A Review of the Effectiveness of Spatial Information used in Urban Land Cover Classification of VHR Imagery

Salehi, B.[1], Zhang, Y.[2], Zhong, M.[3] and Dey, V.[4]
University of New Brunswick, 15 Dineen Drive, Fredericton, New Brunswick, Canada, E3B 5A3
E-mail: bsalehi@unb.ca[1], yunzhang@unb.ca[2], ming@unb.ca[3], vivekde@gmail.com[4]

Abstract

Land cover classification of very high resolution (VHR) imagery in urban areas is an extremely challenging task, because of the low intra-class (within-class) and high inter-class (between-classes) spectral similarities of impervious land cover types (such as buildings and traffic areas). Over the past decade, a significant amount of research has been conducted on the incorporation of spatial information along with spectral information of VHR imagery into urban land cover classification. The spatial information includes textural, morphological and contextual measures extracted from VHR imagery, as well as LiDAR- and photogrammetrically-derived DSM and existing GIS data layers. In this paper, a comprehensive review of recent literature was conducted to evaluate the effectiveness of such measures in land cover classification of urban areas using VHR imagery. For each measure, a comprehensive list of papers for both pixel-based and object-based classification is provided. In addition, the classification results of representative publications are reported and its advantages and limitations in both pixel-based and object-based approaches are discussed. It has been found that, in general, object-based classification performs better than pixel-based approaches, since it facilitates the use of spatial measures by segmenting the image. Moreover, utilizing spatial measures significantly improves the classification performance for impervious land cover types, while may have no effect or even lower the classification accuracy for classes of vegetation and water surfaces. Textural measures are more commonly utilized in pixel-based approaches, while morphological measures have better performance in object-based classification. The effect of contextual measures on classification is enhanced when these measures are used in conjunction with two other measures, particularly in object-based approaches. Although ancillary data shows a very high potential to address the problem of spectral-based classifiers in separating spectrally similar impervious land cover types, incorporating such data, particularly photogrammetrically-derived DSM, in classification is still in a very early stage and requires significant exploration and development.

1. Introduction

Land cover classification is one of the most important topics in remote sensing both for researchers and practitioners, because of its broad applications in almost all geo-related domains. Remotely sensed images are the major, and sometimes the only, input in land cover classification. The spatial resolution of the image is one of the most important factors that affect land cover classification performance (Chen et al., 2004). Previous research has explored the impact of spatial resolution on classification of remotely sensed data (e.g. Price, 1997 and Quattrochi and Goodchild, 1997). Because of the sub-meter ground resolution, VHR images unveil a very high potential for more detailed and accurate mapping of the urban environment (Pacifici et al., 2009). New VHR digital aerial cameras provide an excellent data source for the mapping and classification of urban areas, but their images are expensive and not easy to collect in a short period of time. With the advent of very high spatial resolution (≤1m) satellite sensors since 1999, such as IKONOS, QuickBird, OrbView, WorldView-1, GeoEye-1 and WorldView-2, urban land cover classification has rapidly gained interest within the remote sensing community. However, the increased spatial resolution of VHR imagery does not automatically yield improved accuracy of urban land cover classification, if classifiers just employ the spectral information of the image (spectral-based classifiers). This is mainly due to the high spectral variation within the same land cover (intra-class spectral variation; e.g. buildings with different roof types) and the spectral confusion between different land covers (inter-class spectral confusion) (Lu et al., 2010, Xu and Li, 2010 and Huang et al., 2011). To compensate for the limitations of spectral-based classifiers, many researchers have attempted to develop techniques to incorporate spatial

ISSN 1686-6576 / © Geoinformatics International

information extracted from VHR imagery and/or from ancillary data into classification. For the sake of convenience, we categorize the spatial information into four types: textural, contextual and morphological measures (extracted from the VHR image), and ancillary data such as digital elevation/surface model (DEM/DSM) derived from LiDAR or stereo photographs and existing GIS data layers (e.g. road network and buildings' footprint). This paper aims to review the most recent research on urban land cover classification using VHR satellite imagery. Specifically, the objective is to evaluate the effectiveness of incorporating the four aforementioned types of spatial measures in both pixel-based and object-based classification approaches. Comprehensive lists of major publications (including peer-reviewed journal papers) in which these four types of spatial measures have been utilized in both pixel-based and object-based classification approaches are reported. In addition, the effect of each type of measures on increasing the classification accuracy is quantitatively evaluated by reporting the accuracies achieved in some recent literature. To date, we have not found a comprehensive review of different spectral and spatial measures used in classification of VHR imagery, particularly over urban areas. Lu and Weng (2007) conducted a survey of image classification methods and techniques for improving classification performance. Their survey includes a brief description of general process of image classifications, with the citation of a large amount of previously published literature. However, it did not focus on the classification of VHR imagery in urban environment. Liu et al., (2006) briefly reviewed classification patterns of remotely sensed imagery based on object-oriented approaches. Their study, however, is limited to describing basic steps of object-oriented image analysis along with reviewing a few selected publications. Gamba et al., (2005) presented a bibliographic review of the state-of-the-art of urban remote sensing using multiple data sets. They briefly reviewed the data fusion issues in urban areas without taking into consideration the capabilities of VHR imagery in urban analysis.

2. Spatial Information used in Classification

The aforementioned spatial measures are utilized in either pixel-based or object-based classification in order to help the classifier distinguish different land cover classes. Since the nature of such measures and their effects on classification are different in pixel-based and object-based image analysis, it is useful to give a brief review of these two generic types of classification before proceeding to the review of the measures.

2.1 Pixel-Based vs. Object-Based Classification

In general, image classification approaches can be grouped into different categories such as supervised or unsupervised, parametric or non-parametric, hard or soft (fuzzy) (Lu and Weng, 2007). In each category the basic processing unit could be pixel or object; accordingly approaches are described as pixel-based (or per-pixel) and object-based. Although pixel-based approaches are still widely used for mapping particular urban impervious land cover types such as large commercial parcels, the distribution and shape of such cover types in heterogeneous areas may be more accurately mapped by object-based classification approaches (Hester et al., 2008). A serious problem associated with pixel-based classifiers is the so-called "salt and pepper" effect or "structural clutter" (Van de Voorde et al., 2007), which produces a noisy classification result due to the spectral heterogeneity of classes in an urban environment and the lack of topological information used to classify pixels (Sims and Mesev, 2007). This effect is more significant in the VHR image classification especially over complex urban environments, because of the high intra-class spectral heterogeneity of some surface types. To reduce this negative effect of pixel-based classification approaches, some studies have proposed post-classification techniques (e.g. Van de Vorde et al., 2007, Hester et al., 2008).These techniques, however, may remove small land cover types such as single-family houses and single trees. The object-based classification approach, on the other hand, decreases variance within the same land cover type by averaging the pixels within the object, which prevents the "salt and pepper" effect of pixel-based classification approaches (Chen et al., 2009a). Starting from around the year 2000, studies of object-based images analysis have sharply increased (Blaschke, 2010). A comprehensive list of recent literature concerning object-based urban land cover classification of VHR imagery is provided in tables 1, 4, 6, and 8. Other sources of object-based image analysis research include three online archives of conference proceedings (Object-based Image Analysis (OBIA, 2006), GEographic Object-based Image Analysis (GEOBIA, 2008 and GEOBIA, 2010)), a book published on object-based image analysis by Blaschke et al., (2008) and a literature review paper by Blaschke, (2010). It is noteworthy that much of the work referring object-based image analysis originated around the "eCognition" software (Benz et al., 2004). The next section reviews the spatial measures utilized in classification of land covers in a typical urban environment.

The measures incorporated into both pixel-based and object-based classification are discussed and the advantages and limitations of each type of measure are evaluated.

2.2 Spectral-Based Classifiers

Spectral-based classifiers have promising performance when applied to the medium and high spatial resolution images with several spectral bands for mapping relatively large homogeneous areas such as vegetation, forest, water and soil (Lu and Weng, 2007, McMahon, 2007 and Xu and Li, 2010). However, because of similarities in the spectral response of land cover types in an urban scene, along with the low spectral resolution of VHR imagery (the limited number of spectral bands and the wide wavelength range covered by them), the classification of such imagery is a challenging task (Zhang and Couloigner, 2006). Despite the fine spatial resolution of VHR imagery, its spectral resolution is limited to four multispectral bands (except the newly launched WorldView-2, which has eight multispectral bands) and a panchromatic band. Moreover, most of the bands that are suitable for separating urban land cover types lie outside or near the boundaries of the wavelength range of the multispectral bands of VHR imagery (Herold et al., 2003b) (e.g. at wavelength around 580 nm which lies in the boundaries of the Green band or at wavelengths around 740nm which lies outside of Red and NIR bands of VHR images). Ben-Dor (2001), Herold et al., (2004) and Warner and Nerry (2009) concluded that the shortwave and thermal infrared spectral regions are important for urban applications. These bands, however, are not present in the VHR imagery. Thomas et al., (2003) stated that the lack of mid-infrared bands in VHR images hinders the ability of traditional spectral-based classifiers to accurately distinguish detailed land cover types. Herold et al., (2003b) also concluded that some land cover classes such as asphalt road, tar roof and parking lot have a very similar and constant low reflectance over the whole spectral range such that even AVIRIS hyperspectral sensor has limitation in mapping these classes.

2.3 Spatial Measures Extracted From the Image

Spectral-based classifiers detect land cover classes exclusively according to spectral information while the large amount of valuable image spatial information is neglected. Moreover, in an urban landscape, impervious classes are spectrally too similar to be distinguished using only spectral information of the image. Hence, for the mapping of such classes, it is necessary to incorporate spatial information together with spectral information in the classification process. Two distinct types of method which utilize spatial information from an image are region-based and window-based methods (Gong et al., 1992). The region-based method is usually used in object-based, whereas the window-based method is used in pixel-based approaches. In the following sub-sections, the performance of each type of spatial measures, i.e. textural, contextual and morphological measures of the image and spatial measures of ancillary data in both pixel-based and object-based approaches is reviewed. According to the results achieved in individual publications, the strengths and limitations of each group are discussed. The increase of classification accuracy resulting from the incorporation of spatial information into classification over conventional spectral-based classifiers is reported. The classification accuracies reported through this review paper are all based on *Error* or *Confusion matrix* (Congalton, 1991 and Richards and Jia, 2006) unless otherwise specified. For most literature, mentioned in tables, the Overall Accuracy (OA) and/or Kappa coefficient (KA) are used. In some cases, where available, the Producers' Accuracy (PA) of the classification of land covers such as buildings and roads is reported as well.

2.3.1Textural measures

Many researchers have attempted to employ texture measures as additional spatial information in the urban land cover classification of VHR images to overcome the lack of spectral information (Carleer and Wolff, 2006 and Myint, 2007) in both pixel-based and object-based classification approaches (Table 1).

Table 1: List of papers which have utilized textural measures in pixel-based or object-based urban land cover classification of VHR imagery

Pixel-based classification	Object-based classification
Zhang (1999); Pesaresi (2000); Maillard (2003); Shackelford and Davis (2003a); Shackelford and Davis (2003b); Chen et al. (2004); Myint et al. (2004); Walter (2004); Mena and Malpica (2005); Myint and Lam (2005); Puissant et al. (2005); Zhang and Couloigner (2006); Alonso et al. (2007); Myint (2007); Agüera et al. (2008); Aksoy et al. (2009); Pacifici et al. (2009); Luo and Mountrakis (2010); Lu et al. (2010); Ouma et al. (2010); Tassetti et al. (2010).	Herold et al. (2003a); Thomas et al. (2003); Walter (2004); Song et al. (2005); Carleer and Wolf (2006); Su et al. (2008); Chan et al. (2009); Hermosilla et al. (2011); Pu et al. (2011); Salehi et al. (2011a).

The result of our literature review indicates that texture features extracted from gray level statistics, especially those of co-occurrence gray level matrix (GLCM), are the most useful texture measures used in the classification of VHR images over urban areas, particularly when pixel-based approaches are utilized. Several recent publications have benefited from the GLCM for the purpose of mapping urban areas using VHR imagery. For instance, Maillard (2003) concluded that GLCM gives superior results over the semivariogram and the Fourier-based texture extraction for the scenes where objects are distinguishable visually by their textures' characteristics. Buildings and traffic areas are well-textured classes in VHR images and they can be easily distinguished by visual interpretation. Consequently, the GLCM works better than the two others in this case. Among the 14 GLCM texture measures, originally proposed by Haralick (1979), some of them are strongly correlated with each other (Cossu, 1988). Thus, the choice of optimal texture measure is an important issue in GLCM texture extraction (Jensen, 2005). Maillard (2003) reported that the most commonly used GLCM texture features in literature are, in decreasing order of popularity, the angular second moment (ASM), entropy (ENT), the inertia (initially contrast (CON)), the correlation (COR) and the inverse difference moment (IDM). Pacifici et al., (2009) believe that energy (ENR) (which is the square root of ASM), CON, variance (VAR), COR, ENT and IDM are the most relevant measures used in literature. Based on the prototype performance approach and its application in urban areas (Pratt, 2007), Puissant et al., (2005) and Su et al., (2008) concluded that four GLCM texture features used in mapping urban areas are homogeneity (HOM), ENT, dissimilarity (DIS) and the ASM. We conducted a broad search on the major publications (mostly peer reviewed journals) in the area of urban land cover classification using VHR imagery to find which GLCM measures have mostly been utilized. The results are summarized in Table 2. Based on the number of papers listed in Table 2 for each measure, a graph was plotted and is shown in Figure 1. ENT and ASM (or ENG) are the most frequently used measures. Nineteen and sixteen papers have utilized ENT and ASM respectively. HOM and CON ranked third (figure 1). These two measures have been utilized in thirteen papers. Having determined the appropriate GLCM measure(s), four important factors including window size and shape, quantization level, inter-pixel distance and direction of spatial relationship, must be defined for each measure (Su et al., 2008), since they influence the effectiveness of extracted measures in the classification process. The success of classification using texture features depends largely on the selected window size (Su et al. 2008). The window size and shape are automatically identified in object-based texture extraction inasmuch as the measure is calculated for each individual object resulting from the segmentation stage. In pixel-based approaches, however, the optimum of window size and three other factors must be determined. Window size is related to image resolution and content. Puissant et al., (2005) believe that it would be interesting to choose different window sizes according to the size of features to be extracted. For the choice of direction, the literature proposes calculation of texture measures in four directions (0, 45, 90 and 135 degrees) and then taking the average of them (Haralick, 1979 and Anys et al., 1994). Several researchers suggest that it is a good idea to reduce the quantization level (e.g. from 8 bits to 5 bits) of the input data so that the GLCM to be computed for each pixel does not become too large (Jensen, 2005). Table 3 shows different texture measures with their corresponding texture factors used in some recent papers. Because of space limitations, we listed the results four journal papers in which different VHR imagery, different urban land cover types, and different classification methods are utilized.

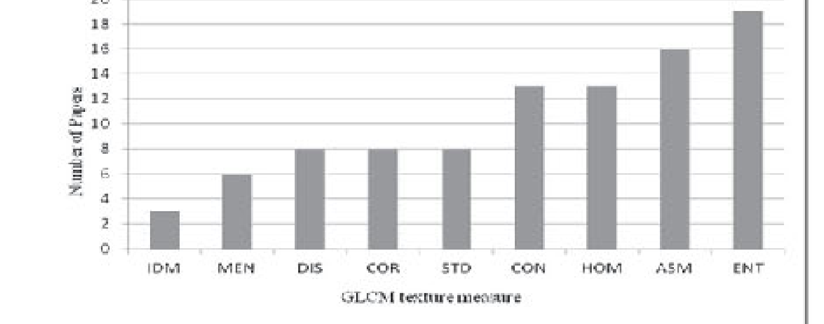

Figure 1: The most used GLCM texture measures for urban land cover classification of VHR imagery

Table 2: List of papers which have utilized different GLCM textural measures in urban land cover classification of VHR imagery

GLCM	Paper
HOM	Zhang (1999); Herold et al. (2003a); Puissant et al. (2005); Carleer and Wolf (2006); Agüera et al. (2008); Su et al. (2008); Pacifici et al. (2009); Lu et al. (2010); Luo and Mountrakis (2010); Ouma et al. (2010); Tassetti et al. (2010); Pu et al. (2011); Salehi et al. (2011a).
ASM (ENG)	Zhang (1999); Pesaresi (2000); Herold et al. (2003a); Maillard (2003); Myint et al. (2004); Puissant et al. (2005); Carleer and Wolf (2006); Myint (2007); Agüera et al. (2008); Su et al. (2008); Pacifici et al. (2009); Luo and Mountrakis (2010); Lu et al. (2010); Ouma et al. (2010); Pu et al. (2011); Salehi et al. (2011a).
ENT	Zhang (1999); Pesaresi (2000); Herold et al. (2003a); Maillard (2003); Shackelford and Davis (2003a); Shackelford and Davis (2003b); Myint et al. (2004); Puissant et al. (2005); Carleer and Wolf (2006); Alonso et al. (2007); Myint (2007); Agüera et al. (2008); Su et al. (2008); Chan et al. (2009); Pacifici et al. (2009); Ouma et al. (2010); Tassetti et al. (2010); Pu et al. (2011); Salehi et al. (2011a).
CON (Inertia)	Zhang (1999); Pesaresi (2000); Herold et al. (2003a); Maillard (2003); Carleer and Wolf (2006); Myint (2007); Agüera et al. (2008); Su et al. (2008);Pacifici et al. (2009); Luo and Mountrakis (2010); Ouma et al. (2010); Pu et al. (2011); Salehi et al. (2011a).
COR	Maillard (2003); Alonso et al. (2007); Myint (2007); Agüera et al. (2008); Pacifici et al. (2009); Ouma et al. (2010); Pu et al. (2011); Salehi et al. (2011a).
IDM	Pesaresi (2000); Maillard (2003); Myint (2007).
VAR (STD*)	Herold et al. (2003a); Shackelford and Davis (2003b); Agüera et al. (2008); Chan et al. (2009); Ouma et al. (2010); Tassetti et al. (2010); Pu et al. (2011); Salehi et al. (2011a).
DIS	Herold et al. (2003a); Puissant et al. (2005); Carleer and Wolf (2006); Agüera et al. (2008); Pacifici et al. (2009); Lu et al. (2010); Ouma et al. (2010); Pu et al. (2011).
MEN†	De Martino et al. (2003); Agüera et al. (2008); Chan et al. (2009); ; Lu et al. (2010); Ouma et al. (2010); Tassetti et al. (2010).

*STD: Standard Deviation ; † MEN:Mean

Table 3: Examples of using textural measures (with their corresponding optimal factors) for improving the urban land cover classification of VHR imagery. The textural measures are incorporated as additional band into the classification. Accuracy improvement (Imp) is over the case where only spectral band(s) of the image is utilized in the classification process.

Input data	Land cover types	Texture measures	Optimal Texture factors	Classification approach	Accuracy	Paper
VHR – Simulated image (1 pan and 3 MS bands at 1m Spatial res.)	Water ,shadow, tree, grass, road, built-up	HOM of pan band	W.S: :7x7 Dir.: average of (0,45,90, 135) LP: 1 ; Q.L: NM	Pixel-based Discriminant Analysis	OA:92.2% Imp:4.4%	Puissant *et al.* (2005)
VHR-QB (Quickbird) (Pansharp bands)	Shrub, grassland, water, road, building, vacant land, shadow	ASM over objects from segmented image	Dir.: 45 Q.L:NM, LP:NM	Object-based Maximum Likelihood	OA:83.7% Imp:2.1%	Su *et al.* (2008)
		CON over image*	W.S:7x7 Q.L:NM, LP:NM		OA:87.3% Imp:5.7%	
VHR-IK (Ikonos) (Pansharp bands)	Road, building, tree, water, grass, shadow, bare soil	1st order statistic (ENT) of Pan band	W.S: 20x20 Dir., LP, Q.L: Not Applicable	Pixel-based Hierarchical Fuzzy	Road-Buil.:73% Imp :1.5% Grass-Tree: 97.3% Imp :11%	Shackelford and Davis (2003b)
VHR-IK (Pan & MS)	Vegetation, soil, asphalt, metallic roof, shadow	ENT and COR of Pan band	W.S:NM Dir. PC1 of 8 PCs for (0, 45, 90,135) Q.L:NM, LP:NM	Pixel-based Maximum Likelihood	OA: 82.7% Imp: 3.4%	Alonso *et al.* (2007)

First, the texture measure was calculated for four MS bands resulting four textural bands, then Principal Component Analysis (PCA) was applied to four textural bands and the 1st PC was chosen as the final texture measure.

Table 4: List of papers which have utilized contextual measures in pixel-based or object-based urban land cover classification of VHR imagery

Pixel-based classification	Object-based classification
Gurney and Townshed (1983); Binaghi *et al.* (2003); Melgani and Sebastiano (2003); Shakelford and Davis (2003a); Shakelford and Davis (2003b); Khedam and Belhadj (2004); Jin and Davis (2005); ; Bellens *et al.* (2008a); Miller *et al.* (2009).	Herold *et al.* (2003a); Shakelford and Davis (2003a); Thomas *et al.* (2003); Nghi and Mai (2008) ; Chan et al. (2009); Hermosilla *et al.* (2011) ;

From Table 1, it can be seen that the texture measures and their optimal factors differ from each other for different input data and different land cover types. Furthermore, the improvement in OA, as the result of incorporating texture into classification, ranges from 2% to 11% for papers listed in this table.

2.3.2 Contextual measures

Whereas texture is the spatial variation within a small group of pixels, the context of a pixel (or a group of pixels) refers to its spatial relationship with the local and global configuration of neighbouring pixels (Gurney and Townshedn, 1983 and Binaghi et al., 2003). In VHR imagery, adjacent pixels are related or correlated (Khedama and Belhadj-Aissaa, 2004). The spatial correlation or dependency arises due to the fact that spatial resolution of the sensor is finer than the size of objects being classified (Shekhar et al., 2002). Information from neighbouring pixels (contextual information) plays an important role in the identification and extraction of urban features from remote sensing imagery (Jin and Davis, 2005) by increasing the discrimination capabilities of spectral pixel-based measured data. Moreover, in object-based approaches, when multi-scale segmentation is utilized, over-segmentation occurs in the lower levels of segmentation, resulting in too many boundaries such that real objects such as buildings and roads are split into two or more smaller objects. Hence, the spectral information of the object of interest is correlated with that of adjacent objects in the same level or with the spectral information of the super objects from the upper level. This contextual information can be incorporated into classification compensating for the spectral confusion between spectrally similar objects. To date, little research has been conducted on incorporating contextual information in classification and object extraction of VHR imagery (compared to textural and morphological information). Table 4 lists the papers which have utilized contextual measures in the pixel-based or object-based classification of VHR imagery of urban areas. Shadows of high-rise features such as buildings and trees are among the widely used contextual measures. Several studies have utilized it in the pixel-based (e.g. Jin and Davis, 2005 and Bellens et al., 2008b) or in the object-based (e.g. Shackelford and Davis, 2003a and Nghi and Mai, 2008) classification processes in order to separate buildings from roads and streets in urban areas. Structural measures such as the length and width of a connected group of spectrally similar pixels are considered as contextual measures in pixel-based image analysis methods, while they are categorized as morphological measures in object-based methods. In the former case, these measures are calculated based on the spatial relationship between neighbouring pixels whereas in the latter case they are directly related to shape and size of each individual object. For example, Shackelford and Davis (2003b) used the length and width of a connected group of pixels, calculated based on spatial relationships within that group of pixels as two additional contextual bands in the classification. The length and width bands have high value for road and building classes, respectively. Nghi and Mai (2008) utilized contextual relation in object-based classification. The contextual relation of an object is the number of objects that are adjacent with the object of interest. For instance, road objects have stronger relation than building objects. Thus, the relation can be used as additional information in the object-based classification to separate roads from buildings or shadows from large water bodies. In object-based approaches, the difference between the mean brightness value of the object of interest to its neighbour objects and to super objects for different bands are other types of contextual measures (Thomas et al., 2003). Table 5 presents the results of some literature that have used contextual information as additional bands in the urban land cover classification process of VHR imagery.

2.3.3 Morphological measures

The effectiveness of the aforementioned spatial measures, particularly texture measures, is highly dependent on the choice of optimal window size in pixel-based approaches and the level of segmentation in object-based ones. The optimal window size differs for different classes in an urban scene, but defining an optimal window size or segmentation level for extracting textural and contextual measures is not an easy task. Moreover, members of same class can have different spectral reflectance values, such as black and gray building roofs (Miller et al., 2009). Consequently, they may have different textural and contextual measures. On the other hand, buildings, roads and parking lots possess specific morphological characteristics such as shape (smoothness and compactness) and size (length, width, perimeter, area, etc), especially in VHR images. A list of morphological measures can be found in the literature (e.g. Pratt, 2007 and Definiens Developer, 2007). Incorporation of these morphological features into classification compensates for the lack of spectral information of VHR images and facilitates the discrimination process of spectrally similar classes. Subsequently, the classification accuracy of such classes increases. Incorporating morphological measures in the pixel-based classification is usually done by applying a morphological filter to the image. Having applied the filter to the image, the spatial form or structure of objects within the image is modified. This modified image, then, is used as an additional band to the original bands of the image in classification. Dilation, erosion and skeletonization are three fundamental morphological operations (Pratt, 2007). However, opening and closing, which are composed by combination of erosion and dilation operations, are the most common morphological operators used in literature in order to modify the form and structure of the objects within the image (Shackelford and Davis, 2003b). Morphological measures are more meaningful and applicable in object-based classification approaches (Bellens et al., 2008b) and many researchers have benefited from their use. Table 6 lists the papers in which morphological measures have been utilized in pixel-based or object-based classification of VHR imagery of urban areas. In object-based approaches the shape and size, as morphological measures of segmented regions, are directly utilized in classification. Table 7 as well as Table 5 report some examples of the use of morphological measures along with the classification approach and their corresponding accuracies in the urban land cover classification of VHR imagery.

Table 5: Examples of using spatial measures (mainly contextual measures) for improving the urban land cover classification of VHR imagery. The spatial measures are incorporated as additional band into the classification. Accuracy improvement (Imp) is over the case where only spectral band(s) of the image is utilized in the classification process.

Input data	Land cover types	Spatial measures	Classification approach	Accuracy	Paper
VHR-IK (Pansharp bands)	Road, building, tree, water, grass, shadow, bare soil	Texture: 1st order ENT for grass-tree Context: Length-width of a connected group of pixels for road-building and water-shadow	Pixel-based hierarchical fuzzy classification	OA: 92.7%,Imp:11% Road: PA:88%,Imp:17% Building: PA:84%, Imp:11%	Shackelford and Davis (2003b)
VHR-IK (Pansharp bands)	Road, building, tree, water, grass, shadow, bare soil, impervious surf.	Context: Shadow of the objects Morphology: Skeleton of the objects	Pixel/Object-based Fuzzy classification	Building: PA:76%, Imp:3%, Road: PA:99%Imp:28%	Shackelford and Davis (2003a)
VHR-QB	Water, grass, tree, roof(dark, red, white) road, other man made, shadow	Context: Shadow Proximity Feature(SPF), Shadow Distance Feature(SDF)	Pixel-based Maximum Likelihood	OA: 72% Imp:4% In the case where only one content, either SPF or SDF is used	Bellens *et al.* (2008a)
VHR-IK (Pan band)	Agricultural land, road networks, industrial plants, quarries, urban components (large buildings , small houses)	Context: A Multi-window set (6x6,20x20,36x36 pixels) of image data was used directly as the input of contextual information for MLP classifier	Pixel-based Multi Layer Perceptron (MLP) Neural Network	Correlation coefficient[†]:0.77 Imp: 0.28 Standard error:25.1 Imp: 28.8	Binaghi *et al.* (2003)

[*]Feature Analyst software was released by Visual Learning System (VLS) in 2001and later on was developed as an extension for ESRI's Arc GIS and ERDAS Imaging (Miller et al. 2009).

[†]Accuracy assessments are based on regression analysis used for soft classification. Improvement is over the case when only a single window of 3x3, as oppose to multi-window set, was used with the same classification strategy.

Table 6: List of papers which have utilized morphological measures in pixel-based or object-based urban land cover classification of VHR imagery

Pixel-based classification	Object-based classification
Benediktsson *et al.* (2003); Cablk and Minor (2003); Jin and Davis (2005); Mena and Maplica (2005); Bruzzone and Carlin (2006); Inglada (2007); Bellens *et al.* (2008b); Fauvel *et al.* (2008); Tuia *et al.* (2009); *Chan et al* (2009); Lhomme *et al.* (2009).	Bauer and Steinnocher (2001);Herold *et al.* (2003a); Shakelford and Davis (2003a); Thomas *et al.* (2003); Carleer and Wolf (2006); Nghi and Mai (2008); Walker and Blashke (2008); van der Werff and van der Meer (2008); Chen *et al.* (2009a); Zhou *et al.* (2009); Xu and Li (2010); Hermosilla *et al.* (2011); Pu *et al.* (2011); Salehi *et al.* (2011b); Salehi *et al.* (2011c).

Table 7: Examples of using morphological measures (with their corresponding optimal factors) for improving urban land cover classification of VHR imagery. The morphological measures are incorporated as additional band into the classification. Accuracy improvement (Imp) is over the case where only spectral band(s) of the image is utilized in the classification process.

Input data	Land cover types	Morphological measures	Classification approach	Accuracy	Paper
VHR-QB (Full band)	Building and Transportation area	Length, width, length/width, area excluding inner regions, area including inner regions, perimeter, compactness	Object-based Nearest Neighbour	KC: Urban area:0.51 Suburban area:0.71	Carleer and Wolff (2006)
VHR- QB (Pan band)	Building, apartment blocks, road, railway, veg. ,trees, bare soil, soil, tower	Morphological operators: Opening, closing, reconstructed opening, reconstructed closing, with SE of the size of 9-25 pixels	Pixel-based SVM	OA:86.5%,Imp:52.5% Build.:89%,Imp:62% Road:89%,Imp:35%	Tuia *et al.* (2009)
VHR-QB (Pan and Pansharp)	Gray roof, Red roof, Road, Shadow, Rural area , Grass, Tree	From the 6th level of segmented image: Width/length ratio, Shape index, Rectangular fit	Pixel-based SVM classification of segmented image	OA:92.6% Imp:4%	Bruzzone and Carlin (2006)
VHR-IK (Full band)	Water ,grass, tree, buildings (dark, red and white roof), road, shadow, other man-made objects	Geometric Activity(GA) features including Ridge features based on facet model and Morphological features(closing)	Pixel-based Multi Layer Perception Neural Network	OA:71%,Imp[*]:5.4% Average PA of Man-made objects:72.2%, Imp[*]:5.3%	Chan *et al.* (2009)
VHR-QB (Full band)	Bare soil ,grass, tree, roof, road	Morphological operators: Eight disk shaped SEs and eight line shaped SEs morphological profiles(MP) with partial reconstruction	Pixel-based Multilayer Perceptions Neural Network	OA:89%, Imp[†]:9% Road:72%, Imp:18.4% Roof:93%, Imp:17.6%	Bellens *et al.* (2008b)
VHR-QB (Pansharp bands)	Building, road, tree, grass, soil, shadow, other impervious surface	7 Hu's, invariant moments 10 Zernike invariant moments 17 Wavelet invariant moments	Object-based SVM	OA: 80.5%, Imp[‡]:6.8%, Build:74.6%,Imp:6.4% OA: 79.5%, Imp:5.8% Build:70.2%,Imp:2% OA: 80.2%, Imp: 6.5% Build: 70.2%,Imp:2%	Xu and Li (2010)

[*] Improvement is over pixel-based MLP neural network when only spectral information (i.e. 4 MS bands + Pan and NDVI bands) is used
[†] Improvement is over the same pixel-based neural network when only 1st component of 40 principal components of the original data is used
[‡] Improvement is over the same object-based SVM when only spectral information (four Pan-sharp bands) is used.

2.4 Spatial Measures Extracted from Ancillary Data

Spatial information can be derived from the image itself (e.g. texture, context and morphology), which was broadly discussed in previous sections and/or from other data sources, the so-called ancillary data. Ancillary data layers are key components of accurate image classification (Thomas et al., 2003). Particularly, with the widespread availability of VHR imagery, digital elevation/surface model (DEM/DSM) extracted from LiDAR data or stereo images and existing GIS data layers, the importance of integrating these data for detailed mapping of urban environments becomes highly significant. Heinl et al., (2009) showed that the use of ancillary data improves the classification accuracy independent of classification method. When ancillary data are utilized along with the image in the classification process, usually the misregistration between the ancillary data and the image is a problematic issue.

Precise geometric registration of corresponding data layers is often very difficult to achieve, particularly for VHR imagery. Since the basic mapping unit in object-based approaches is a group of connected pixels (i.e. object instead of pixel), the misregistration between multisource data (i.e. VHR image, LiDAR and GIS data layers) is not as serious as for pixel-based approaches (Justice et al., 1989 and Zhou et al., 2008). In fact, object-based approaches facilitate the use of ancillary data (Kim et al., 2010) and since they require less precise registration of the image, they are highly desirable for multisource image analysis (Kim et al., 2010*)*. Several methods may be used to incorporate ancillary data into the classification process. The most common used is the stacked vector or the logical channel method (Jensen, 2005, Watanachaturaporn, et al., 2008 and Huang et al., 2011), which considers the ancillary data as an extra channel (band) to the original channels of the image in pixel-based or object-based classification. In object-based approaches, in addition to the stacked layer method, ancillary data are also used in a rule-based manner during segmentation (e.g. Bouziani et al., 2010) and classification. Two groups of ancillary data are frequently used: DSM produced from LiDAR (LiDAR-derived DSM) data or from aerial or satellite stereo images (photogrammetrically-derived DSM) and GIS data layers such as the map of parcels, the road centreline network, etc. Table 8 lists the papers in which ancillary data have been incorporated into classification (both pixel-based and object-based) of VHR imagery of urban areas. As seen in this table, object-based approaches have utilized ancillary data significantly more than have pixel-based approaches.

2.4.1 DSM Derived from LiDAR and stereo photographs

DSM generated by LiDAR data or by stereo aerial/satellite images (photogrammetrically-derived DSM) gives information about the height of objects (e.g. buildings and trees); thus it is very helpful in separating spectrally similar objects with different heights (e.g. buildings from streets and tree from shadows). LiDAR- and photogrammetrically-derived DSM represent the height of each pixel with respect to the reference datum. Consequently, when they are intent to be incorporated in the classification process, the first stage involves removal of the underlying terrain, so called digital elevation model (DEM), from the DSM to produce the heights of objects above the local ground (e.g. building's height) Ma, 2005).The resultant height is referred to as normalized DSM (nDSM) (Chen et al., 2009b). The nDSM could then be integrated along with the spectral bands of the image as an additional channel in both segmentation and classification processes or as additional information in rule-based segmentation and classification. The influence of DSM on classification can be controlled by adjusting channels' weights in the former method (Hofmann, 2001). There is a considerable amount of research accomplished in the integration of LiDAR DSM with the image during classification, particularly in object-based analysis of VHR images over urban areas (Table 8). Regardless of the classification approach, results show a significant improvement of classification accuracy (e.g., Sohn and Dowman, 2007, Watanachaturaporn et al., 2008 and Huang et al., 2011).Very few studies, however, have benefited from the incorporation of the photogrammetrically-derived DSM and VHR imagery for classification over urban areas. This is mainly due to unavailability of precise DSM of urban areas and misregistration between the DSM and VHR imagery. However, recent development of stereo satellite and aerial VHR imagery in conjunction with advancements in object-based image analysis methods has facilitated the integration of the photogrammetrically-derived DSM and VHR data. Hussain and Shan (2010) integrated photogrammetrically-derived nDSM and VHR imagery in a rule-based object-based classification method in order to separate buildings from transportation areas and the result was very satisfactory (Hussain and Shan, 2010). They further employed the nDSM to separate different buildings according to their heights (i.e single and double story houses and apartments). Table 9 presents a summary of some recent papers in which ancillary data (mainly LiDAR nDSM) have been utilized together with spectral bands of the VHR imagery for object-based and pixel-based classification.

2.4.2 GIS data layers

Despite the growing attention toward use of LiDAR data for classification, few studies have taken the advantages of existing GIS data layers for improving urban land cover classification (Table 8). Thomas et al., (2003) developed strategies including spatial modelling techniques to deal with the problem of confusion between spectrally similar classes in VHR data. Their spatial modelling was based on the integration of GIS ancillary data layers with the image bands. They used the distance from road centrelines to differentiate buildings and parking lots from roads Zhou and Troy (2008) employed building footprint data in a rule-based hierarchical object-oriented classification to separate buildings from non-building objects. They showed that the classification accuracy increased when building footprint data together with LiDAR DSM are incorporated into the classification process.

3. Discussion

3.1 Spectral Measures

The result of this review shows that spectral information of VHR imagery plays a major role in classifying vegetation (including grass and trees), water surfaces and even shadow areas. In particular, pixel-based approaches give superior results over object-based classification for mapping such land cover types in a typical urban environment. However, impervious land cover types (e.g. buildings, roads and parking areas) are spectrally too similar to be separated only according to their spectral information in VHR imagery. Between pixel-based and object-based classification approaches, a serious problem common to both approaches arises in the classification of spectrally similar and heterogeneous classes.

Table 9: Examples of using ancillary data (mainly LiDAR nDSM) in addition to VHR imagery for improving urban land cover classification. The ancillary measures are incorporated as additional band into the classification and/or segmentation. Accuracy improvement (Imp) is over the case where only spectral band(s) of the image is utilized in the classification process.

Input data	Ancillary data used	Land cover types	Classification approach	Accuracy	Paper
VHR-QB Pansharp bands	LiDAR nDSM: HA*=30cm, VA†=15cm SR=1m	Water, shadow, grass, shrub, building, road, vacant areas	Hierarchical object-based	OA:89.4%,Imp:20% Building:PA:92.5%, Imp:9% Road:PA:86%, Imp,47.5%	Chen *et al.* (2009b)
VHR-Aerial image (0.6m SR in G,R, NIR bands)	LiDAR nDSM: First and last pulse data, SR=1m, GIS data: Parcel boundary, Building footprints	Building, pavement, coarse texture vegetation, fine texture vegetation, bare soil	Hierarchical rule-based object-oriented	OA: 92.3% Building PA:94.4% Pavement PA: 88.3% Imp: NM	Zhou and Tro (2008)
LiDAR intensity	LiDAR nDSM, DSM, Echo code (Multiple return)	Water, low veg., road structure, deciduous, coniferous, intertidal	Hierarchical rule-based object-oriented	OA: 98.1%, Structure PA:94.2% Road PA:94.5% Imp: NM	Brennan and Webster (2006)
VHR-Aerial orthophoto (0.4 m in R,G,B bands)	LiDAR DSM HA: 0.5 m, VA: 0. 15 m A Max-Min band resulted from moving a window of 13x13 size over LiDAR DSM	Ground, grass, shadow, building, tree	Pixel-based SVM	OA: 94.7%, Imp‡: 12.2% Build: 96.4% Imp: 29.6%	Huang *et al.* (2011)

* HA: Horizontal Accuracy; † VA: Vertical Accuracy

‡ Improvement is over pixel-based SVM when only spectral information (3 multispectral bands) is used.

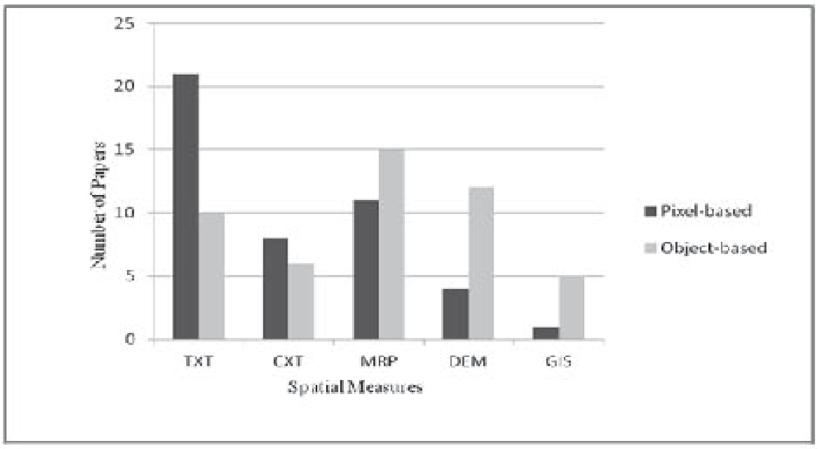

Figure 2: The number of papers which have utilized spatial measures (TXT: texture, CXT: context, MRP: morphology, DEM, height data, and GIS: GIS data) in pixel-based and object-based classification of VHR imagery over urban areas.

3.2 Spatial Measures

3.2.1 Pixel-based vs. object-based classification

Figure 2 presents the number of papers in which spatial measures have been utilized in pixel-based or object-based classification approaches. This figure is a summary of Tables 1, 4, 6, and 8. Texture features have been utilized in pixel-based approaches significantly more than in object-based approaches. On the other hand, ancillary data, especially LiDAR DEM, have been used in object-based classification more often than in pixel-based classification.

The number of papers that have used morphological measures is almost equal in both pixel-based and object-based classification (with two papers more in object-based).Ten papers have used contextual measure in pixel-based classification and seven papers used them in object-based approaches. Compared to the other measures, GIS data have been employed in a very few papers (five object-based and one pixel-based) indicating the very early stage of combining GIS data and VHR imagery for classification purposes.

3.2.2 Textural measures

From the accuracies reported in the literature, texture has a significant effect on improving the classification accuracy of urban areas using VHR imagery. The improvement ranges from 2% to11% in terms of overall accuracy. Some of the major findings on the performance of texture measures are as follows:

- Texture measure, in general, has better performance in pixel-based than object-based approaches.
- Texture does not necessarily increase the classification accuracy for all classes. It is desirable to incorporate texture measures in classification using a fuzzy-based classification with different membership for different sets of classes.
- For urban impervious surfaces, integrating texture increases the classification accuracy, whereas for vegetation and water, texture does not have significant effect or even lowers the classification accuracy.
- In general larger window sizes of texture work better in classifying heterogeneous land cover types for VHR imagery, whereas smaller window size is preferred for lower resolution imagery.
- In object-based image analysis, when multilevel segmentation is used, texture measures of objects at higher levels of segmentation would be more appropriate for classification of heterogeneous land cover types such as impervious surfaces, while for homogeneous classes, texture measures of objects at lower levels of segmentation are preferred.

3.2.3Contextual measures

The amount of literature in which contextual information is utilized in the classification process is relatively small compared to that using texture and morphology (figure 2). Shadow is the most used contextual measure in classification, especially in pixel-based methods. The spectral and spatial relation between objects in the same or different levels of segmentation is the major source of contextual measures in object-based methods. Furthermore, contextual measures are rarely used as the only spatial measure in classification. In other words, contextual measures are often used in conjunction with textural and/or morphological measures during classification. Nevertheless, in all cases where contextual measures have been incorporated, classification accuracies have been increased.

3.2.4 Morphological measures

The role of morphological information in object-based classification is more significant than in pixel-based. Indeed, a key advantage of object-based over pixel-based image analysis is that the results of segmentation in object-based approaches are sets of meaningful regions for which a number of morphological features can be directly measured and used during the classification process. Moreover, some segmentation algorithms are able to create objects in different scales with different sizes and shapes. This multi scale or multi level image segmentation allows the classifier to utilize morphological properties of objects in various scales, resulting in higher classification accuracy for classes such as roads and buildings, which present in different sizes and shapes. The quantitative results of some studies showed an average accuracy of around 90% in terms of producer's accuracy (PA) for buildings and roads in a typical urban environment, when both spectral and morphological measures of VHR imagery are utilized in object-based classification.

3.2.5 DSM derived from LiDAR and stereo photographs

Recent developments in object-based image analysis and increasingly available LiDAR data and VHR imagery have directed researchers' attention toward their integration for classification purposes. Nevertheless, we learned that the use of LiDAR data along with VHR imagery for detailed land cover classification is still at an early stage of development, although they have exhibited potential in urban land cover mapping. Many of the published papers in recent years are conference papers and the results show that the incorporation of LiDAR data into the classification of VHR imagery can significantly solve the problem of differentiating high-rise and low-rise objects with similar spectral reflectance. For the case of integration of VHR imagery and photogrammetrically-derived DSM, very few papers have been found. With the availability of stereo VHR imagery from satellites such as QuickBird, GeoEye-1 and WorldView-2, the advances in object-based image analysis and precise DSM generation methods, the integration of VHR imagery and its photogrammetrically-derived DSM for mapping of complex urban landscapes is more feasible than before. It is worth mentioning that it is important to avoid misregistration problems between DSMs and VHR images. Almost all papers avoided the use of off-nadir VHR images in the classification. In the real world, however, more than 90% of VHR images are collected off nadir.

3.2.6 GIS data layers

Traditionally, existing GIS map layers such as road/street networks and building footprints have been used as reference data to evaluate the performance of classification. Recently, the potential of utilizing this vector data for improving classification accuracy has drawn increased attention from many researchers. Because GIS data layers are somewhat consistent with object-based classification input, i.e. object resulting from segmentation may share the same boundaries with GIS data, these data have been increasingly used to improve object-based classification. Nonetheless, the use of GIS data in classification, in the literature, is far less than that of other spatial measures (figure 2). Misregistration between GIS vectors and images still poses a great challenge for integration. Most papers used either low resolution images or nadir VHR images in the classification to avoid the errors introduced by misregistration.

4. Conclusion

Over the past decade, there have been ever-growing numbers of researchers studying on detailed land cover classification of urban areas. This is partly due to the fact that the differentiation amongst three major impervious classes in an urban scene (roads, buildings and parking lots) becomes more feasible with the availability of VHR imagery. Due to the complex nature of urban landscapes as well as the spatial and spectral characteristics of VHR imagery, the classification of such landscapes requires not only spectral but also spatial information of the image. For this reason, the amount of literature using spatial information includes texture, context, morphology and information extracted from ancillary data such as DSM and archived GIS layers has grown since the launch of first VHR satellite in 1999. Although spatial measures have been used in both pixel-based and object-based classification approaches, the employment of them in the latter case is more effective mainly due the following reasons: 1) Determining the optimal window size of spatial measures, which is a critical issue in pixel-based classification methods, has been solved in object-based approaches by segmenting the image to individual objects with meaningful boundaries. 2) Contextual and especially morphological measures are more meaningful in the object-based image analysis. 3) Object-based approaches require less precise geometric registration between different data layers when ancillary data are employed in the classification.

However, the major drawback of object-based image classification is the lack of a high degree of automation. In fact, different rule sets must be developed for different image and applications. Among the three spatial measures extracted from the images, texture has most often employed in classification, especially in pixel-based approaches. Multiresolution segmentation of object-based approaches, on the other hand, enhances the capability of morphological measures to improve the classification of land covers such as buildings and traffic areas. The classification accuracies reported in literature indicate that morphological measures have significantly higher effect on differentiating classes such as buildings and roads than textural and contextual measures. The use of contextual measures in classification of VHR imagery is less than the use of textural and morphological measures. In addition, contextual measures have rarely been used as the only source of spatial measure in classification. Indeed, the effectiveness of such measures is enhanced when they are used in conjunction with other spatial measures (textural and/or morphological measures). The employment of LiDAR- and especially photogrammetrically-derived DSM in urban land cover classification of VHR imagery is in the very early stages perhaps because of non availability of LiDAR data and the rather low precision of photogrammetrically-derived DSM. Nonetheless, the results of our literature review reveal the very high potential of this type of data in conjunction with VHR imagery for land cover mapping of urban environments. With the widespread availability of VHR stereo aerial/satellite imagery and the development in object-based image analysis and precise DSM generation methods, the use of photogrammetrically-derived DSM in classification of VHR imagery over urban areas is more feasible than before and has a very high potential for future research. Also despite the availability of archived GIS data layers, this type of ancillary data has not been well utilized for classification purposes. The development of spatial modelling of available GIS data layers for incorporation into the classification process, particularly object-based classification approaches, would be another interesting topic for future research.

Acknowledgements

This research is supported by the NSERC (Natural Science and Engineering Research Council of Canada) Strategic Project Grants and NSERC Discovery Grants Programs.

References

Aksoy, S., Koperski, K., Tusk, C. and Marchisio, G., 2009, Land Cover Classification with Multi-Sensor Fusion of Partly Missing Data. *Photogrammetric Engineering and Remote Sensing*, Vol. 75, No. 5, 577-598.

Agüera, F., Aguilar, F. J. and Aguilar, M. A., 2008, Using Texture Analysis to Improve Per-Pixel Classification of Very High Resolution Images for Mapping Plastic Greenhouses. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 63, No. 6, 635-646.

Alonso, M. C., Sanz, M. A. and Malpica, J. A., 2007, Classification of High Resolution Satellite Images using Texture from the Panchromatic Band. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 4842 LNCS, Part 2, 499-508.

Anys, H., Bannari, A., He, D. C. and Morin, D., 1994, Texture Analysis for the Mapping of Urban Areas using Airborne Meis-II images. *In Proceeding of the First International Airborne Remote Sensing Conference and Exhibition*: Applications, Technology, and Science, Strasbourg, France, Vol. 3, 231-245.

Aubrecht, C., Steinnocher, K., Hollaus, M. and Wagner, W., 2009, Integrating Earth Observation and Giscience for High Resolution Spatial and Functional Modeling of Urban Land Use. *Computers, Environment and Urban Systems*, Vol. 33, No. 1, 15–25.

Bauer, T. and Steinnocher, K., 2001, Per-parcel Land Use Classification in Urban Areas Applying a Rule-Based Technique. *Geo-Informations-Systeme*, Vol. 14, No. 6, 24-27.

Bellens, R., Douterloigne, K., Gautama, S. and Philips, W. 2008a. Per Pixel Contextual Information for Classification of VHR Images of Urban Areas. *In IEEE International Geoscience and Remote Sensing Symposium*, Boston.

Bellens, R., Gautama, S., Martinez-Fonte, L., Philips, W., Chan, J. C. and Canters, F., 2008b, Improved Classification of VHR Images of Urban Areas Using Directional Morphological Profiles. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No. 10, 2803-2813.

Ben-Dor, E., 2001, Imaging Spectroscopy for Urban Applications in Imaging Spectroscopy F. Van Der Meer and S.M. de Jong (Eds), 243-281(Kluwer Academic Press).

Benediktson, J. A., Pesaresi, M. and Arnason, K., 2003, Classification and Feature Extraction for Remote Sensing Images from Urban Areas Based on Morphological Transformations. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 9(Part1), 1940-1949.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I. and Heynen, M., 2004, Multi-Resolution, Object-Oriented Fuzzy Analysis of Remote Sensing Data for GIS-Ready Information. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 58, No. 3-4, 239-258.

Binaghi, E., Gallo, I. and Pepe, M., 2003, A Cognitive Pyramid for Contextual Classification of Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 4, No. 12, 2906-2922.

Blaschke, T. 2010. Object-Based Image Analysis for Remote Sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 65, No. 1, 2-16.

Blaschke, T., Lang, S. and Hay, G. J., 2008, Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications. Series: XVII, Lecture Notes in Geoinformation and Cartography (Berlin: Springer-Verlarge).

Bouziani, M., Goita, K. and He, D. C., 2010, Rule-Based Classification of a Very High Resolution Image in an Urban Environment using Multispectral Segmentation Guided by Cartographic Data. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, No. 8, 3198-3211.

Brennan, R., and Webster, T. L. 2006, Object-Oriented Land Cover Classification of Lidar-Derived Surfaces. *Canadian Journal of Remote Sensing*, Vol. 32, No. 32, 162-172.

Bruzzone, L. and Carlin, L., 2006, A Multilevel Context-Based System for Classification of Very High Spatial Resolution Images. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 44, No. 9, 2587-2600.

Cablk, M. E. and Minor, T. B., 2003, Detecting and Discriminating Impervious Cover with High-Resolution IKONOS Data using Principal Component Analysis and Morphological Operators. *International Journal of Remote Sensing*, Vol. 24, No. 23, 4627-4645.

Carleer, A. P. and Wolff, E., 2006, Urban Land Cover Multi-Level Region-Based Classification of VHR Data by Selecting Relevant Features. *International Journal of Remote Sensing*, Vol. 27, No. 6, 1035-1051.

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 Objects using Very High Resolution Satellite Imagery. *Photogrammetric Engineering and Remote Sensing*, Vol. 75, No. 4, 397-411.

Chen, D., Stow, D. A. and Gong, P., 2004, Examining the Effect of Spatial Resolution and Texture Window Size on Classification Accuracy: An Urban Environment Case. *International Journal of Remote Sensing*, Vol. 25, No. 11, 2177-2192.

Chen, M., Su, W., Li, L., Zhang, C., Yue, A. and Li, H., 2009a, Comparison of Pixel-Based and Object-Oriented Knowledge-Based Classification Methods using SPOT5 Imagery. *WSEAS Transactions on Information Science and Applications*, Vol. 6, No. 3, 477-489.

Chen, Y., Su, W., Li, J. and Sun, Z., 2009b, Hierarchical Object-Oriented Classification using Very High Resolution Imagery and LIDAR Data Over Urban Areas. *Advances in Space Research*, Vol. 43, No. 7, 1101-1110.

Congalton, R. G., 1991, A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing of Environment*, Vol. 37, No. 1, 35-46.

Cossu, R., 1988, Segmentation by Means of Textural Analysis. Pixel , Vol. 1, No. 2, 21-24.

Definiens, 2007, eCognition Professional User Guide. Avialable online at: http://www.definiens.com (accessed March 2011).

Fauvel, M., Bnediktsson, J. A., Chanussot, J. and Sveinsson, J. R., 2008, Spectral and Spatial Classification of Hyperspectral Data using SVMs and Morphological Profiles. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No. 11, 3804-3814.

Gamba, P., Dell'Acqua, F. and Dasarathy, B. V., 2005, Urban Remote Sensing using Multiple Data Sets: Past, Present and Future. *Information Fusion*, Vol. 6, No. 4, 319-326.

GEOBIA, 2008, The Proceedings of GEOgraphic Object Based Image Analysis for the 21 Century. Calgary, Canada. Available online at: http://homepages.ucalgary.ca/~gjhay/geobia/Proceedings/Sessions/ GEOBIA_Proceedings_RP10_ FINALc.pdf (accessed March 2011).

GEOBIA, 2010, The Proceedings of GEOgraphic Object Based Image Analysis. Ghent, Belgium 29 June-2July, 2010 Available online at: http://geobia.ugent.be/proceedings/html/papers.html (accessed March, 2011).

Gong, P., Marceau, D. J. and Howarth, P. J., 1992, A Comparison of Spatial Feature Extraction Algorithms for Land-Use Classification with SPOT HRV Data. *Remote Sensing of Environment*, Vol. 40, No. 2, 137-151.

Gurney, C. M. and Townshend, J. R. G., 1983, The use of Contextual Information in the Classification of Remotely Sensed Data. *Photogrammtric Engineering and Remote Sensing*, Vol. 49, No. 1, 55–64.

Haralick, R. M., 1979, Statistical and Structural Approaches to Texture. *In Proceedings of IEEE*, Vol. 67, No. 5, 786-804.

Heinl, M., Walde, J., Tappeiner, G. and Tappeiner, U., 2009, Classifiers vs. Input Variables-The Drivers in Image Classification for Land Cover Mapping. *International Journal of Applied Earth Observation and GeoInformation*, Vol. 11, No. 6, 423-430.

Herold, M., Liu, X. and Clarke, K. C., 2003a, Spatial Metrics and Image Texture for Mapping Urban Land Use. *Photogrammtric Engineering and Remote Sensing*, 69, 991–1001.

Herold, M., Gardner, M. E. and Roberts, D. A., 2003b, Spectral Resolution Requirements for Mapping Urban Areas. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No.9 (Part 1), 1907-1919.

Herold, M., Roberts, D. A., Gardner, M. E. and Dennison, P. E., 2004, Spectrometry for Urban Area Remote Sensing-Development and Analysis of a Spectral Library from 350 to 2400 nm. *Remote Sensing of Environment*, Vol. 91, No. 2/4, 304-319.

Hermosilla, T., Ruiz, L. A., Recio, J. A. and Estornell, J., 2011, Evaluation of Automatic Building Detection Approaches Combining High Resolution Images and LiDAR Data. *Remote Sensing*, Vol. 3, No. 6, 1188-1210.

Hester, D. B., Cakir, H. I., Nelson, S. A. C. and Khorram, S., 2008, Per-Pixel Classification of High Spatial Resolution Satellite Imagery for Urban Land-Cover Mapping. *Photogrammetric Engineering and Remote Sensing*, Vol. 74, No. 4, 463-471.

Hodgson, M. E., Jensen, J. R., Tullis, J. A., Riordan, K. D. and Archer, C. M., 2003, Synergistic use of Lidar and Color Aerial Photography for Mapping Urban Parcel Imperviousness. *Photogrammetric Engineering and Remote Sensing*, Vol. 69, No. 9, 973-980.

Hofmann, P., 2001, Detecting Buildings and Roads from IKONOS Data using Additional Elevation Information. *Geo-Informations-System*, Vo.14, No. 6, 28-33.

Huang, X., Zhang, L. and Gong, W., 2011, Information Fusion of Aerial Images and LIDAR Data in Urban Areas: Vector-Stacking, Re-Classification and Post-Processing Approaches. *International Journal of Remote Sensing*, Vol. 32, No. 1, 69-84.

Hussain, E. and Shan, J., 2010, Rule Inheritance in Object-Based Image Classification for Urban Land Cover Mapping. *In: Proceedings of ASPRS Annual Conference*, April 26-30, San Diego, California.

Inglada, J., 2007, Automatic Recognition of Man-Made Objects in High Resolution Optical Remote Sensing Images by SVM Classification of Geometric Image Features. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 62, No.3, 236-248.

Jensen, J. R., 2005, Introductory Digital Image Processing: A Remote Sensing Perspective, 3rd ed. (Upper Saddle River, N.J.: Prentice Hall).

Jin, X. and Davis, C. H., 2005, Automated Building Extraction from High-Resolution Satellite Imagery in Urban Areas using Structural, Contextual and Spectral Information. *Eurasip Journal on Applied Signal Processing*, Vol. 14, 2196-2206.

Justice, C. O., Markham, B. L., Townshend, J. R. G. and Kennard, R. L. 1989, Spatial Degradation of Satellite Data. *International Journal of Remote Sensing*, Vol. 10, No. 9, 1539-1561.

Khedama, R. and Belhadj-Aissaa, A., 2004, Contextual Classification of Remotely Sensed Data using MAP Approach and MRF. *In Proceeding of ISPRS 2004 Congress*, Istanbul, Turkey.

Kim, M., Madden, M. and Xu, B., 2010, GEOBIA Vegetation Mapping In General Smoky Mountains National Park with Spectral and Non-Spectral Ancillary Information. *Photogrammetric Engineering and Remote Sensing*, Vol. 76, No. 2, 137-148.

Liu, Y., Li, M., Mao, L., Xu, F. and Huang, S., 2006, Review of Remotely Sensed Imagery Classification Patterns Based on Object-Oriented Image Analysis. *Chinese Geographical Science*, Vol. 16, No. 3, 282-288.

Lhomme, S., He, D. C., Weber, C. and Morin, D., 2009, A New Approach to Building Identification from Very High- Spatial-Resolution Images. *International Journal of Remote Sensing*, Vol. 30, No. 5, 1341-1354.

Lu, D., Hetrick, S. and Moran, E., 2010, Land Cover Classification in a Complex Urban-Rural Landscape with Quickbird Imagery. *Photogrammetric Engineering and Remote Sensing*, Vol. 76, No. 10, 1159-1168.

Lu, D. and Weng, Q., 2007, A Survey of Image Classification Methods and Techniques for Improving Classification Performance. *International Journal of Remote Sensing*, Vol. 28, No. 5, 823-870.

Luo, L. and Mountrakis, G., 2010, Integrating Intermediate Inputs from Partially Classified Images within a Hybrid Classification Framework: An Impervious Surface Estimation Example. *Remote Sensing of Environment*, Vol. 114, No. 6, 1220-1229.

Ma, R., 2005, DEM Generation and Building Detection from Lidar Data. *Photogrammetric Engineering and Remote Sensing*, Vol. 71, No. 7, 847-854.

Maillard, P., 2003, Comparing Texture Analysis Methods through Classification. *Photogrammetric Engineering and Remote Sensing*, Vol. 69, No. 4, 357-367.

McMahon, G., 2007, Consequences of Land-Cover Misclassification in Models of Impervious Surface. *Photogrammetric Engineering and Remote Sensing*, Vol. 73, No.12, 1343-1353.

Mena, J. B. and Malpica, J. A., 2005, An Automatic Method for Road Extraction in Rural and Semi-Urban Areas Starting from High Resolution Satellite Imagery. *Pattern Recognition Letters*, Vol. 26, No. 9, 1201-1220.

Melgani, F. and Serpico, S. B., 2003, A Markov Random Field approach to Spatio-Temporal Contextual Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 11(Part 1) 2478-2487.

Miller, J. E., Nelson, S. A. C. and Hess, G. R., 2009, An object Extraction Approach For Impervious Surface Classification with Very-High-Resolution Imagery. *Professional Geographer*, Vol. 61, No. 2, 250-264.

Myint, S. W., 2007, Urban Mapping with Geospatial Algorithms. In: Urban Remote Sensing Q. Weng, D.A Quattrochi (Eds.), 109-135 (Boca Raton, Florida: Taylor & Francis).

Myint, S. W. and Lam, N. S. 2005, A Study of Lacunarity-Based Texture Analysis Approaches to Improve Urban Image Classification. *Computer, Environment and Urban System*, Vol. 29, No. 5. 501-523.

Myint, S. W., Lam, N. S. and Tyler, J. M., 2004, Wavelets for Urban Spatial Feature Discrimination: Comparisons with Fractal, Spatial Autocorrelation and Spatial Co-Occurrence Approaches. *Photogrammetric Engineering and Remote Sensing*, Vol. 70, No. 7, 803-812.

Nghi, D. L. and Mai, L. C. 2008, An Object-Oriented Classification Techniques for High Resolution Satellite Imagery. *In the Proceedings of International Symposium on Geoinformatics for Spatial Infrastucture Development in Earth and Applied Science*.

OBIA, 2006, The of Proceedings 1st International Conference on Object-based Image Analysis, July 2006 Salzburg, Austria. Available online at: http://www.commission4.isprs.org/obia06/index.html (Accessed April, 2010).

Ouma, Y. O. R. Tateishi, R. and Sri-Sumantyo, J. T., 2010, Urban Features Recognition and Extraction from Very-High Resolution Multi-Spectral

Satellite Imagery: A Micro–Macro Texture Determination and Integration Framework. *IET Image Processing*, Vol. 4, No. 4, 235-254

Pacifici, F., Chini, M., and Emery, W. J., 2009, A Neural Network Approach using Multi-Scale Textural Metrics from Very High-Resolution Panchromatic Imagery for Urban Land-Use Classification. *Remote Sensing of Environment*, Vol. 113, No. 6, 1276-1292.

Pesaresi, M. 2000, Texture Analysis for Urban Pattern Recognition Using Fine-Resolution Panchromatic Satellite Imagery. *Geographical & Environmental Modelling*, Vol. 4, No.1, 43- 63.

Pratt, W. K., 2007, Digital Image Processing, 4th ed. Wiley.

Price, J. C. 1997, Spectral Band Selection for Visible-Near Infrared Remote Sensing: Spectral-Spatial Resolution Tradeoffs. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 35, No. 5, 1277-1285.

Pu, R., Landry, S. and Qian, Y., 2011, Object-Based Urban Detailed Land Cover Classification with High Spatial Resolution IKONOS Imagery. *International Journal of Remote Sensing*, Vol. 32, No. 12, 285-3308.

Puissant, A., Hirsch, J. and Weber, C., 2005, The Utility of Texture Analysis to Improve Per-Pixel Classification for High To Very High Spatial Resolution Imagery. *International Journal of Remote Sensing*, Vol. 26, No. 4, 733-745.

Quattrochi, D. A. and Goodchild, M. F. 1997, Scale in Remote Sensing and GIS (Boca Raton, Florida: Lewis Publishers).

Richards, J. A., and Jia, X. 2006, Remote Sensing Digital Image Analysis: An Introduction, 4th ed. (Berlin: Springer).

Rottensteiner, F., Trinder, J., Clode, S. and Kubik, K., 2003, Building Detection using LiDAR Data and Multi-spectral Images. *In Proc. of VIIth Digital Image Computing: Techniques and Applications*, Sydney, Australia.

Rottensteiner, F., Summer, G., Trinder, J., Clode, S. and Kubik, K., 2005, Evaluation of a Method for Fusing Lidar Data and Multispectral Images for Building Detection. In: Stilla U, Rottensteiner F, Hinz S (Eds) CMRT05. IAPRS, Vol. XXXVI, Part 3/W24 - Vienna, Austria.

Salehi, B., Zhang,Y. and Zhong, M., 2011a, Combination of Object-Based and Pixel-Based Image Analysis for the Classification of VHR Imagery Over Urban Areas. *In the Proceedings of ASPRS 2011 Annual Conference*, May 1-5, Milwaukee, Wisconsin, USA.

Salehi, B., Zhang, Y. and Zhong, M., 2011b, Object-based Land Cover Classification of Urban Areas using VHR Imagery and Photogrammetrically-Derived DSM. *In the Proceedings of ASPRS 2011 Annual Conference*, May 1-5, Milwaukee, Wisconsin, USA.

Salehi, B. Zhang, Y. and Dey, V., 2011c, A Hierarchical Rule Based Urban Land Cover Classification of WorldView-2 Imagery, *In Proceedings of 32th Canadian Symposium on Remote Sensing*, Sherbrook, Canada.

Schiewe, J. and Elers, M., 2005, A Novel Method for Generating 3d City Models from High Resolution and Multi Sensor Remote Sensing Data. *International Journal of Remote Sensing*, Vol. 26, No. 4, 683-698.

Shackelford, A. K., and Davis, C. H., 2003a, A Combined Fuzzy Pixel-Based and Object-Based Approach for Classification of High-Resolution Multispectral Data Over Urban Areas. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 10 (Part1), 2354-2363.

Shackelford, A. K. and Davis, C. H. 2003b, A Hierarchical Fuzzy Classification Approach for High-Resolution Multispectral Data Over Urban Areas. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 9 (Part 1), 1920-1932.

Shekhar, S., Schrater, P. R., Vatsavai, R. R., Wu, W. and Chawla, S., 2002, Spatial Contextual Classification and Prediction Models for Mining Geospatial Data. *IEEE Transactions on Multimedia*, Vol. 4, No. 2, 174-188.

Sims, F. M. and Mesev, V., 2007, Use of Ancillary Data in Object-Based Classification of High Resolution Satellite Data. *In Proceedings of Urban Remote Sensing Joint Event*. April 11-13, 2007, Paris, France.

Sohn, G. and Dowman, I., 2007, Data Fusion of High-Resolution Satellite Imagery and LiDAR Data for Automatic Building Extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 62, No. 1, 43-63.

Song, M., Civco, D. L. and Hurd, J. D., 2005, A Competitive Pixel-Object Approach for Land Cover Classification. *International Journal of Remote Sensing*, Vol. 26, No. 22, 4981-4997.

Su, W., Li, J., Chen, Y., Liu, Z., Zhang, J., Low, T. M., Suppaih, I. and Hashim, S. A. M., 2008, Textural and Local Spatial Statistics for the Object-Oriented Classification of Urban Areas using High Resolution Imagery. *International Journal of Remote Sensing*, Vol. 29, No. 11, 3105-3117.

Tassetti, A. N., Malinverni, E. S., and Hahn, M., 2010, Texture Analysis to Improve Supervised Classification in Ikonos Imagery. In: Wagner W., Székely, B. (eds.): ISPRS TC VII Symposium – 100 Years ISPRS, Vienna, Austria, IAPRS, Vol. XXXVIII, Part 7A.

Thomas, N., Hendrix, C. and Congalton, R. G., 2003, A Comparison of Urban Mapping Methods using High-Resolution Digital Imagery. *Photogrammetric Engineering and Remote Sensing*, Vol. 69, No. 9, 963-972.

Tuia, D., Pacifici, F., Kanevski, M. and Emery, W. J., 2009, Classification of Very High Spatial Resolution Imagery using Mathematical Morphology and Support Vector Machines. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 47, No. 11, 3866-3879.

Van de Voorde, T., De Genst, W. and Canters, F., 2007, Improving Pixel-Based VHR Land-Cover Classifications of Urban Areas with Post-Classification Techniques. *Photogrammetric Engineering and Remote Sensing*, Vol. 73, No. 9, 1017-1027.

Van der Werff, H. M. A. and Van der Meer, F. D., 2008, Shape-Based Classification of Spectrally Identical Objects. *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 63, No. 2, 251–258.

Walker, J. and Blaschke, T., 2008, Object-Based Land Cover Classification for the Phoenix Metropolitan Area: optimization vs. transportability. *International Journal of Remote Sensing*, 29, 2021-2040.

Walter, V., 2004, Object-based Classification of Remote Sensing Data for Change Detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 58, No. 3-4, 225– 238.

Warner, T. A., and Nerry, F., 2009, Does Single Broadband or Multispectral Thermal Data add Information for classification of visible, near and Shortwave Infrared Imagery of Urban Areas? *International Journal of Remote Sensing*, Vol. 74, No. 7, 2155-2171.

Watanachaturaporn, P., Arora, M. K. and Varshney, P. K., 2008, Multisource Classification using Support Vector Machines: An Empirical Comparison with Decision Tree and Neural Network Classifiers. *Photogrammetric Engineering and Remote Sensing*, Vol. 74, No. 2, 239-246.

Wurm, M., Taubenböck, H., Schardt, M., Esch, T., and Dech, S. 2011, Object-Based Image Information Fusion using Multisensor Earth Observation Data over Urban Areas. *International Journal of Image and Data Fusion*, Vol. 2, No. 2, 121 - 147

Xu, H. and Li., P., 2010, Urban Land Cover Classification from Very High Resolution Imagery using Spectral and Invariant Moment Shape Information. *Canadian Journal of Remote Sensing*, Vol. 36, No. 3, 248-260.

Yu, B., Liu, H., Wu, J., Hu, Y. and Zhang, L., 2010, Automated Derivation of Urban Building Density Information using Airborne LiDAR Data and Object-Based Method. *Landscape and Urban Planning*, Vol. 98, No. 3-4, 210–219.

Zhang, Y., 1999, Optimisation of Building Detection in Satellite Images by Combining Multispectral Classification and Texture Filtering. *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 54, No.1, 50–60.

Zhang, Q. and Couloigner, I., 2006, Benefit of the Angular Texture Signature for the Separation of Parking lots and roads on High Resolution Multi-Spectral Imagery. *Pattern Recognition Letters*, Vol. 27, No. 9, 937-946.

Zhou, W., Huang, G., Troy, A. and Cadenasso, M. L., 2009, Object-Based Land Cover Classification of Shaded Areas in High Spatial Resolution Imagery of Urban Areas: A Comparison Study. *Remote Sensing of Environment*, Vol. 113, No. 8, 1769-1777.

Zhou, W. and Troy, A., 2008, An Object-Oriented Approach for Analyzing and Characterizing Urban Landscape at the Parcel Level. *International Journal of Remote Sensing*, Vol. 29, 3119-3135.

Zhou, W., Troy, A. and Grove, M., 2008, Object-Based Land Cover Classification and Change Analysis in the Baltimore Metropolitan Area using Multitemporal High Resolution Remote Sensing Data. *Sensors*, Vol. 8, No. 3, 1613-1636.