Revealing the Anatomy of Cities through Spectral Mixture Analysis of Multispectral Satellite Imagery: A Case Study of the Greater Cairo Region, Egypt.

Tarek Rashed, John R. Weeks and **M. Saad Gadalla**
Department of Geography, San Diego State University
San Diego, CA 92182-4493, U.S.A.
E-mail: trashed@mail.sdsu.edu
E-mailjohn.weeks@sdsu.edu
E-mail: gadalla@mail.sdsu.edu

Allan G. Hill
Harvard School of Public Health, Harvard University
Cambridge, MA 02138, U.S.A.
E-mail: ahill@hsph.harvard.edu

Abstract

This paper examines the feasibility of spectral mixture analysis (SMA) in deriving comparable physical measures of urban land cover that describe the morphological characteristics of cities. SMA offers a way of analyzing satellite imagery of urban areas that may be superior to more standard methods of classification. Mixing models are based on the assumption that the remotely measured spectrum of a given pixel can be modeled as a combination of pure spectra, called endmembers. SMA, using four image endmembers (vegetation, impervious surface, soil, and shade), was applied to an IRS-1C multispectral image in order to extract measures that describe the anatomy of the Greater Cairo region, Egypt, in terms of endmember fractions. The resulting fractions were then used to classify the urban scene into eight classes of natural and human-built features through a decision tree (DT) classifier. The accuracy of the DT classification was compared to the accuracies of two per-pixel supervised classifications of the IRS-1C image employing maximum likelihood (ML) and minimum distance-to-means (MDM) classifiers. Overall KAPPA accuracies were 0.88 for the DT classification based on SMA fractions, and 0.60 and 0.45 for the classifications conducted through ML and MDM respectively.

Introduction

How to describe patterns of the urban landscape is a fundamental question that has attracted the attention of geographers, ecologists, and other scientists interested in various urban phenomena. The traditional approach to addressing this question has been based on adopting a classification scheme by which the urban fabric is logically arranged in systems of discrete objects based on observable characteristics (Jensen *et al.*, 1983). As remotely sensed (RS) data have become increasingly incorporated in urban analyses, this approach has not significantly deviated from its classical origins. Rather, users of RS data have often seemed to maintain what Mather (1999, pp. 7) calls a "hard classification" view. That is, the world is viewed as a set of contiguous rectangular pixels, each of which is allowed to have only a single label representing one discrete land use or land cover category (Mather, 1999).

A review of the limited literature on urban remote sensing confirms the preoccupation with creating this kind of "hard" land-use/land-cover classification from imagery. In some cases, the delimitation of land cover and land use types is the goal of the analysis (Barnsley and Barr, 1996; Ryherd and Woodcock, 1996; Berberoglu *et al.*, 1999; Bibby and Shepherd, 1999; Couloigner and Ranchin, 2000). In others, the classification is considered as an intermediate step toward examining such phenomena as energy and moisture flux (Deguchi and Sugio, 1994), urban change (Chen *et al.*, 2000; Ward *et al.*, 2000), and urban heat islands (Lo *et al.*, 1997; Quattrochi *et al.*, 2000), or toward developing empirical models to estimate biophysical, demographic, and socioeconomic variables (Lo, 1995; Thomson and Hardin, 2000).

In spite of accomplishments in these applications, the accuracy with which urban land-use/land-cover classification may be derived from RS data using conventional, per-pixel techniques is often judged to be too low for operational use (Wilkinson, 1996; Foody, 1999). This is especially true in

Geocarto International, Vol. 16, No. 4, December 2001
Published by Geocarto International Centre, G.P.O. Box 4122, Hong Kong.

the context of multispectral images with medium spatial resolution such as those provided by Landsat TM, SPOT, and Indian satellites. Several reasons may be cited for why the "hard" classification approach limits the potential of remote sensing as a research tool for urban analysis. In this paper, however, we focus our attention on two main observations that are evident in the literature. The first concerns the classification schemes being adopted in the analysis of urban imagery, while the second is directly related to the nature of urban landscapes.

The first observation is that the majority of currently available classification schemes do not provide a clear distinction between land cover and land use (Ridd, 1995; Ward *et al.*, 2000). Ridd (1995, pp. 2165) notes that "success [in classification] is typically measured by the ability to match spectral signatures to the Anderson classification which, in the urban arena, is simply land use". This creates problems because land use is an abstract concept—a combination of economic, social, and cultural factors that is defined in terms of function rather than physical characteristics (Barr and Barnsley, 2000). Yet, to compare urban systems, whether between cities or even between various districts within a particular city, comparable descriptive parameters are required. RS imagery can only record land cover, which describes the physical state of features in urban lands (e.g., vegetation types, water bodies). Thus, land cover is more objectively measured than land use, which cannot directly be linked to RS data and is prone to interpretation error because different users will have different perspectives on the classification procedure (Anderson *et al.*, 1976). One way to overcome this problem is to base the measures on biological and physical structures of urban landscapes because these measures do not depend on human interpretation, nor are they influenced by such criteria as the economic development of the city (Ridd, 1995).

The second observation concerning the limitations imposed by "hard" classification techniques on urban remote sensing applications is directly related to the nature of urban landscapes (see Jensen and Cowen, 1999, for a more comprehensive discussion of urban landscape characteristics). One of the most important characteristics is the heterogeneity of urban features in relation to the spatial resolution of the sensor (Weber, 1994). Because the urban environment includes a complex mix of natural and human-built urban features often interwoven with one another, there is a need to deal with a complex mixture of spectral responses (Forster, 1985). With the presence of spectral mixing in the pixels of available satellite images, the identification of land cover using per-pixel classification techniques becomes very difficult since the continuum of land cover cannot be divided readily into discrete classes as required by the "hard" classification view. More recently there has been a move toward a "softer" way of describing the spatially varying character of land cover in terms of probability surfaces (Mather, 1999). In the "soft classification" approach, each pixel is assigned a class membership probability for each land cover type, representing it as a continuous surface of variation. Fuzzy classification and spectral mixture analysis (SMA) are two families of techniques designed to provide a "soft" classification of mixed pixels. The basic difference between them is that SMA is based on a physical model of the mixture of discrete spectral response patterns (Roberts *et al.*, 1998a), thus providing a deterministic method to addressing the spectral mixing problem rather than relying on statistical methods as in the case of the fuzzy approach (Mather, 1999).

In this paper, we present a replicable procedure to analyze the anatomy of cities using spectral mixture analysis (SMA) of multispectral images with medium spatial resolution. The present study represents part of a larger ongoing project undertaken by the International Population Center at San Diego State University and is directed toward applying remote sensing and GIS techniques to the analysis of demographic processes in Arab cities. We have favored the use of SMA over fuzzy classification techniques because it better serves our purpose of deriving standardized and comparable RS measures that can be utilized with census data in a GIS to study demographic dynamics in the Greater Cairo region, Egypt. Our specific objectives in this paper are as follows:

(1) Establish the feasibility of SMA in deriving comparable physical measures of urban land cover that describe the morphological characteristics of the study site; and
(2) Compare the accuracy of land-use/land-cover classifications derived by two different approaches. In the first, two conventional per-pixel classifiers (maximum likelihood and minimum distance-to-means) are applied directly to an IRS-1C multispectral image of the study area. In the second approach, an SMA model is first applied to the image and the model output is then used to derive a discrete land-use/land-cover classification for the study area through a decision tree classifier. We hypothesize that the second approach in which "hard" classification techniques operate in a subservient role to "soft" classification techniques may be more accurate (and hence more effective) than a direct per-pixel classification of the imagery.

Background

A Scene Model of Urban Land Cover Composition

In general, extracting information from RS data can be accomplished by the use of models that involve the earth's surface (the scene model), the atmosphere that lies between it and the spacecraft (the atmosphere model), and the image-forming sensors on board the spacecraft (the sensor model) (Graetz, 1990). Therefore, a discussion of a remote sensing scene model of urban landscapes can only be made with reference to a particular sensor model and how it produces the measurements that structure the image. Our discussion of the urban scene model is based on its relation to multispectral remote sensors with medium spatial resolutions such as IRS-1C (24m), SPOT (20m), and Landsat TM (30m).

Strahler *et al.* (1986) divide scene models into two types,

H- and L- resolution models, depending on the relationship between the size of elements (e.g., vegetation) in the scene and the resolution cell of the sensor. In the H-resolution model, scene elements are larger than resolution cells, and therefore the spatial arrangement of scene elements can directly be detected. The L-resolution model is the opposite, where scene elements are not individually detectable because they are smaller than the resolution cells. Detecting the spatial arrangement of objects may require a resolution cell size several times smaller than their size. Accordingly, for multispectral images with medium spatial resolution, the scene model of urban landscapes can be regarded as an L-resolution model. Further, as the size of objects in the urban scene becomes increasingly small relative to the resolution cell size, it may no longer be possible to consider objects individually (Strahler *et al.*, 1986). Instead, the urban scene model can be regarded as a *continuous* model, in which the measurement of each pixel can be treated as a sum of spectral interactions between various scene elements weighted by their concentration or relative aerial proportions within the resolution cell (i.e., a mixture model).

The implication of this reasoning is significant because it determines a new pathway for using RS imagery in the analysis of urban landscapes. Even though H-resolution models have been dominating urban remote sensing analyses for decades, they have a reduced role to play in the inference of urban structures (Graetz, 1990). Rather, it can be asserted that H-resolution models in future urban applications will operate in a subservient role to L-resolution models—an approach that we attempt to test in this study.

Spectral Mixture Analysis (SMA)

The past few years have witnessed an increasing use of SMA within the remote sensing community (Adams *et al.*, 1986; Adams *et al.*, 1993; Novo and Shimabukuro, 1994; Tompkins *et al.*, 1997; Gross and Schott, 1998; Roberts *et al.*, 1998b; Mather, 1999; Peddle *et al.*, 1999). The majority of SMA applications have been directed toward the natural environment. However, SMA has obvious applications to the urban environment. The advantage of SMA over traditional classification techniques lies in two major areas (Roberts *et al.*, 1998a): (1) SMA conforms well to the scene model because it is a physically based model that transforms radiance or reflectance values to physical variables that are linked to the sub-pixel abundance of endmembers within each pixel, and (2) SMA provides quantitative results that can be incorporated into models of the processes governing the distribution of materials within the urban scene.

Mixing models are based on the assumption that the landscape is formed from continuously varying proportions of idealized types of land cover with pure spectra, called endmembers. Endmembers are features recognizable in the scene as being abstractions of land cover materials with uniform properties. The pure spectra of endmembers can be measured in the laboratory, in the field, or extracted from the image itself. In an urban environment, these may include impervious surfaces, vegetation covers, water bodies, and bare soils (Ridd, 1995). Through SMA, the areal fractions of the endmembers are quantified at the sub-pixel level, allowing the inference of the morphological characteristics of an urban landscape in terms of endmember composition. Linear SMA is the process of solving for endmember fractions, assuming that the spectrum measured for each pixel represents a *linear* combination of endmember spectra that corresponds to the physical mixture of some components on the surface weighted by surface abundance (Tompkins *et al.*, 1997). Spectral mixture of endmembers also has the potential to become nonlinear (i.e., when radiations interact with more than one component). Although nonlinear mixing can become significant for some types of analysis, the effects of multiple scattering in the majority of applications are assumed to be negligible (Roberts *et al.*, 1998a; Mather, 1999; Lillesand and Kiefer, 2000). The analysis and results presented here are based on the assumption of linear mixing. Assuming a linear mixing helps establish a direct link between SMA and Ridd's (1995) conceptual model of urban landscape composition which we utilize to guide our analysis.

The Study Area and Data

The selected area of study is the metropolitan area of Cairo, Egypt, and its surroundings (Figure 1). The area covers 22.9 km X 22.2 km, encompassing major parts of the governorates of Cairo and Giza. The Nile forms the administrative division between these two governorates, with Cairo on the east bank of the river and Giza on the west bank. The area includes a variety of land uses associated with a complex mix of land cover. The Mukatim desert occupies the southern part of the scene whereas the northern part includes a green belt comprising agricultural fields that are continuously being intruded by urbanization. The urbanized areas are located at the center of the scene. In these areas, residential use is often mixed with commercial, public, and sometimes "light" industrial uses within the same block. However, variations can easily be observed by the naked eye between: (a) higher social status residential areas (sites 1 to 4 in Figure 1) with low population density (6,300 people/km^2); (b) lower social status residential areas (sites 7 to 11) with high density (44,800 people/); (c) the central business district (CBD) of the capital (site 5); and (d) newly developed lands (site 6).

Two satellite images have been used in our analysis. The first is an Indian Remote Sensing multispectral image (IRS-IC LISS-III) acquired on June 12th, 1996 covering three bands in VNIR (520-590, 620-690 and 770-860 nm - 23.6 m spatial resolution), and one band in SWIR (1550-1700 nm - 70.8 m spatial resolution). The second is a panchromatic image acquired from the same sensor on June 26th, 1998, which covers a spectral range between 500 and 750 nm at 10 m spatial resolution. While the SMA primarily relied upon the multispectral image, the panchromatic image was utilized in conjunction with a polygon coverage for the 1996 land use survey of Cairo for the selection of test sites used for assessing the accuracy of the final classification.

Methods

SMA of the IRS -1C Multispectral Image

The SMA approach is summarized in Figure 2. The analysis begins with the selection of a set of endmembers, followed by applying an SMA model in order to estimate endmember fractions. A good model is one that produces physically realistic fractions (i.e., between 0% and 100%) and measures of error less than a certain threshold (e.g., RMS < 5 DN). That may not happen if the model uses an improper set of endmembers (e.g., the exclusion of an endmember which is represented in the scene, or the addition of an endmember which is not represented). In such a case, the selection of endmembers would have to be refined. This process is repeated until the optimum set of endmembers is identified. Finally, a direct analysis of the fractions or classification is conducted according to the purpose of the application.

Successful SMA application relies on the accuracy of endmember selection. If the endmembers are incorrect in the physical sense, then the fractional abundances are also incorrect and the results of SMA become meaningless. The selection of endmembers can be done in two ways (Adams *et al.*, 1993): (1) by deriving them directly from the image (*image endmembers*), or (2) from field or laboratory spectra of known materials (*reference endmembers*) (see Roberts *et al.*, 1998a for a comparison between the two). In the present study, we have relied exclusively upon image endmembers extracted from the IRS image for two reasons. First, the study is exploratory in nature and only utilizes a single-date image. Second, we did not have reference endmembers collected from the study area in Cairo at the time of research.

The conceptual model selected to extract image endmembers from the RS data is Ridd's VIS model (Ridd, 1995). The VIS model represents the composition of an urban environment as a linear combination of three types of land cover, namely green Vegetation, Impervious surfaces, and bare Soil. Just as soils may be described in terms of their proportions of salt, silt, and clay using the traditional triangular diagram, so various subdivisions of urban areas may be described in terms of proportions of vegetation, soil, and impervious surface (Figure 3). Ridd's VIS model provides an intuitive link to the spectral mixing problem, because the spectral contribution of its three

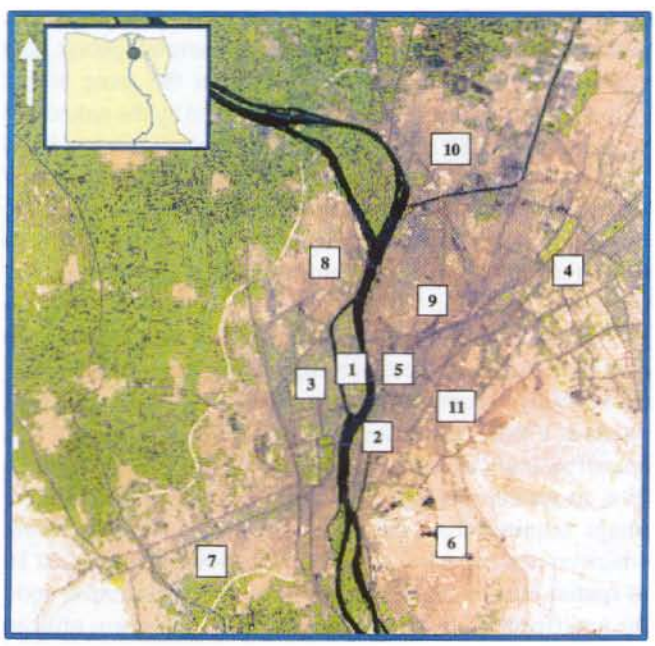


Figure 1 Satellite image for metropolitan Cairo and its surroundings. The numbers show different aspects of the urban scene (see text for discussion)

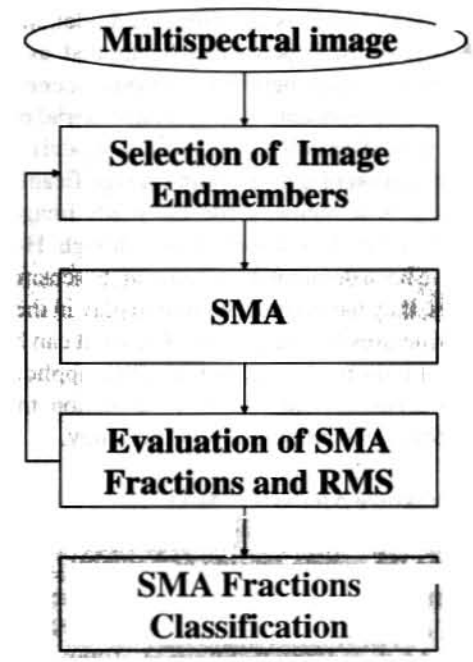


Figure 2 Flow chart for SMA process using image endmembers.

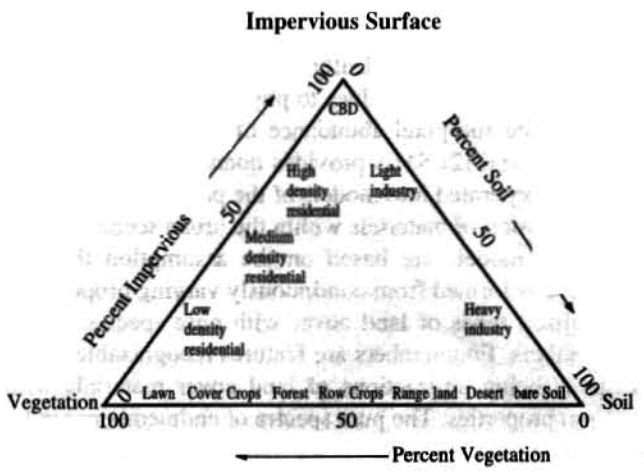


Figure 3 Ridd's Vegetation-Impervious surface-Soil (VIS) model for urban composition (after Ridd, 1995)

main components can be resolved at the sub-pixel level using the SMA technique. The model was originally introduced with reference to the contemporary urban realm of American cities, but has been tested in other urban areas such as Queensland, Australia (Ward *et al.*, 2000) and the Bangkok metropolitan area in Thailand (Madhavan *et al.*, 2001). The findings of these studies confirm that the model exhibits a general applicability to various cities. However, for those cities that differ in their urban fabric from that of the American cities, the model may require an additional component (e.g., shade) to achieve a better characterization of their morphological patterns.

Image endmembers were selected using the Pixel Purity Index (PPI) (Boardman *et al.*, 1995). The PPI method allocates to each pixel a score based on the number of times it is found to occupy a near-vertex position in the repeated projections of the n-dimensional data onto a randomly-oriented vector passing through the mean of a data cloud. The resulting score helps identify image endmembers because those pixels that hold pure spectra will have a high score (i.e., will be found repeatedly at the extremes of the data distribution). The final five endmembers selected for our data included two endmembers for bare-soil, and one each for vegetation, impervious surface, and shade. Two endmembers were required to represent bare-soil in order to ensure having representative soil endmembers for the diverse land cover of the study area (i.e., desert soil, agricultural soil and urban soil). Figure 4 shows the spectral profiles of these endmembers and their locations on a 2D scatter plot (red/NIR).

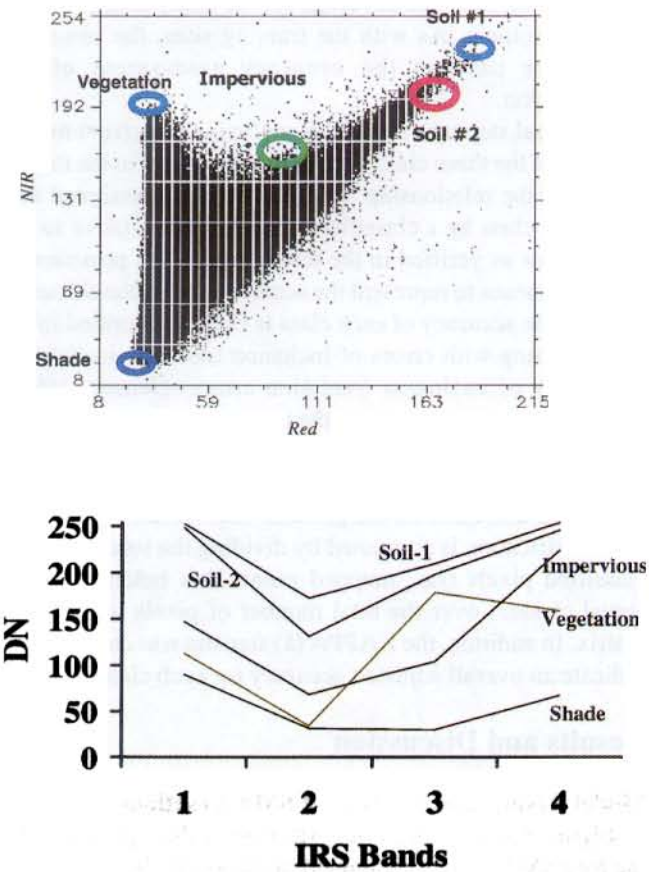


Figure 4 At the top, NIR to red scatter plot showing the spectral regions for the selected endmembers. At the bottom, the spectral profiles of endmembers.

Having identified representative endmembers, a computer program including a linear unmixing code was applied to the multispectral IRS-1C image. In this code, we employed the unconstrained Modified Gram-Schmidt least square method (Roberts *et al.*, 1998a), in which fractions are constrained to sum to 1 while individual fractions are allowed to be less than 0 or greater than 1. Given a mixture and a set of endmembers, the Gram-Schmidt method attempts to solve for the fractions through a series of linear equations. The specific formulation can be found in Adams *et al.* (1993), Tompkins *et al.* (1997), and Roberts *et al.* (1998a). When the equations are applied to an image consisting of N spectral bands using a number of endmembers less than or equal to N, the output is a fraction image for each endmember and some measure of fit.

We examined four SMA models: a 3-endmember model of vegetation, impervious surface, and bare-soil#1; a 3-endmember model of vegetation, impervious surface, and bare-soil#2; a 4-endmember model of vegetation, impervious surface, bare-soil#1, and shade; and finally, a 4-endmember model of vegetation, impervious surface, bare-soil#2, and shade. The impervious surface endmember represented the average spectrum of tile roofs and asphalt (i.e., few "big" buildings and wide roads were identifiable from the image). The shade endmember was used in our analysis as an indicator of building heights—a factor that appears to be significant in characterizing the morphological patterns of Cairo. A water endmember was used as a surrogate for a pure shade spectrum since both water and shade exhibit the same characteristics of dark objects in the visible spectrum.

Model fit, as discussed earlier, was assessed in terms of two criteria: (1) whether the fractions provide realistic abundance and (2) in an error term. The error term was expressed as a root-mean-square (RMS) error, which provides an estimate of the average error calculated for each pixel across all bands using the following equation:

$$RMS = \sqrt{\frac{\sum_{1}^{N}(\varepsilon_{i\lambda})^2}{(N-1)}} \qquad (1)$$

where N is the number of bands $\varepsilon_{i\lambda}$ and represents a residual term calculated for each pixel i at wavelength λ as follows:

$$\varepsilon_{i\lambda} = R_{i\lambda} - \sum_{m=1}^{M} r_{m\lambda} \times f_{mi} \qquad (2)$$

where $R_{i\lambda}$ is the mixed spectrum (e.g. DN, radiance, reflectance) of the pixel *i* at wavelength λ, and $r_{m\lambda}$ is the pure spectral response of an endmember *m* (of total endmembers *M*) at wavelength λ, weighted by the fractional abundance of the endmember f_{mi}, within that pixel.

Classification

Experiments were carried out in order to examine the basic hypothesis that SMA-derived fractions provide a way to classify the urban scene better than other conventional per-pixel classification methods. To do so, the urban scene of the study area was classified using three different algorithms. The first classification applied a decision tree (DT) classifier (Hansen *et al.*, 1996; Friedl and Brodley, 1997) using endmember fractions and calculated RMS errors as input. A tree was constructed in S-plus software by recursively partitioning the dataset into purer, more homogeneous subsets using a measure called *deviance* (i.e., a likelihood ratio statistic) to compare all possible splits of the data in order to find one split that maximizes the dissimilarity among resulting subsets. The resultant tree represented hierarchical, nonlinear relationships within the data, composed of a root (representing the first splitting rule with the maximum deviance), a set of nodes (representing the internal splits), and a set of leaves (representing various classes at the terminal nodes). The robustness of the resultant tree was examined in terms of a residual mean deviance and misclassification error rate of the tree (number of misclassified points/total number of points) which indicate whether or not the training samples used in creating the tree are representative. Finally, the splitting rules of the resultant tree were entered in a separate classification program applied to the SMA output.

The other two classifications applied maximum likelihood (ML) and minimum distance-to-means (MDM) supervised classifiers (currently implemented in ERDAS Imagine software) to the original multispectral image bands. We chose these two techniques because the literature indicates that they have frequently been used in the classification of urban areas. In the three cases, the scene was classified into 8 target classes, namely: desert (DS); water body (WB) which includes the Nile River and its tributaries; active agricultural areas (AG); urban parks and recreation areas (UG); residential areas with higher social-status (HC); residential areas with lower social-status (LC); the central business district (BD); and newly developed lands (DL). These classes, particularly the last four, correspond to the long-term objective of our project concerning the analysis of demographic dynamics in the region.

The same training sites were used for all three classifiers so that the final results could be compared. The approach we followed to select the training samples was based on the "guided clustering" steps suggested by Bauer *et al.* (1994), and currently adopted by the GAP analysis program (Lillesand and Kiefer, 2000). The approach is very efficient in the urban context because it allows the analyst to delineate numerous training sets that are not perfectly homogeneous for each class in a scene. The following steps summarize the guided clustering approach (Bauer *et al.*, 1994; Lillesand and Kiefer, 2000):

1. Delineate "initial" training areas for a target class X from the multispectral image.
2. Cluster all pixels of class X into spectral subclasses X1,.....,Xn using ISODATA automated clustering algorithm.
3. Examine resultant subclasses and merge or delete signatures as appropriate.
4. Repeat steps 1 to 3 for the rest of the 8 target classes.
5. Perform maximum-likelihood classification using all spectral subclasses on the entire image.
6. Aggregate (RECODE) subclasses back to the original 8 target classes.
7. Select "final" training sites from the original multispectral image by utilizing information from the classified image produced in step 6 in conjunction with the land use coverage of the study area. The total number of points selected for the final training sample was 1,589 with an average of 175 to 200 points for each of the 8 classes.

Accuracy Assessment

An accuracy assessment process was conducted to quantitatively compare the DT based classification of SMA output bands to ML and MDM classifications of the original image bands. We used a number of test sites that were collected independently from the training samples used in the classifications. Using the panchromatic image of the study site, in combination with our familiarity with the area, a total sample of 1,424 points was identified as test sites for the 8 classes (each class contained between 160-190 points). These sites were further checked against information provided by the digital land-use coverage for the study area to ensure their correctness. As with the training sites, the same test sites were used in the accuracy assessment of the classification.

The final step was to build a confusion (*or error*) matrix for each of the three classified images. The confusion matrix expresses the relationship between the pixels assigned to a particular class by a classification algorithm relative to the actual class as verified in the test sample. This provides an effective means to represent the accuracy of the classification because the accuracy of each class is clearly described in the matrix, along with errors of inclusion (commission errors) and errors of exclusion (omission errors) (Jensen, 1996). The omission error indicates the probability that a test pixel is correctly classified, while the commission error indicates the probability that a pixel classified on the image actually represents that class on the ground. The overall accuracy of the classification is computed by dividing the total correctly classified pixels (i.e., mapped points that belong to their actual classes) over the total number of pixels in the error matrix. In addition, the KAPPA (k) statistic was computed to indicate an overall adjusted accuracy for each classification.

Results and Discussion

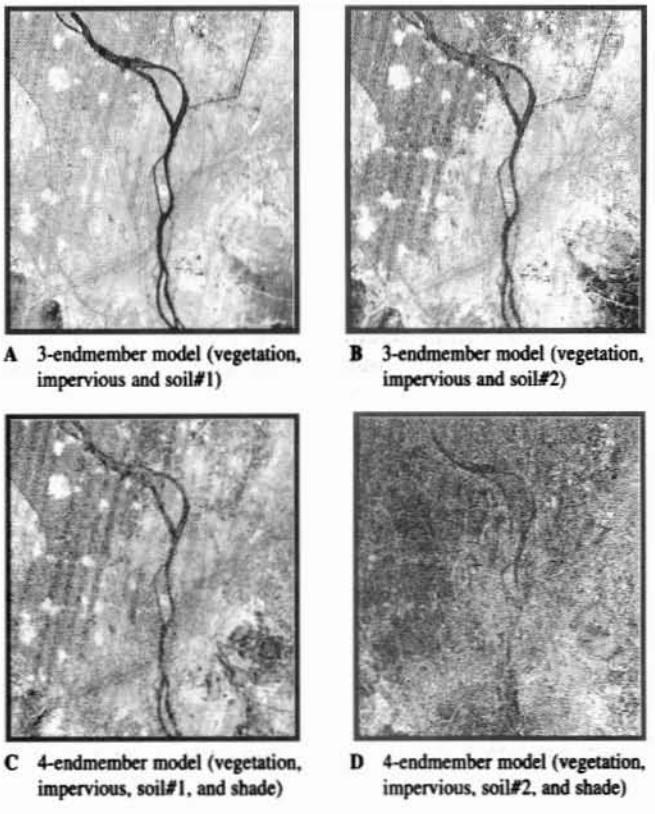
Model Results and Analysis of SMA Fractions

Figure 5 shows maps of RMS error values produced by the four SMA models examined in this study. Brighter areas indicate high RMS errors while darker areas indicate low errors. Among the four models, the 4-endmember model

which utilizes vegetation, impervious surface, bare-soil #2, and water (used as a surrogate for the shade) was found to produce the best results with realistic fraction values and a mean RMS error of 4.3%. As shown in map "D" of Figure 5, the RMS errors for this model do not show any systematic pattern in comparison to the other three maps. In addition, significant reductions in error values can be observed in the urbanized areas of the scene. This conforms well to our observation that Ridd's VIS model may need to utilize an additional component when applied to other cities that differ in their morphological patterns from those of American cities. In the case of Cairo, shade appears to be important in the distinction between various urban subdivisions.

The SMA fractions of vegetation, impervious surface, bare-soil (#2), and water/shade endmembers are shown in Figure 6. Brighter areas indicate a higher fractional abundance of the endmember while darker areas indicate lower abundance. These fractions provide a measure of the physical properties of the dominant classes in the scene, thus helping to reveal the morphological patterns of the Cairo metropolitan area and its surroundings. For example, the active agricultural fields in the NW quadrant of the scene can be characterized as consisting primarily of vegetation and shade with a minor amount of soil, consistent with the types of crops cultivated in these areas (map A of Figure 6). In contrast, urban vegetated areas such as recreational parks and lawns include a lower shade content, and higher vegetation (i.e., higher green leaf) and soil, consistent with the small trees and gaps of exposed soil that exist in such areas. The Nile River and its tributaries can be characterized as having a high water/shade content due to the low reflectance (map C), while the Mukatim desert at the southern portion of the scene consists exclusively of bare-soil (map B).

Both the impervious surface and shade endmembers play a more important role than bare-soil endmember as we move to the urbanized area of the scene. The central business district (CBD) of the city can be described as having a high content of impervious surface and shade, with very low vegetation and bare-soil fractions (map D). The shade fractions are also very effective in revealing the pattern of wide street networks due to the darkness of pavement materials and shadows from buildings and trees. Further, patterns of vegetation, impervious surface and shade fractions display the physical variability between different residential districts. For example, the dominant components of residential areas with higher social strata include a considerable amount of impervious surface, high shade fractions, and some vegetation. This is consistent with the fabric of these areas which include a variety of high-rise structures with different building and roofing materials (e.g., steel, concrete) mixed with recreational areas, sport clubs, and relatively wide boulevards. In contrast, the less affluent residential districts with a lower social-status can be characterized as having lower shade and vegetation fractions, which reflect the "higgledy-piggledy burrows" of Cairo's popular quarters (Rodenbeck, 1999, pp. 224) and are associated with narrow streets and low-lying buildings made of local materials (i.e.,



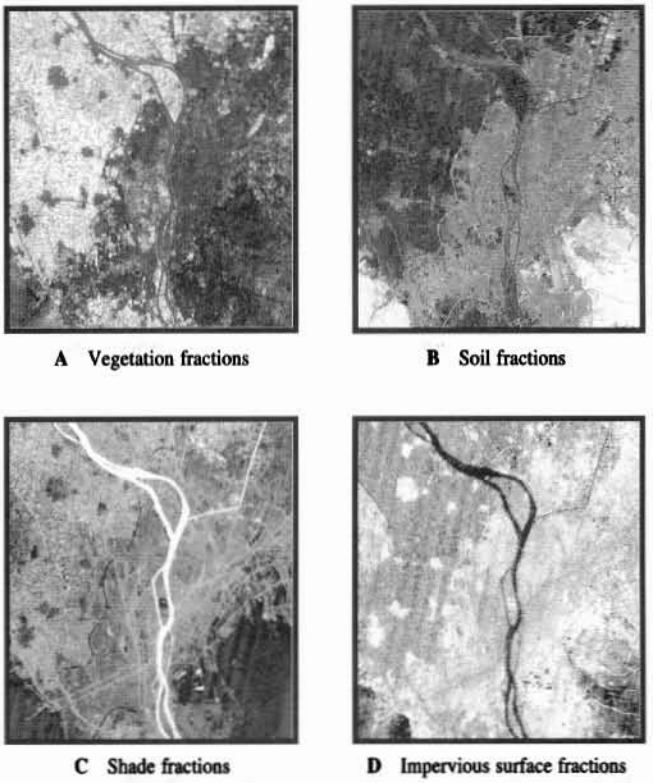
A 3-endmember model (vegetation, impervious and soil#1)

B 3-endmember model (vegetation, impervious and soil#2)

C 4-endmember model (vegetation, impervious, soil#1, and shade)

D 4-endmember model (vegetation, impervious, soil#2, and shade)

Figure 5 RMS error maps for the four models examined in the study. Brighter areas indicate higher errors while darker areas indicate lower errors.



A Vegetation fractions

B Soil fractions

C Shade fractions

D Impervious surface fractions

Figure 6 Fraction images produced by the 4-endmember model utilizing vegetation, impervious, soil#3, and shade. Brighter areas indicate higher abundance while darker areas indicate lower abundance.

3 to 5 stories on average). Finally, the newly developed land located at the SE quadrant can be distinguished by its high soil and impervious fractions, and low shade and vegetation fractions.

Analysis of the RMS errors is also very useful. The errors indicate that vegetated areas, the Nile River, and urbanized areas are well characterized by the endmembers (low RMS with an average less than 3.5%). On the other hand, some portions of the Mukatim desert are poorly characterized (high RMS) due to variations in the soil reflectance. These variations can mainly be attributed to the saturation of pixel values because of the strong reflectance that is beyond the range of the sensor to detect. Table 1 summarizes the characteristics of various urban classes in terms of SMA fractions and RMS values.

Classification Results

The training sample was used as input to a decision tree classifier. This technique was chosen because it requires a minimum number of assumptions about statistical properties of the classes, yet has the potential to provide a set of decision rules based on physical properties (Roberts *et al.*, 1998a). The final tree consisted of 19 nodes with a misclassification rate of 0.07, indicating a high overall classification accuracy of the training sites. The classification tree, shown in Figure 7, can further our understanding of the anatomy of the study area in terms of its physical patterns. The first splitting rule, occurring at the root node, is based on vegetation, implying that greenness is the variable that produces the largest deviance measure. The classes under the low-vegetation category (< 15%) include desert (DS), water bodies (WB), newly developed land (DL), lower social-status residential areas (LC), the CBD (BD), while those with a high-vegetation component (> 15%) include active agricultural fields (AG) and urban parks and recreational areas (UG). The class of higher social-status residential areas (HC) is located under both categories, confirming a high heterogeneity and physical variability of this class in Cairo as discussed earlier. The desert class is classified through low vegetation, low impervious and low shade filter, very similar to the water class but the latter has a higher shade component. Both newly developed land (DL) and lower social status residential areas (LC) classes can be reached through a high impervious and low shade filter. However, the bare-soil plays a vital part in the distinction between the DL and the LC classes, with DL higher in the percentage of exposed soil and lower in vegetation. In the case of the business district class (BD), the shade becomes more important (> 40%). The multiple paths for some classes reflect variability in class composition: either natural, as in the case of active agricultural (AG) and urban parks (UG) classes (e.g., different types of cultivated crops, grass), or human-induced as in the case HC, LC and BD classes (e.g., building heights, roofing materials, pavement conditions). The tree confirms the diminished role of soil in the context of long-established residential areas of Cairo since the distinction between HC and LC classes is achieved in terms of vegetation, shade, and impervious surface. However, soil is still important for the distinction between other classes such as AG versus UG, as well as DL versus LC.

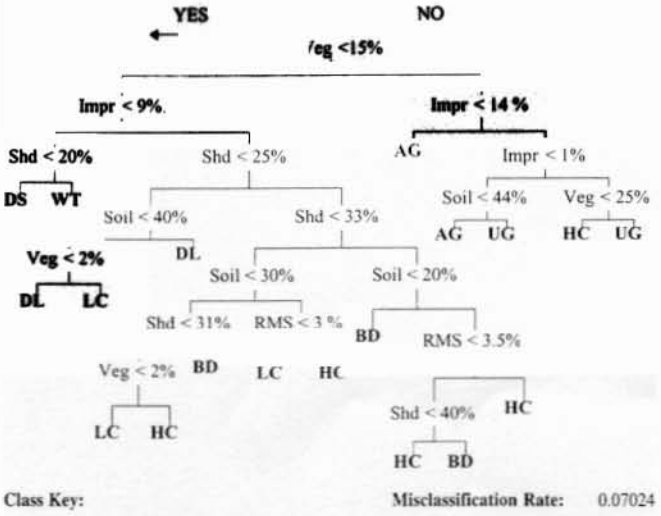


Figure 7 Binary decision tree used to classify the images based on spectral fractions and RMS errors. The numbers in the decision tree indicate the percentage of endmember fractions at which splits occur as calculated by S-plus using the deviance measure.

Accuracy Assessment Results

Applying the splitting rules to the SMA output bands, the scene of the study area was classified into the 8 classes discussed above. The final classification was then assessed in terms of its individual class accuracy. The commission and omission errors of the 8 classes are reported in Table 2 (in percent). The water bodies (WB) class shows the best result with only 0.88% commission error indicating the representative use of water/shade endmember in the classification. The desert (DS) class also shows a high accuracy (0% commission error), with an error of omission

Table 1 The composition of various urban classes in terms on SMA fractions

Class	SMA Fractions
Active agricultural areas (AG)	High vegetation, some shade, low soil
Urban parks and recreational areas (UG)	High vegetation, high soil, low shade
Desert (DS)	High soil, high RMS
Water bodies (WB)	High shade
Newly developed land (DL)	Some vegetation, some soil, some impervious, low shade
Central business district (BD)	High impervious, high shade, low vegetation
Residential areas with lower social-status (LC)	Low vegetation, low shade, some impervious
Residential areas with higher social-status (HC)	High shade, some impervious, some vegetation

indicating that only 15.31% of the test points were incorrectly excluded from that class. The worst case was the HC class with a 12.30% error of commission and 58.20% error of omission. This indicates that the training areas did not account for the high diversity of the physical settings that exist in the residential areas of higher social-status—a characteristic which we have already discussed above. Because the LC class corresponds to the residential areas with lower social-status, which are characterized by a homogeneous fabric of landscape, the accuracy of that class was much better (26.95% and 2.34% commission and omission errors respectively). As for the other urban classes, business district (BD) and newly developed land (DL), the accuracy scores were quite good (2.23% and 3.74% omission error for BD and DL respectively). The DL class, however, has a higher commission error of 16.82% since some of the test points under the desert class had been misclassified as the DL class because of the high soil fractions. Finally, the accuracy results of classes representing vegetated areas were acceptable, with 7.46% omission error for active agricultural areas (AG) class and 16.33% for urban parks (UG) class. The analysis of commission errors indicates a degree of confusion between these two classes due to the similarity in their physical attributes.

As mentioned earlier, we also assessed the classification accuracy conducted through two per-pixel supervised classifiers: maximum likelihood (ML) and minimum distance-to-means (MDM). The purpose was to test our basic hypothesis that a classification of the urban scene based on SMA-derived measures may be superior to other traditional per-pixel classifications techniques. Results of the comparison between the overall accuracy of the three classifications are shown in Table 3. As reported in the table, the overall accuracy of the classification based on SMA fractions and RMS errors was 89.52% (with a KAPPA (k) coefficient = 0.88) indicating that the technique performs well. The overall accuracy was severely reduced in the case of the other two classifiers applied directly to the original image bands (64.51% and $k = 0.60$ in the case of ML, and 52.69% and $k = 0.45$ for the MDM).

These results suggest that, in the case of multispectral images with medium spatial resolution, a classification based on SMA-derived fractions would be recommended over the other two traditional per-pixel classifiers. The results also confirm that a decision tree (DT) model is robust and well suited to representing the complexity of interactions between diverse urban classes through its hierarchical, nonlinear structure. Vegetation and impervious surface fractions that operate at regional scale (e.g., distinguish broadly between vegetated classes versus urbanized classes) were used as splitting criteria early in the model, while shade and soil fractions that have local influence (e.g., distinguish between various urbanized classes) were used near the terminal nodes.

Table 2 Error of commission and error of omission for different classes derived through a decision tree classification based on SMA-derived fractions.

Class	Commission (%)	Omission (%)
Active agricultural areas (AG)	17.41	7.46
Urban parks and recreational areas (UG)	5.98	16.33
Desert (DS)	0.00	15.31
Water bodies (WB)	0.88	0.00
Newly developed land (DL)	16.82	3.74
Central business district (BD)	3.63	3.23
Residential areas with lower social-status (LC)	26.95	2.34
Residential areas with higher social-status (HC)	12.30	58.20
	Overall Accuracy = 89.51%	
	Kappa = 0.8793	

Summary and Conclusion

In this paper, we have described a remote sensing methodology for analyzing the anatomy of cities using the Greater Cairo metropolitan area as an example. The methodology adopted is based on applying the SMA technique, using endmembers derived from the image. We have utilized Ridd's VIS model as a conceptual framework to guide us through the selection procedure of the endmembers. However, the results of the analysis indicate that a 4-endmember model that utilized an additional endmember (i.e., shade/water) will provide better fractions and lower RMS error than a 3-endmember model based on the three main components of the VIS model. This implies that Ridd's model may require some modification when it is applied to other settings that differ in their morphological patterns from the American cities.

Using a 4-endmember model, we extracted four fractional bands, which provided a "soft" classification of the urban scene that describes what materials, and how much, are present on the ground. We used the results of the soft classification to perform a "hard" classification by which the urban scene was classified into 8 discrete classes of natural and human-built features. Accuracy results validate our hypothesis that an approach in which the hard classification complements the soft analysis of imagery has the potential to provide improved discrimination of urban classes over other traditional per-pixel classification techniques.

SMA addresses the spectral mixture problem which implicitly exists in all urban imagery with a medium spatial

Table 3 A comparison between the overall classification and KAPPA accuracies for the three classifications applied to the study area.

	Decision tree (based on SMA fractions	Maximum likelihood	Minimum distance-to-means
Overall accuracy (%)	89.5194	64.5126	52.6875
KAPPA	0.8793	0.5950	0.4450

resolution. Future investigations are still needed to examine the feasibility of applying SMA to other urban settings, as well as to explore potential uses of the technique other than traditional land-use/land-cover mapping. Of course, the SMA technique does have its own limitations, specifically the identification of the required number of endmembers and their spectral characteristics. Further, endmembers derived from a single-date image cannot be used to analyze other images from different dates. This imposes a major limitation for applications such as urban change detection. Directly related to this problem is the conflict that exists between the number of endmembers that can be used in the analysis (4 in the case of IRS images) and achieving a successful model of the diverse patterns of urban landscapes. We offer the following two suggestions for future research concerning the use of SMA in the urban arena:

(1) Investigate the use of reference endmembers for conducting SMA. This can be achieved by building a region-specific spectral library. This step is essential if temporal relationships between urban biophysical variables and other phenomena such as the development of the city or social variations are to be investigated.

(2) Employ techniques that allow each pixel in the image to be modeled as different endmember combinations. This technique is known as Multiple Endmember Spectral Mixture Analysis (MESMA) (Roberts *et al.*, 1998b). MESMA incorporates a large number of endmembers in the analysis while meeting the constraints regarding the relationship between the number of endmembers and image bands. Thus, MESMA can account for variations between different materials constituting the built environment (since the assumption that a mixture between constant endmembers for all pixels in an image is unlikely to be valid).

In terms of the operational value of the SMA approach in connection with our ongoing project in Egypt, the SMA-derived classification of the urban scene seems capable of revealing the anatomy of the metropolitan area of Cairo and describing areal differences between various urban districts (i.e., LC, HC, BD, DL), which can be linked to variations in wealth or social class. An analysis of Landsat images for Detroit, Michigan, has shown that in that city a classification of the change in different types of vegetation occurring in urban areas is associated with socioeconomic changes occurring in these areas (Ryznar, 1998). Our analysis takes the classification scheme beyond that, to examine not only vegetation (which was indeed an important variable in our data analysis), but other features of the urban scene including buildings, shades, and bare soils. The implication of this is significant as it suggests that remote sensing imagery can be used to compensate for deficiencies in the range of data collected in the census. For example, it is rare in developing countries such as Egypt for censuses to include questions on income and wealth. However, the results of the present research indicate that features of the urban environment that are observable "from the top" can be classified and quantified to represent patterns of urban morphology that are associated with characteristics of the people living on the ground—a hypothesis which remains to be tested.

Acknowledgments

The authors would like to thank D. Roberts and R. Powell, University of California Santa Barbara, and J. Rogan, San Diego State University, for their constructive comments. This research has been supported by grants from the Andrew Mellon Foundation and the National Science Foundation (Grant No. BCS-0095641). An earlier version of this paper was presented at the Annual Meeting of the Association of American Geographers, New York City, 2001.

References

Adams, J. B., M. O. Smith and A. R. Gillespie, 1993. Imaging Spectroscopy: Interpretation based on Spectral Mixture Analysis. *In Remote Geochemical Analysis: Elemental and Mineralogical Composition*, C. M. Pieters and P. Englert, Cambridge, Cambridge University Press: 145-166.

Adams, J. B., M. O. Smith and P. E. Johnston, 1986. Spectral Mixture Modeling: A New Analysis of Rock and Soil Types at the Viking Lander Site. *Journal of Geophysical Research* **91**(8098-8112).

Anderson, D. E., E. Hardy, R. J and R. Witmer, 1976. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. Washington D. C., U.S. Geological Survey.

Barnsley, M. J. and S. L. Barr, 1996. Inferring Land Use from Satellite Sensor Images using Kernel-Based Spatial Re-classification. *Photogrammetric Engineering & Remote Sensing* **62**(8): 949-958.

Barr, S. and M. Barnsley, 2000. Reducing Structural Clutter in Land Cover Classification of High Spatial Resolution Remotely-Sensed Images for Urban Land Use Mapping. *Computers & Geosciences* **26**: 433-449.

Bauer, M. E., T. E. Burk, A. R. Ek, P. R. Coppin, S. D. Lime, *et al.*, 1994. Satellite Inventory of Minnesota Forest Resources. *Photogrammetric Engineering and Remote Sensing* **60**(3): 287.298.

Berberoglu, S., C. D. Lloyd, P. M. Atkinson and P. J. Curran, 1999. The Integration of Spectral and Texture Information using Neural Networks for Land Cover Mapping in the Mediterranean. *Computers & Geosciences* **26**: 385-396.

Bibby, P. and J. Shepherd, 1999. Monitoring Land Cover and Land Use for Urban and Regional Planning. *In Geographical Information Systems*, P. A. Longley *et al.* New York, John Wiley & Sons, Inc. **V2: Management Issues and Applications**: 953-965.

Boardman, J. W., F. A. Kruse and R. O. Green, 1995. Mapping Target Signatures via Partial Unmixing of AVIRIS Data. Summaries, *Fifth JPL Airborne Earth Science Workshop, JPL Publications 95-1* **1**: 23-26.

Chen, S., S. Zheng and C. Xie, 2000. Remote Sensing and GIS for Urban Growth in China. *Photogrammetric Engineering & Remote Sensing* **66**(10): 593-598.

Couloigner, I. and T. Ranchin, 2000. Mapping of Urban Areas: A Multiresolution Modeling Approach for Semi-Automatic Extraction of Streets. *Photogrammetric Engineering & Remote Sensing* **66**(7): 867-874.

Deguchi, C. and S. Sugio, 1994. Estimations of Percentage Impervious Area by the Use of Satellite Remote Sensing Imagery. *Water Science and Technology* **29**: 134-144.

Foody, G. M., 1999. Image Classification with a Neural Network: From Completely-Crisp to Fully-Fuzzy Situation. *In Advances in Remote Sensing and GIS Analysis*, P. M. Atkinson and N. J. Tate. Chichester, West Sussex, John Wiley & Sons Ltd.: 17-37.

Forster, B. C., 1985. An Examination of Some Problems and Solutions in Monitoring Urban Areas from Satellite Platforms. *International Journal of Remote Sensing* **6**(1): 139-151.

Friedl, M. A. and C. E. Brodley, 1997. Decision Tree Classification of Land Cover from Remotely Sensed Data. *Remote Sensing of Environment*, **61**: 399-409.

Graetz, R. D., 1990. Remote Sensing of Terrestrial Ecosystem Structure: An Ecologist's Pragmatic View. *In Ecological Studies: Remote Sensing of Biosphere Functioning*, R. J. Hobbs and H. A. Moony.: 79-85.

Gross, H. N. and J. R. Schott, 1998. Application of Spectral Mixture Analysis and Image Fusion Techniques for Image Sharpening. *Remote Sensing of Environment* **63**: 85-94.

Hansen, M., R. Dubayah and R. Defries, 1996. Classification Trees: an Alternative to Traditional Land Cover Classifiers. *International Journal of Remote Sensing* **17**(5): 1075-1081.

Jensen, J. R., 1996. *Introductory Digital Image Processing: a Remote Sensing Perspective*. Upper Saddle River, N.J., Prentice Hall.

Jensen, J. R., M. L. Bryan, S. Z. Friedman, F. M. Henderson, R. K. Holz, *et al.*, 1983. Urban/Suburban Land Use Analysis. *In Manual of Remote Sensing: Interoretation and Applications*, J. E. Estes and G. A. Thorley. Falls Church, Virginia, American Society of Photogrammetry. **2**: 1571-1666.

Jensen, J. R. and D. C. Cowen, 1999. Remote Sensing of Urban/ Suburban Infrastructure and Socio-Economic Attributes. *Photogrammetric Engineering & Remote Sensing* **65**(5): 603-610.

Lillesand, T. M. and R. W. Kiefer, 2000. *Remote Sensing and Image Interpretation*. New York, John Wiley & Sons, Inc.

Lo, C. P., 1995. Automated Population and Dwelling Unit Estimation from High-Resolution Satellite Images: a GIS Approach., *International Journal of Remote Sensing* **16**(1): 17-34.

Lo, C. P., D. Quattrochi and J. Luvall, 1997. Application of High-Resolution Thermal Infrared Remote Sensing and GIS to Assess the Urban Heat Island Effect. *International Journal of Remote Sensing* **18**(2): 287-304.

Madhavan, B. B., S. Kubo, N. Kurisaki and T. V. L. N. Sivakumar, 2001. Appraising the Anatomy and Spatial Growth of the Bangkok Metropolitan Area Using a Vegetation-Impervious-Soil Model through Remote Sensing. *International Journal of Remote Sensing* **22**(5): 789-806.

Mather, P. M., 1999. Land Cover Classification Revisited. *In Advances in Remote Sensing and GIS Analysis*, P. M. Atkinson and N. J. Tate. Chichester, West Sussex, John Wiley & Sons Ltd.: 7-16.

Novo, E. M. and Y. E. Shimabukuro, 1994. Spectral Mixture Analysis of Inland Tropical Waters. *International Journal of Remote Sensing* **15**(6): 1351-1356.

Peddle, D. R., F. G. Hall and E. F. LeDrew, 1999. Spectral Mixture Analysis and Geometric-Optical Reflectance Modeling of Boreal Forest Biophysical Structure. *Remote Sensing of Environment* **65**: 288-297.

Quattrochi, D. A., J. C. Luvall, D. L. Rickman, M. G. Estes, C. A. Laymon, *et al.*, 2000. A Decision Support Information System for Urban Landscape Management using Thermal Infrared Data. *Photogrammetric Engineering & Remote Sensing* **66**(10): 1195-1207.

Ridd, M., 1995. Exploring a V-I-S (Vegetation-Impervious Surface-Soil) Model for Urban Ecosystem Analysis through Remote Sensing: Comparative Anatomy of Cities. *International Journal of Remote Sensing* **16**(12): 2165-2185.

Roberts, D. A., G. T. Batista, J. L. G. Pereira, E. K. Waller and B. W. Nelson, 1998a. Change Identification Using Multitemporal Spectral Mixture Analysis: Applications in Eastern Amazonia. *In Remote Sensing Change Detection: Environmental Monitoring Applications and Methods*, R. S. Lunetta and C. D. Elvidge. Ann Arbor, MI, Ann Arbor Press: 137-161.

Roberts, D. A., M. Gardner, R. Church, S. Ustin, G. Scheer, *et al.*, 1998b. Mapping Chaparral in the Santa Monica Mountains using Multiple Endmember Spectral Mixture Model. *Remote Sensing of Environment* **65**: 267-279.

Rodenbeck, M., 1999. *Cairo: The City Victorious*. New York, Alfred A. Knopf.

Ryherd, S. and C. Woodcock, 1996. Combining Spectral and Texture Data in the Segmentation of Remotely Sensed Imagery. *Photogrammetric Engineering & Remote Sensing* **62**(2): 181-194.

Ryznar, R. M., 1998. Urban Vegetation and Social Change: an Analysis using Remote Sensing and Census Data. *In*., University of Michigan.

Strahler, A., C. Woodock and J. A. Smith, 1986. On the Nature of Models in Remote Sensing. *Remote Sensing of Environment* **20**: 121-139.

Thomson, C. N. and P. Hardin, 2000. Remote Sensing/GIS Integration to Identify Potential Low-income Housing Sites. *Cities* **17**(2): 97-109.

Tompkins, S., J. F. Mustard, C. M. Pieters and D. W. Forsyth, 1997. Optimization of Endmembers for Spectral Mixture Analysis. *Remote Sensing of Environment* **59**: 472-489.

Ward, D., S. R. Phinn and A. T. Murray, 2000. Monitoring Growth in Rapidly Urbanization Areas Using Remotely Sensed Data. *The Professional Geographer* **52**(3): 371-385.

Weber, C., 1994. Per-Zone Classification of Urban Land Cover for Urban Population Estimation. *In Environmental Remote Sensing from Regional to Global Scales*, G. M. Foody and P. J. Curran. Chichester, John Wiley & Sons Ltd.: 142-148.

Wilkinson, G. G., 1996. Classification Algorithms - Where Next? *In Soft Computing in Remote Sensing Data Analysis*, E. Binaghi *et al.* Singapore, World Scientific: 93-99.