landscape in Indianapolis with Landsat ETM+ Imagery. *Photogrammetric Engineering and Remote Sensing*, *70*, 1053−1062.

Lucieer, A. (2004). Uncertainties in segmentation and their visualisation, Ph D., *International Institute for Geo-Information Science and Earth Observation (ITC) and the University of Utrecht*, Netherlands, 177 pp.

Möller, M., Lymburner, L., & Volk, M. (2007). The comparison index: A tool for assessing the accuracy of image segmentation. *International Journal of Applied Earth Observation and Geoinformation*, *9*, 311−321.

Mueller, M., Segl, K., & Kaufmann, H. (2004). Edge- and region-based segmentation technique for the extraction of large, man-made objects in high-resolution satellite imagery. *Pattern Recognition*, *37*, 1619−1628.

Munoz, X., Freixenet, J., Cufi, X., & Marti, J. (2003). Strategies for image segmentation combining region and boundary information. *Pattern Recognition Letters*, *24*, 375−392.

Myint, S. W. (2006). A new framework for effective urban land use land cover classification: A wavelet approach. *GIScience and Remote Sensing*, *43*, 155−178.

Myint, S. W., Giri, C. P., Wang, L., Zhu, Z., & Gillette, S. (2008). Identifying mangrove species and their surrounding land use and land cover classes using an object oriented approach with a lacunarity spatial measure. *GIScience and Remote Sensing*, *45*, 188−208.

Myint, S. W., & Lam, N. S. N. (2005). Examining lacunarity approaches in comparison with fractal and spatial autocorrelation techniques for urban mapping. *Photogrammetric Engineering and Remote Sensing*, *71*, 927−937.

Myint, S. W., May, Y., Cerveny, R., & Giri, C. P. (2008). Comparison of remote sensing image processing techniques to identify tornado damage areas from Landsat TM data. *Sensors*, *8*, 1128−1156.

Myint, S. W., Mesev, V., & Lam, N. S. N. (2006). Texture analysis and classification through a modified lacunarity analysis based on differential box counting method. *Geographical Analysis*, *38*, 371−390.

Myint, S. W., Wentz, E., & Purkis, S. (2007). Employing spatial metrics in urban land use/land cover mapping: Comparing the Getis and Geary indices. *Photogrammetric Engineering and Remote Sensing*, *73*, 1403−1415.

Navulur, K. (2007). *Multispectral image analysis using the object-oriented paradigm*. Boca Raton, FL: CRC Press, Taylor and Frances Group.

Okin, G. S., Roberts, D. A., Murray, B., & Okin, W. J. (2001). Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sensing of Environment*, *77*, 212−225.

Purkis, S. J., Myint, S. W., & Riegl, B. M. (2006). Enhanced detection of the coral *Acropora cervicornis* from satellite imagery using a textural operator. *Remote Sensing of Environment*, *101*, 82−94.

Roberts, D. A., Dennison, P. E., Gardner, M., Hetzel, Y. L., Ustin, S. L., & Lee, C. (2003). Evaluation of the potential of Hyperion for fire danger assessment by comparison to the Airborne Visible Infrared Imaging Spectrometer. *IEEE Transactions on Geoscience and Remote Sensing*, *41*, 1297—1310.

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R. O. (1998). Mapping Chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment*, *65*, 267—279.

Schowengerdt, R. A. (1995). Soft classification and spatial-spectral mixing. *Proceedings of international workshop on soft computing in remote sensing data analysis. 4–5 December 1995. Milan, Italy*.

Stow, D., Hamada, Y., Coulter, L., & Anguelova, Z. (2008). Monitoring shrubland habitat changes through object-based change identification with airborne multi-spectral imagery. *Remote Sensing of Environment*, *112*, 1051—1061.

Stow, D., Lopez, A., Lippitt, C., Hinton, S., & Weeks, J. (2007). Object-based classification of residential land use within Accra, Ghana based on QuickBird satellite data. *International Journal of Remote Sensing*, *28*, 5167—5173.

Walter, V. (2004). Object-based classification of remote sensing data for change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, *58*, 225—238.

Weidner, U. (2008). Contribution to the assessment of segmentation quality for remote sensing applications. *International archives of photogrammetry and remote sensing, XXXVII, part B7, Beijing* (pp. 479—484).

Weng, Q., & Hu, X. (2008). Medium spatial resolution satellite imagery for estimating and mapping urban impervious surfaces using LSMA and ANN. *IEEE Transactions on Geosciences and Remote Sensing*, *46*, 2397—2406.

Winter, S. (2000). Location similarity of regions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *55*, 189—200.