# A FLEXIBLE METHOD FOR URBAN VEGETATION COVER MEASUREMENT BASED ON REMOTE SENSING IMAGES

Guobin Zhu, Fuling Bian, and Mu Zhang

*The Research Center of Spatial Information and Digital Engineering, Wuhan University, Luoyu Road 129#, Wuhan, 430079, China*

**Abstract:**
Traditional methods use NDVI to investigate vegetation cover from remote sensing imagery. These methods provide per-pixel vegetation distribution, and cause a modifiable areal unit problem (MAUP), when a meaningful statistical result is issued. In this paper, a new method based on advanced segmentation techniques and classification is proposed for urban vegetation investigation extraction. This method utilizes ASTER data to build a hierarchical multi-resolution structure, so as to reflecting the inherent relationship between ground features under various scale levels. By analyzing the hierarchical structure, a flexible measurement of urban vegetation cover index (VCI) is issued based on remote sensing imageries.

**Key Words:** NDVI, ASTER, segmentation, classification, vegetation cover index (VCI)

## INTRODUCTION

Urban vegetation investigation plays a very important role in urban planning, environmental protection, and consistency development policy making. Precise, reliable, and meaningful measurement of urban vegetation cover helps decision-makers and urban researchers to reach their goals.

The earliest method for urban vegetation investigation was manual statistics within a city boundary (Francis, 1987). The investigation results were usually shown by a total index or distribution indices in terms of a whole city or neighborhoods of the city, respectively. Obviously this method is low efficient, imprecise, and time consuming.

While remote sensing imagery is introduced into applied areas, NDVI, the normalized difference vegetation index, which is representative of plants' photosynthetic efficiency, has been widely used for presenting vegetation cover from different data sources (Lo and Faber, 1997; Roderick et al., 1996; van Leeuwen et al., 1999) and for various applications (Csiszar and Kerenyi, 1995; Dalezios et al., 2002; Gutman and Ignatov, 1997; Hess et al., 1996; Stow and Chen, 2002; Teillet et al., 1997). Conceptually, however, NDVI extraction from remotely sensed imageries is based on a per-pixel mechanism, which means the resulting NDVI is a regular grid-distributed expression associated with the specific resolution of the data source.

In reality, however, the ground features group themselves by means of inner characteristics, which are demonstrated in remote sensing imageries by digital numbers (DNs), textures, shapes, and so forth. This fact motivates researchers to extract more meaningful (or semantic) objects, rather than physical pixels, by using these spectral and spatial features. For example, Kressler et al. (2001) use the object-oriented techniques based on image segmentation and data fusion to perform land use classification. Kaminsky et al. (1997) build a neural network based classifier by using texture knowledge in remote sensing imagery.

In this paper, a new method based on advanced segmentation techniques and classification is proposed for urban vegetation investigation extraction. This method utilizes ASTER data to build a hierarchical multi-resolution structure, so as to reflecting the inherent relationship between ground features under various scale levels. By analyzing the hierarchical structure, a flexible measurement of urban vegetation cover index (VCI) is issued based on remote sensing imageries.

## METHODOLOGY

### Segmentation

The term segmentation relates to the grouping of image elements according to homogeneity (Cross and Mason, 1988; de Kok et al., 2000; Willhauck, 2000). According to this definition, the landscape is composed of a mosaic of irregularly shaped objects that are larger than individual pixels and which are homogeneous. The image segmentation procedures based on this kind of scene operate on the basic principle of minimizing the variance, or some other measure of internal homogeneity, within the regions identified (Haralick and Shapiro, 1985).

Through the process of segmentation, one image can virtually be reorganized by image objects, rather than pixels, within a specific resolution scale in which certain image objects will receive the best recognition. From another viewpoint, the segmentation procedure derives more information entropy associated with objects such as average brightness, spectral derivation, distance to other objects, etc., which benefits the image understanding from both spectral and spatial perspectives. The repetition of this processing in the same scene with different resolutions deepens the understanding of the objects organizing the image. In addition, the segmentation decreases the high frequency noise by spatial filtering.

The traditional segmentation on remote sensing imagery takes only account of spectral features referring to mere one image shift (single band) (Sabins, 1987). An advanced segmentation procedure proposed in this paper, however, is formed dependent upon the parameters of scale, DNs, and form concerning to multi-band combination (eCognition, 2001).

The scale parameter $H_{scale}$ is used to control the fractal degree of the segmentation. On the one hand, a segmentation procedure should produce highly homogeneous segments for the best separation and representation of image objects. On the other hand, the average size of image objects generated by segmentation must be adaptable to the scale of the features of interest in terms of specific image resolution and data acquisition. To this end, two basic criteria are used:

1. Average heterogeneity of image objects should be minimized.
2. Averagee heterogeneity of image objects weighted by their size should be minimized.

The DNs of pixels are very important elements in pixel-based classification. They also play essential roles in segmenting the homogeneous image objects in the segmentation procedure. Essentially, the assessment parameter of DNs ($H_{DN}$) are expressed by the sum of the standard deviation of spectral values ($\sigma_c$) in each of the bands ($c$) weighted with an empirical variable ($\omega_c$), which ranges between 0 to 1:

$$H_{DN} = \sum_c \omega_c \sigma_c \tag{1.}$$

According to the average heterogeneous minimization principle, the pixels with similar $H_{DN}$ parameter within the given threshold are grouped as an image object.

Form is another criterion dealing with spatial heterogeneity. It is used to assist the segmentation along with the other two criteria, scale and DNs, reducing the deviation from a compact or smooth shape state. Usually, it is described by the ratio of the object's perimeter ($l$) to the square root of the number of pixels ($\sqrt{n}$) forming the object. This parameter is defined as:

$$H_{form} = \frac{l}{\sqrt{n}} \tag{2.}$$

By adjusting the parameters above, hierarchical layers with different resolutions are formed, explicitly recognizing many scales in a landscape. This concept is similar to the piecewise homogeneous model in the assumption that the landscape is a mosaic of discrete objects. However, in the hierarchical multi-resolution structure there is an explicit hierarchy defining the relationship between these objects. The layers in the hierarchy represent the scales at which various processes are executed in the ground scene.

**MAUP**

Segmentation groups image pixels into various objects attached with specific meanings with respect to scale, DNs, and form. Generically, one can imagine that different segmentation results (image mosaics) will be generated by using different combination of these parameters. Thus, the results cannot be separated from the issue of the modifiable areal unit problem (MAUP).

The MAUP was firstly illustrated through the works of Openshaw and Taylor (1979), when they discovered that statistical results which include correlation of coefficients vary with scale. At a given scale level, results will vary when different zonal patterns are used. For example, in the case of residential segregation, differences in scale usually refer to different results from block groups and census tracts, and differences in zonal refer to inconsistent results from different definitions in what a block group or tract is. The former problem is known as the scale problem and the latter is known as the zonal or aggregation problem.

Because both problems are related to how enumeration units can be drawn, they are together labeled as the MAUP.

The advanced segmentation technique involves into the MAUP in two ways. The first one is associated with the scale effects by means of the scale parameter $H_{scale}$. The scale parameter indicates the average size of objects which have been clustered after the segmentation processing. Therefore, it implies the scale effects in the MAUP. Conceptually, when $H_{scale}$ to 1, the segmented image keeps the original forms, in which the segmented objects are pixels (or primitives). The second one is associated with the zoning effects by means of the parameters $H_{DN}$ and $H_{form}$ form. The $H_{DN}$ reflects the spectral characteristics of a remote sensing image with respect to the data sources, while $H_{form}$ influences the spatial forms of segmented objects. When $H_{DN}$ is set as a trigger to motivate the segmentation process, the form parameter $H_{form}$ is used to restrict the zoning effects of the segmented objects.

**Hierarchical structure of image segmentation**

By adjusting the parameters above, hierarchical layers with different resolutions are formed, explicitly recognizing many scales in a landscape. This concept is similar to the piecewise homogeneous model in the assumption that the landscape is a mosaic of discrete objects. However, in the hierarchical multi-resolution structure there is an explicit hierarchy defining the relationship between these objects. The layers in the hierarchy represent the scales at which various processes are executed in the ground scene.

The multi-resolution hierarchy is analyzed in both horizontal and vertical ways.

Vertically, the image objects on the upper layer (which holds coarser resolution) outline their semantic partners on the lower layer (which holds finer resolution). The objects on the lower layer, in turn, refine the semantically corresponding objects on the upper layer. Different ground features, due to their own physical characteristics, derive various interpretive outcomes associated with a variety of resolutions.

The benefits of this multi-resolution hierarchy are obvious. Since the abstract information on the upper layer is much stronger than on the lower one, this layer shields the spatial high frequency noise caused by too much physical detail on the lower layer. Visually, the “salt-and-pepper” effect is weakened. Actually this modification is achieved by semantic threads hidden within the connections between hierarchical layers, and thus forcing the image objects to be clustered, obeying the semantic threads from the bottom up. As a result, the reliability of recognition is promoted. The refining effective contribution by the lower layer, however, compensates for the lack of detailed physical characteristics on the upper layer, promoting the accuracy of the classification.

Horizontally, the segmentation procedure reshapes the pixel arrays into image objects with irregular borders, a completely different arrangement from original grid forms. In each layer, the distribution of the image objects is unique, making colorful scenarios to reduce the number of classes with similar semantic meanings on each layer. The distribution characteristics include neighboring, containing, distance, etc., which contribute greatly to object recognition. The other indices, such as spectral derivations, ratio between bands, shape index, etc., also assist with the object recognition, as well as with information extraction.

**Measurement of VCI**

The hierarchical multi-resolution structure can be used to enhance the connections between image objects, and to generate new information resources. By analyzing the relationship between objects, layers, and hierarchical structures, information such as an object’s neighboring characteristics, index based on relative areas between hierarchical layers, and even proximity based on distance, can be recognized.

There are two threads in a multi-resolution hierarchical structure, inherited from the upper layer to the lower layer. One of these is a physical thread, which is determined by scales, DNs, and form, assigned by operators. Another one is a logical thread, which is intended by operators as well, but with logical composition and decomposition to the desired classes, depending on their inherent relationships. Generally, the size of objects on the upper layer (usually with finer resolution) is less than the size of objects on the lower layer (usually with coarser resolution). The objects on the lower layer, however, have relatively logical connections (simply as containing) to the corresponding objects on the upper layer. This generates a possibility to calculate an index of a class, for example, vegetation cover, on a specific region by using a proportion of the objects of the class on the upper layer among the corresponding objects, which follow the same logical thread, on the lower layer. The formula is:

$$VCI_i = \frac{\sum_{j=1}^{k} A_j; (\forall A_j \in A_i)}{A_i} \qquad (3.)$$

where $VCI_i$ denotes an index of vegetation cover with respect to region $i$, $A_j$ refers to the area occupied by an object $j$, which is logically classified into vegetation class on the upper layer, and $A_i$ represents the area on the lower layer occupied by an semantic object, which contains the object on the upper layer (Fig. 1).

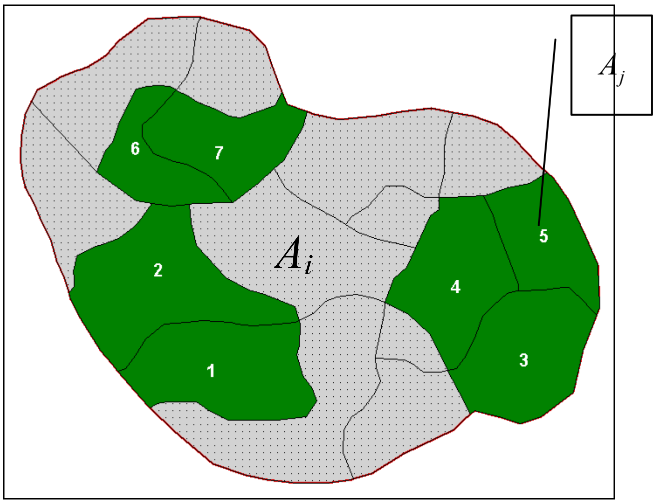
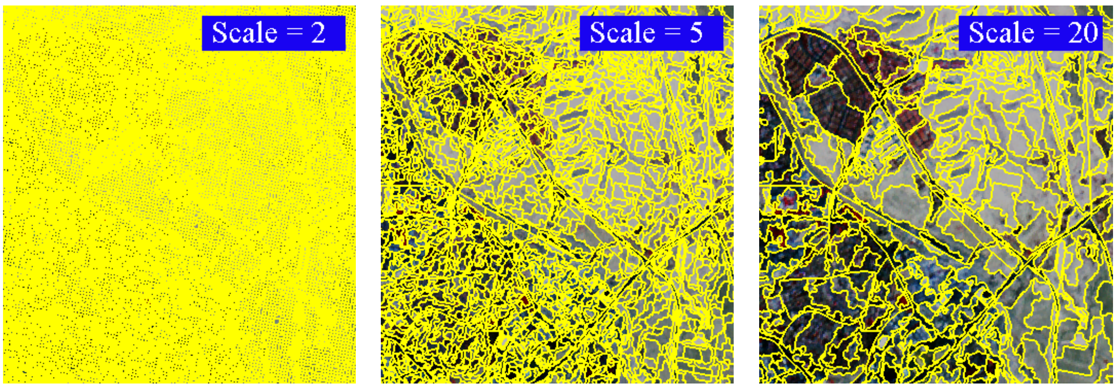



**Fig. 1. The geometric VCI illustration concerning two layers**

Ministry of International Trade and Industry. The ASTER instrument consists of three separate instrument subsystems (Abrams, 2000). Each subsystem operates in a different spectral region. The subsystems are the Visible and Near Infrared (VNIR, 0.52~0.86μm), the Short-wave Infrared (SWIR, 1.6~2.43μm), and the Thermal Infrared (TIR, 8.125~11.65μm) subsystems respectively. The case study area, Beer Sheva, Israel, is centered at 31˚15′46″N/34˚48′20″E (Lat/Long).

According to the aforementioned methodology, a multi-resolution hierarchy was first generated. In this research, four layers with different scales (2, 5, 20) (Fig. 2), smooth factors, and compact degrees were built to meet the different purposes of object extraction. The layer with scale parameter as 2, for example, was designed to realize a refine distinct between various classes, while scale as 5 was designed to fusion physical classification into logical groups in coarser resolution. Layer 3, with scale 20, was designed to separate urban morphologies in semantic ways (for example, residential areas).

The objects in this case were divided into 7 classes: impervious surface (which includes buildings and parking lots), roads and streets (which are generally treated as impervious surface but with obviously different spatial morphologies, for example, they are all linear features), grass area, tree mass area, soils, some special rooftops for industrial usages, and waterbodies. By introducing the municipal




Layer 1 Layer 2 Layer 3

**Fig. 2. The difference between the layers with various segmental parameters**

## EXPERIMENTS AND RESULTS

The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data is chosen as our data source in this study. ASTER is a cooperative effort between NASA and Japan's

control, each class was divided into a further level: one with an urban attribute, another with a non-urban attribute.

The VCI map was generated by using of the Equation (3). The VCI map is shown in Fig. 3.

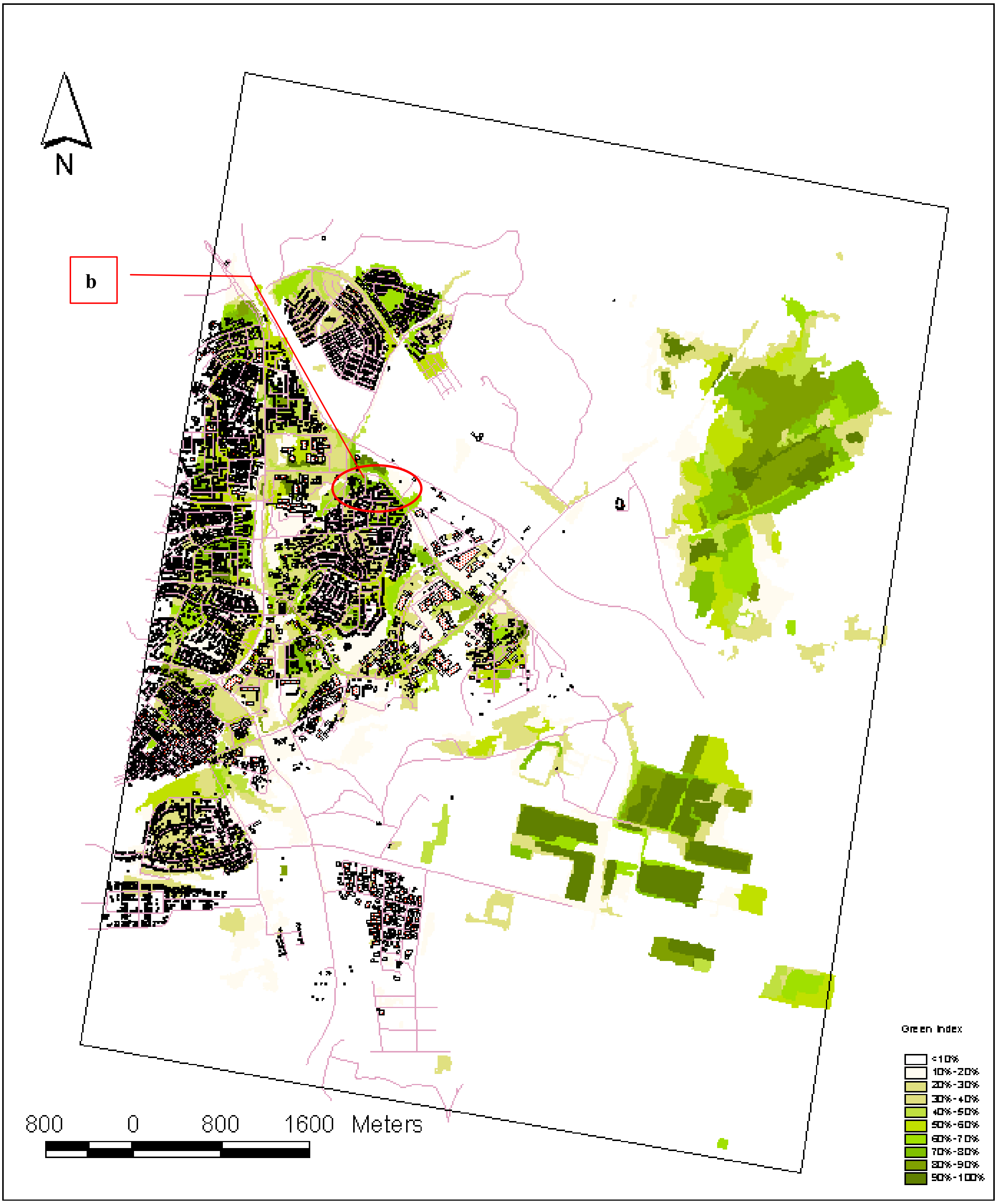



**Fig. 3. The Vegetation Cover Index (VCI) map of Beer Sheva, Israel**

## ANALYSIS, CONCLUSIONS, AND FUTURE WORK

Fig. 3 demonstrates a vegetation cover index distribution in part of Beer Sheva, Israel. To give a clear impression to the urban morphology, the building layer was loaded to overlay the VCI map. By analyzing inherent relationship between segmented layers and connections between logical objects, two points can be drawn as follows.

1. The VCI map generated based on remote sensing is an upscaling process. Generally, one can get statistical sources of vegetation cover of a city from government agencies. These sources are all based on statistical units (for example, neighborhood or sub-neighborhood). The numeric index cited from these materials describes an average degree referring to statistical terms. The index, however, cannot be used to represent a specific point within the statistical area through the use of a downscaling process, because of the MAUP. The VCI generated by the method proposed in this paper, however, avoids the problem. The average size of objects reflects specific semantic meanings in terms of properties of data sources, such as, resolution, wavelength, textures, and etc. By using data aggregation, information based on an object level is collected to indicate the overall degree reaching a neighborhood level. Not only is the result accurate, it is reliable, therefore, safe.

2. The size of objects used for the VCI calculation is flexible. Usually, the statistical indices are based on a fixed geographical unit, for example, a neighborhood. The VCI generated in this paper, however, is based on flexible, sometimes intelligent, processing. The farmlands outside Beer Sheva, for example, by virtue of the fact that they are pure vegetation, are treated as a 100% vegetation index in Fig. 3 (*see* **a**), with both big sizes of objects for calculation. Another example is the grass areas on the campus of Ben-Gurion University, as shown in Fig. 3 (*see* **b**). These areas are scattered among an impervious surface on the campus. In a coarser resolution, they are treated as an impervious surface, while in a finer resolution, they are treated as grass areas. Not surprisingly, when a thematic map (for example, the neighborhood map used in our case) is used as a denominator (Equation (3)), the index will reach a thematically statistical result.

The MAUP describes a generic problem associated with scale and zoning effects. The method proposed in this paper provides a routine to reach the goal of solving the MAUP by remote sensing image processing in the case of urban vegetation cover extraction. The experiments in this paper, however, are not sufficient. To study the scale dependency of VCI in remote sensing imagery, a series of statistical results along with sequence of various average sizes of objects should be issued to generate a pyramid of observation of vegetation cover. By analyzing the inherent relationship of VCI against the scale changing and zoning pattern shifting, a more comprehensive result will be given in the future work.

## ACKNOWLEDGEMENTS

The authors would like to thank Dr. Philip Christensen, Principal Investigator of the ASTER Urban Environmental Monitoring Program at Arizona State University, for providing the newly acquired ASTER data of Beer-Sheva for our study work. The authors also wish to acknowledge the support of the Israel Committee for Higher Education and China Scholarship Council for providing a Ph.D. scholarship for the first author.

## REFERENCES

Abrams, M., 2000. The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER): Data products for the high spatial resolution imager on NASA's Terra platform. International Journal of Remote Sensing, 21(5): 847-859.

Cross, A.M. and Mason, D.C., 1988. Segmentation of remote-sensed images by split-and-merge process. International Journal of Remote Sensing, 9(8): 1329-1345.

Csiszar, I. and Kerenyi, J., 1995. The effect of the vegetation index on the daily variation of the active surface temperature. Advanced Space Research, 16(10): 177-180.

Dalezios, N.R., Loukas, A. and Bampzelis, D., 2002. Assessment of NDVI and agrometeorological indices for major crops in central Greece. Physics and Chemistry of the Earth, 7: 1025-1029.

de Kok, R., Buck, A., Schneider, T. and Ammer, U., 2000. Analysis of Image Objects from VHR Imagery for Forest GIS Updating in the Bavarian Alps, Geoinformation for All. XIX Congress of the International Society for Photogrammetry and remote Sensing (ISPRS). Geomatics Information & Trading Center (GITC) bv, Amsterdam, The Netherlands, pp.

222-229.
eCognition, 2001. eCognition User Guide. Definjens Imaging, Munchen, Germany, 487 pp.
Francis, M., 1987. Urban Open Spaces. In: E.Z.G. Moore (Editor), Advances in Environment, Behavior and Design. Plenum, New York, pp. 71-106.
Gutman, G. and Ignatov, A., 1997. Satellite-derived green vegetation fraction for the use in numerical weather prediction models. Advanced Space Research, 19(3): 477-480.
Haralick, R.M. and Shapiro, L.G., 1985. Image segmentation techniques. Computer Vision, Graphics, and Image Processing, 29: 100-132.
Hess, T., Stephens, W. and Thomas, G., 1996. Modelling NDVI from decadal rainfall data in the North East Arid Zone of Nigeria. Journal of Environmental Management, 48: 249-261.
Kaminsky, E.J., Barad, H. and Brown, W., 1997. Textural neural network and version space classifiers for remote sensing. International Journal of Remote Sensing, 18(4): 741-762.
Kressler, F.P., Bauer, T.B. and Steinnocher, K.T., 2001. Object-oriented per-parcel land use classification of very high resolution images, IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, pp. 164-168.
Lo, C.P. and Faber, B.J., 1997. Integration of Landsat Thermatic Mapper and census data for quality of life assessment. Remote Sensing of Environment, 62: 143-157.
Openshaw, S. and Taylor, P.J., 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In: N. Wrigley (Editor), Statistical Applications in the Spatial Science, Pion, London, pp. 127-144.
Roderick, M., Smith, R. and Lodwick, G., 1996. Calibrating long-term AVHRR-derived NDVI imagery. Remote Sensing of Environment, 58: 1-12.
Sabins, F.F., 1987. Remote Sensing Principles and Interpretation. W. H. Freeman and Company, New York, 449 pp.
Stow, D.A. and Chen, D.M., 2002. Sensitivity of multitemporal NOAA AVHRR data of an urbanizing region to land-use/land-cover changes and misregistration. Remote Sensing of Environment, 80: 297-307.
Teillet, P.M., Staenz, K. and Williams, D.J., 1997. Effects of spectral, spatial, and radiometric characteristics on remote sensing vegetation indics of forested regions. Remote Sensing of Environment, 61: 139-149.
van Leeuwen, W.J.D., Huete, A.R. and Laing, T.W., 1999. MODIS vegetation index compositing approach: A prototype with AVHRR data. Remote Sensing of Environment, 69: 264-280.
Willhauck, G., 2000. Comparison of object oriented classification techniques and standard image analysis for the use of change detection between SPOT multispectral satellite images and aerial photos, XIX Congress of the International Society for Photogrammetry and remote Sensing (ISPRS). Geomatics Information & Trading Center (GITC) bv, Amsterdam, The Netherlands, pp. 214-221.