DETECTING INFORMAL SETTLEMENTS FROM IKONOS IMAGE DATA USING METHODS OF OBJECT ORIENTED IMAGE ANALYSIS – AN EXAMPLE FROM CAPE TOWN (SOUTH AFRICA)

Peter HOFMANN
-
DEFiNiENS AG
Rindermarkt 7
D-80331 Munich
Germany
e-mail: PHofmann@definiens.com
http: www.definiens.com

ABSTRACT

Detecting informal settlements might be one of the most challenging tasks within urban remote sensing. This phenomenon occurs mostly in developing countries. In order to carry out the urban planning and development tasks necessary to improve living conditions for the poorest world-wide, an adequate spatial data basis is needed (see Mason, O. S. & Fraser, C. S., 1998). This can only be obtained through the analysis of remote sensing data, which represents an additional challenge from a technical point of view. Formal settlements by definition are mapped sufficiently for most purposes. However, this does not hold for informal settlements. Due to their microstructure and instability of shape, the detection of these settlements is substantially more difficult. Hence, more sophisticated data and methods of image analysis are necessary, which ideally act as a spatial data basis for a further informal settlement management. While these methods are usually quite labour-intensive, one should nonetheless bear in mind cost-effectivity of the applied methods and tools. In the present article, it will be shown how eCognition™ can be used to detect and discriminate informal settlements from other land-use-forms by describing typical characteristics of colour, texture, shape and context. This software is completely object-oriented and uses a patented, multi-scale image segmentation approach. The generated segments act as image objects whose physical and contextual characteristics can be described by means of fuzzy logic. The article will show methods and strategies using eCognition™ to detect informal settlements from high resolution space-borne image data such as IKONOS. A final discussion of the results will be given.

1 INTRODUCTION

Regarding the occurrences of settlement areas in remote sensing data in the most cases to describe their different types regarding the pixels' spectral information only is insufficient. Furthermore characteristics such as texture, shape or contextual information is necessary to outline and describe them adequately. Especially informal settlement areas hold typical textural and structural information which is mostly determined by their informal status. Mason, O. S. & Fraser, C. S. (1998) describe informal settlements as *"... dense settlements comprising communities housed in self-constructed shelters under conditions of informal or traditional land tenure ... . They are a common feature of developing countries and are typically the product of an urgent need for shelter by the urban poor. As such they are characterised by a dense proliferation of small, makeshift shelters built from diverse materials (such as plastic, tin sheeting and wooden planks), by degradation of the local ecosystem (for example, erosion and poor water quality and sanitation) and by severe social problems."* Principally the physical entities of informal settlements (small, makeshift shelters built from diverse materials), as a result of the social circumstances the inhabitants live in, can be detected from remote sensing data. Hence, depending on the sensor's spatial and spectral resolution, it should be possible in principal to classify and distinguish these settlement areas from other landuse or settlement forms. Additionally the described ecological circumstances could be detectable as well and act as additional contextual information. Depending on the imaging scale and the scale of spatial information to obtain, detecting informal settlements can vary in its aim: on the one hand it is necessary to detect single shacks and on the other hand it is sufficient just to locate and outline the informal settlement areas spatially. Using pure IKONOS image data a single shack detection is not sufficiently feasible due to the sensor's spatial resolution. Therefore additional elevation information gives beneficial ancillary data (see Mason, O. S. & Fraser, C. S., 1998). Thus, the article present will focus on the location of informal settlement areas and their boundaries.

JÜRGENS, Carsten (2001) (Editor): Remote Sensing of Urban Areas/ Fernerkundung in urbanen Räumen.
(= Regensburger Geographische Schriften, Heft 35). Regensburg. ISBN 3-88246-222-1

2 DATA AND PRE-PROCESSING

As data source a subset of an IKONOS scene was used showing parts of Nyanga and Crossroads which are parts of the city of Cape Town (South Africa) and located south to Cape Town's airport. The subset shows several forms of settlement areas beneath other typical urban res. suburban landuse forms such as roads, larger buildings

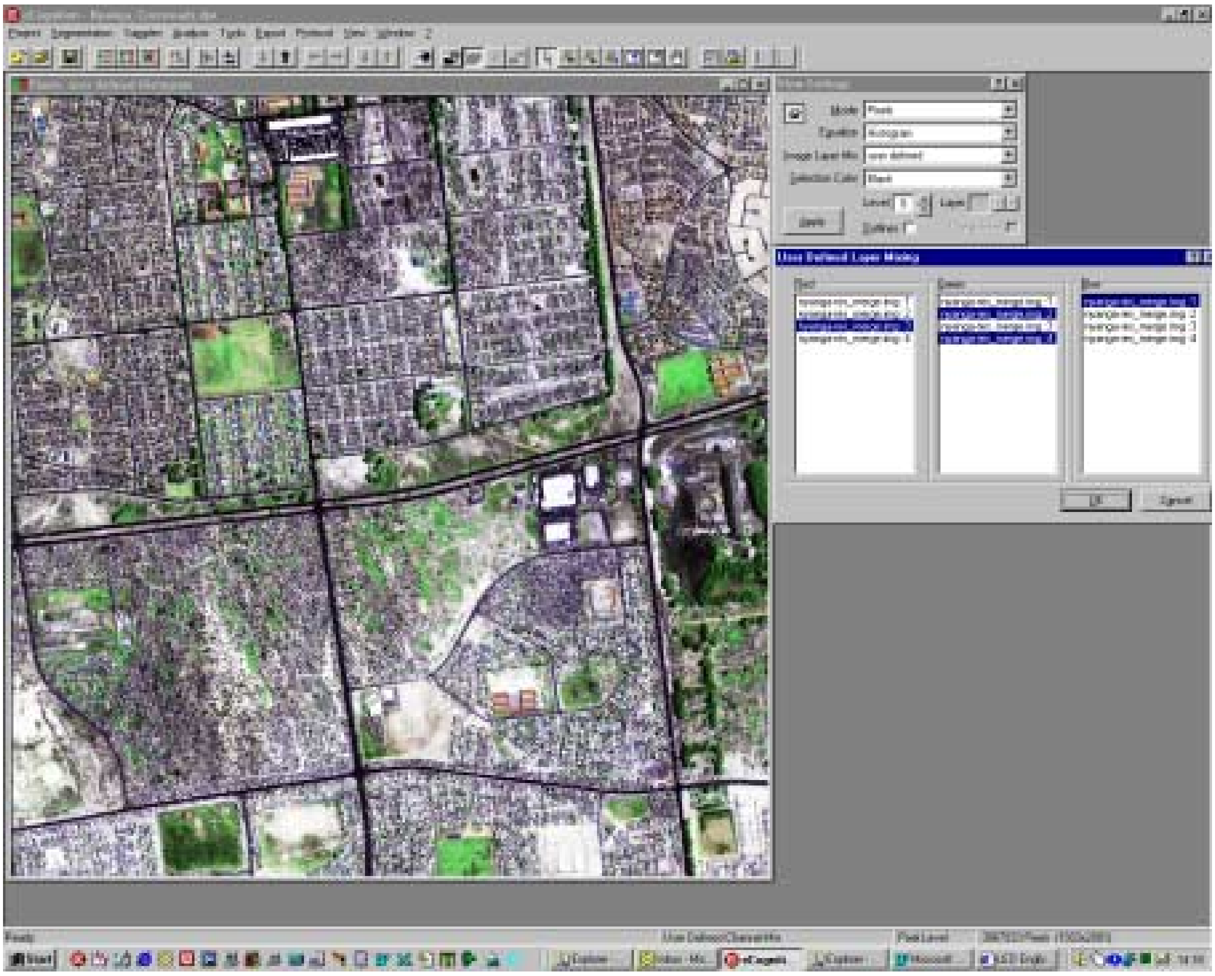

Fig. 1: Enhanced IKONOS subset. Combination of channels: red = red, green = green and near infrared, blue = blue.

and fallow land with more or less vegetation (see Fig. 1).
Before analysing the image with eCognition it was enhanced by applying a principal-component pan-sharpening method. This method makes it possible to benefit from the sensors spectral capabilities simultaneously with its high spatial resolution. Thereby the first principal component of the four spectral IKONOS channels (4m resolution) was substituted by the 1m resolution IKONOS panchromatic channel. This new combination of principal components then was re-transformed applying an inverse principal components transformation. Fig. 1 shows the result of the histogram equalised, re-transformed image. The chosen channel combination denotes active vegetation areas by an exaggerated green since the both vegetation-sensitive channels (IKONOS channel 2 and 4) are represented in green.

Focusing on settlement areas, the following forms can be seen in the image:

- Formal settlement areas with small gardens of mostly inactive vegetation in the north-western and central part of the image. These areas are characterised by a more or less rectangular layout of smaller roads which are obviously following a parcelling scheme given by land ordinance. The single houses are more or less regularly ordered and located in the centre of the parcels. Their roofs vary in colour from dark brown to dark grey. The houses' size can be estimated from $60m^2$ – $80m^2$. In the further proceeding these areas will be called *medium formal*.
- In the north-western part, between the *medium formal* settlement areas, similar formal settlement areas can be seen which contain more and larger areas of active vegetation (obviously gardens). The houses are

JÜRGENS, Carsten (2001) (Editor): Remote Sensing of Urban Areas/ Fernerkundung in urbanen Räumen.
(= Regensburger Geographische Schriften, Heft 35). Regensburg. ISBN 3-88246-222-1

slightly larger then those of *medium formal* settlement areas and the spacing between the houses is larger as that of *medium formal* settlement areas. The roofs' colour varies from a bluish grey to a bright red. These areas will be called *settlement areas with gardens* in the further proceeding.

- In the northern-central and the north-eastern central part of the image settlement areas which are similar to *medium formal* settlement areas can be found. In contrast to the first named settlement forms the houses of these areas stand more closely and are slightly smaller. The roofs' colour is mostly grey. The houses of the north-eastern central part are more elongated and surrounded by irregular ordered, smaller objects[1]. Within these areas only poor vegetation can be seen. These settlement areas will be called *dense formal* in the further proceeding.
- In the north-eastern part a newly created or upgraded settlement area can be recognised. The houses are small and stand mostly narrow and regularly ordered. Between the housing areas some areas of fallow land can be seen as well as some smaller informal settlement areas. The colour of the roofs varies from bright bluish grey to brown and red. In some cases a very bright facade can be seen. These areas will be called *new settlement areas* in the further proceeding.
- In the southern part of the image settlement forms with small houses with an estimated size of $40m^2$ – $70m^2$ can be discovered. The roofs' colour occurs mostly white or bright grey to bright blue. The houses are more or less ordered irregularly but within a clearly structured road network. In some cases small areas of vegetation can be discovered, as well as several mostly small areas of informal settlements. These areas will be called *bright formal settlement* in the further proceeding.
- Larger informal settlement areas can be seen in the northern centre in the south west and in the south east of the image. While the south western informal settlement area is obviously the largest the others are of smaller size. All areas show irregularly ordered, small houses of strongly varying size from estimated $25m^2$ – $40m^2$ in average (shacks). Within the settlement areas no road network can be discovered but within the larger area a system of irregular pathways can be seen. While in the larger informal settlements in many cases single shacks can be seen, in the smaller areas only an agglomeration of shacks is detectable. The shacks' colours vary from dark brown to dark grey. The informal settlements' texture looks more dense and irregular than that of formal settlements. According to their main colour the informal settlement areas will be called *dense*, *medium* or *bright informal*.

It is noticeable that all settlement areas are bordered by roads. Thus roads act as border lines in terms of spectral properties and in terms of semantics to outline different settlement areas.

3 CONCEPTS AND STRATEGIES USING ECOGNITION

eCognition's segmentation approach allows to generate image objects on an arbitrary number of scale-levels taking into account criterions of homogeneity in colour and shape. Thereby a hierarchical network of image objects is generated wherein each object knows its neighbouring objects in horizontal and vertical direction (see Baatz, M. & Schäpe, A., 1999; Baatz, M. & Schäpe, A., 2000; deKOK, R. et al., 1999; Fig. 2).

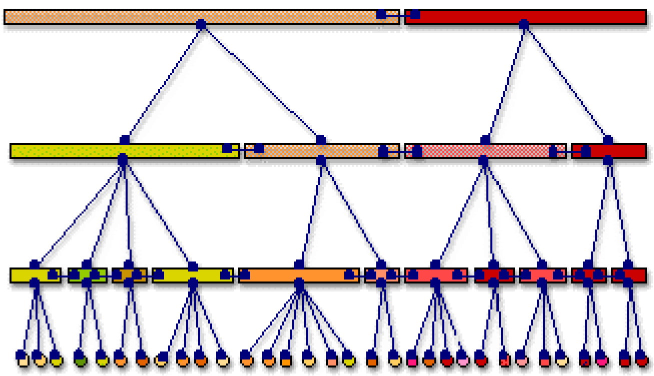

Fig. 2: Hierarchical network of image objects.

Basic aim of the segmentation process should be to generate as meaningful objects as possible. This means that the shape of each object in question should ideally be represented by an according image object. This shape combined with further derivative colour and texture properties can be used to initially classify the image by classifying the generated image objects. Thereby the classes are organised within a class hierarchy. Each class can have a sub- or super-class and thus inherit its properties from one or more super-classes or to its sub-class(es). In a second step additional semantic information can be used to improve the image classification. With

[1] Obviously informal settlements within formal settlement areas.

respect to the multi-scale behaviour of the objects to detect, a number of small objects can be aggregated to form larger objects constructing a semantic hierarchy. Likewise, a large object can be split into a number of smaller objects which basically leads to two main approaches of image analysis: A top-down and a bottom-up approach (see Hofmann, P. & Reinhardt, W., 2000). In eCognition both approaches can be realised performing the following steps:

- Creating a hierarchical network of image objects using the multi-resolution segmentation. The upper-level image segments represent small-scale objects while the lower-level segments represent large-scale objects.
- Classifying the derived objects by their physical properties. This also means that the class names and the class hierarchy are representative with respect to two aspects: the mapped real-world and the image objects' physically measurable attributes. Using inheritance mechanisms accelerates the classification task while making it more transparent at the same time.
- Describing the (semantic) relationships of the network's objects in terms of neighbourhood relationships or being a sub- or super-object. This usually leads to an improvement of the physical classification res. the class hierarchy.
- Aggregating the classified objects to semantic groups which can be used further for a so called 'classification-based' segmentation. The derived contiguous segments then can be exported and used in GIS. The semantic groups can also be used for further neighbourhood analyses.

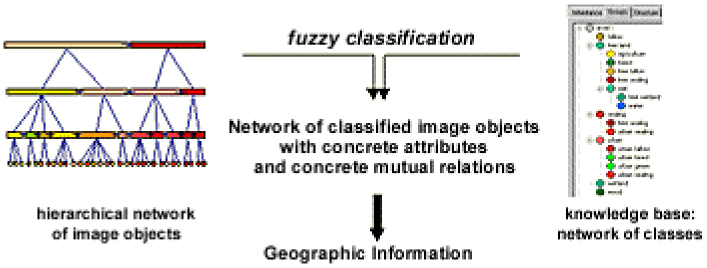

Fig. 3: Usual proceeding when working with eCognition.

These steps describe the usual proceeding when working with eCognition. While the first two steps are a mandatory, the latter two steps may be advisable according to the user's objectives and content of the image.

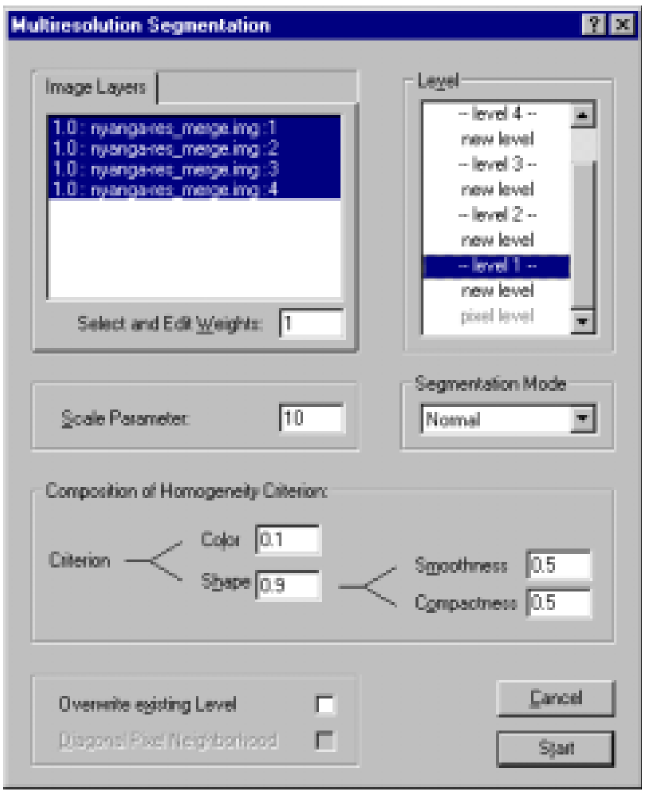

Fig. 4: Dialog for image segmentation with adjustable parameters.

3.1 IMAGE SEGMENTATION

- As the generated image segments act as the basic building blocks for the later-on classification an appropriate image segmentation is crucial. In eCognition image segmentation is controlled by the following parameters (see Fig. 4; for details see Baatz, M. & Schäpe, A., 1999b; DEFiNiNENS AG, 2000): Weight of image channels: this parameter can be used to more or less weight one or more image channels' influence on the object generation. When working with image data of comparable channels in size and content such as IKONOS each channel should be weighted equally.
- Scale parameter: this parameter indirectly influences the average object size. In fact this parameter determines the maximal allowed heterogeneity of the objects. The larger the scale parameter the larger the objects become.
- Color/Shape: with these parameters the influence of colour vs. shape homogeneity on the object generation can be adjusted. The higher the shape criterion the less spectral homogeneity influences the object generation.
- Smoothness/Compactness: when the shape criterion is larger than 0 the user can determine whether the objects shall become more compact (fringed) or more smooth.
- Level: determines whether a new generated image level will either overwrite a current level or whether the generated objects shall become sub- or super-objects of a still existing level. The order of generating the levels affects the objects' shape (top-down vs. bottom-up segmentation).

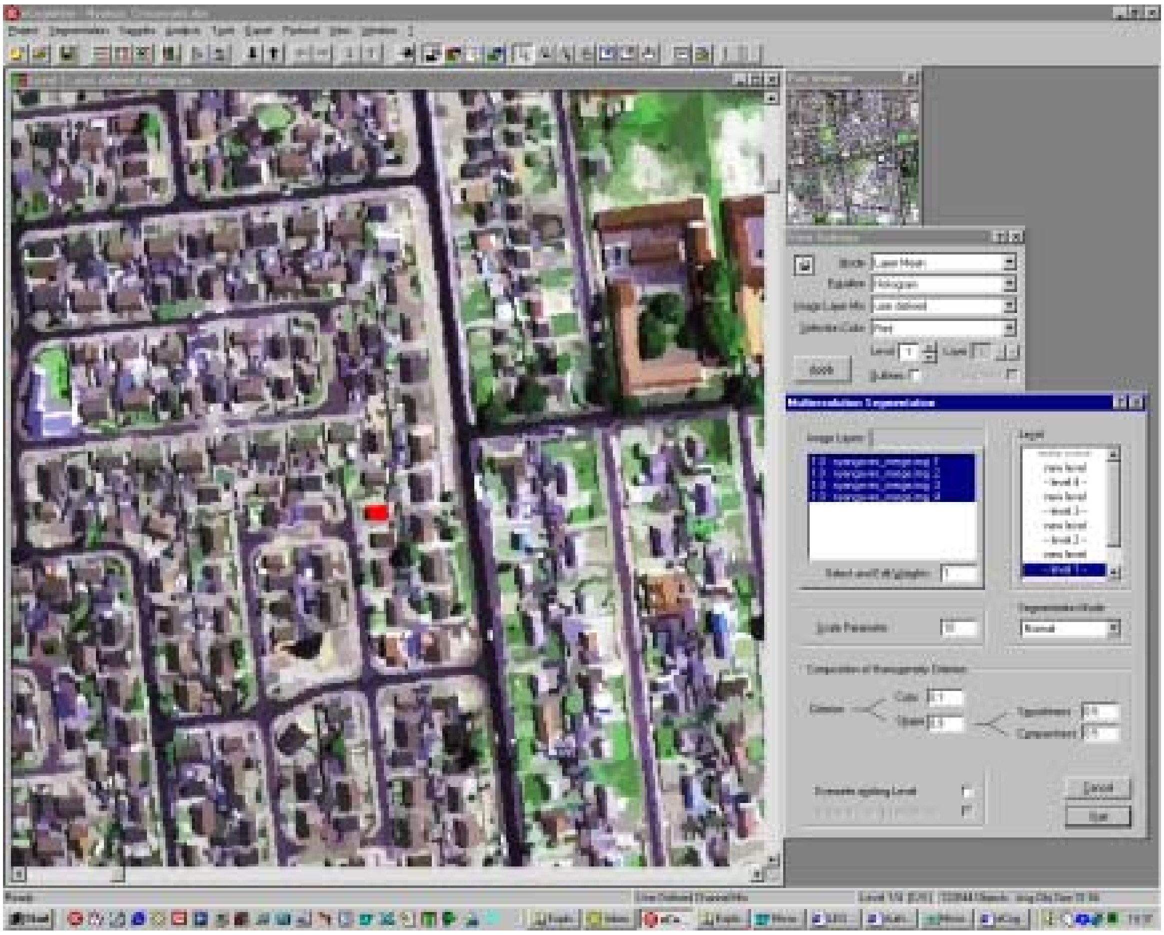

Fig. 5: Segmentation result of the lowermost level (level).

For the example present a bottom-up approach for the image segmentation has been applied. In addition using a value of 0.9 for shape and a value of 0.5 for smoothness and compactness has been experienced as well suited to obtain meaningful objects on the lowermost level (see Fig. 5). Within formal settlement areas many houses or meaningful parts of houses (e.g. brighter and darker roofs) are outlined by the segments of the lowermost level. Beside these house objects a variety of other objects is generated which can more be seen as less meaningful sub-objects of larger objects (e.g. parts of a garden or impervious areas). To obtain larger meaningful objects, further segmentations with coarser objects have been performed. With respect to the classification aim only the top-most level showed meaningful objects (see Fig. 6). On this level the generated large image objects mostly represent different settlement areas or parts of them. Other landuse forms (fallow lands and larger buildings) are extracted as well. Roads are mostly represented by dark, large and elongated objects.

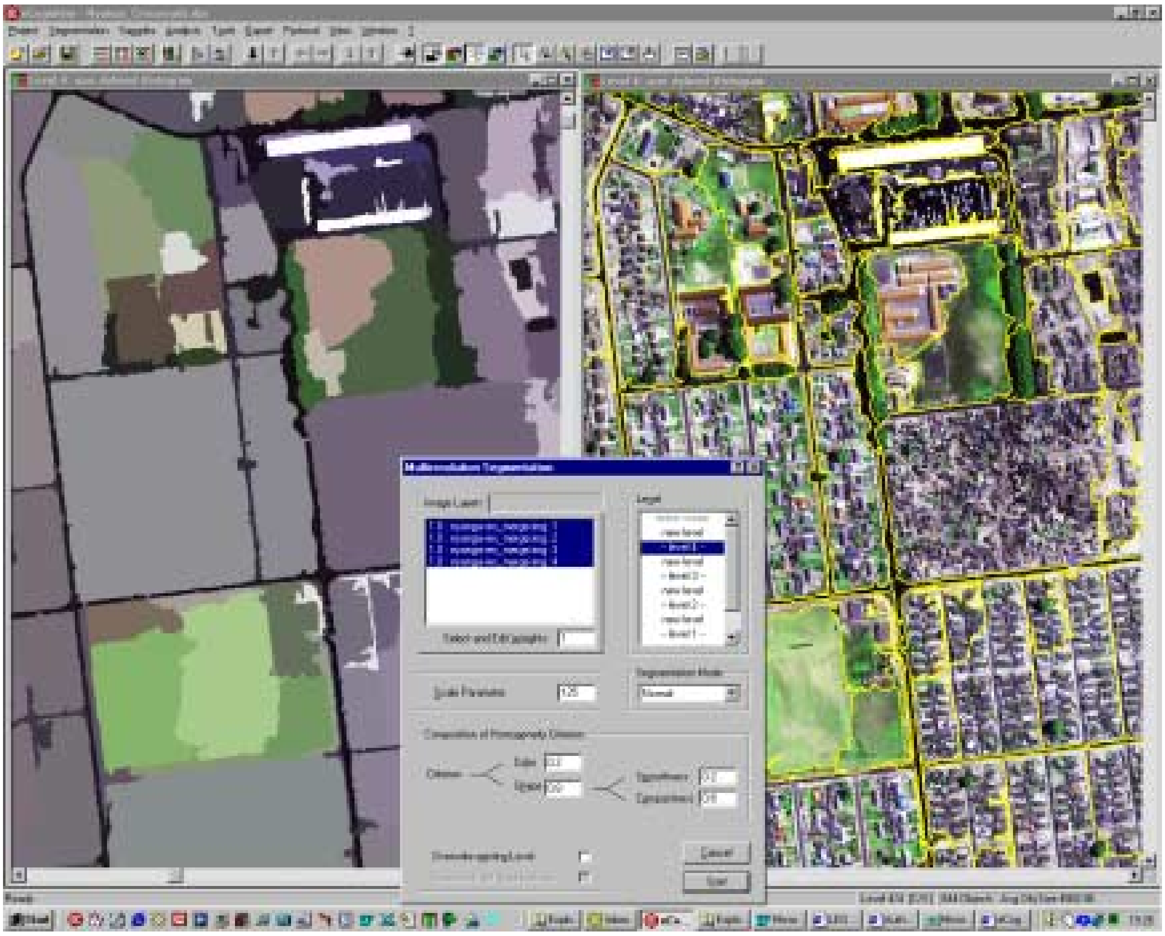

Fig. 6: Segmentation of the topmost level (level 4). Left: Segments coloured by their spectral mean values. Right: Segments' outlines. Middle: Used segmentation parameters.

3.2 CLASS HIERARCHY AND IMAGE CLASSIFICATION

Generating a class hierarchy in eCognition can be understood as generating a rule base wherein the user determines physical and semantic properties typical for the objects of a certain class. Therefore the software offers two basic classifiers: a nearest neighbour classifier and fuzzy membership functions. Both act as class descriptors. While the nearest neighbour classifier describes the classes to detect by sample objects for each class which the user has to determine, fuzzy membership functions describe intervals of feature characteristics wherein the objects do belong to a certain class or not by a certain degree (see Fig. 8). Thereby each feature offered by eCognition can be used either to describe fuzzy membership functions or to determine the feature space for the nearest neighbour classifier. A class then is described by combining one or more class descriptors by means of fuzzy-logic operators or by means of inheritance or a combination of both (see Fig. 8). As the class hierarchy should reflect the image content with respect to scale the creation of level classes is very useful. These classes represent the generated levels derived from the image segmentation and are simply described by formulating their belonging to a certain *level*. Classes which only occur within these levels inherit this property from the level classes. This technique usually helps to clearly structure the class hierarchy. In the example present the class *Level 1* acts as the level-super-class for all classes occurring in the lowermost level. In addition classes occurring in the lowermost level can be seen as semantic sub-objects of the top-level classes (e.g. houses are sub-objects of settlement areas). Since on the top-level several settlement areas are outlined, describing the relationships between (classified) sub- and super-objects can be beneficial for identifying different settlement

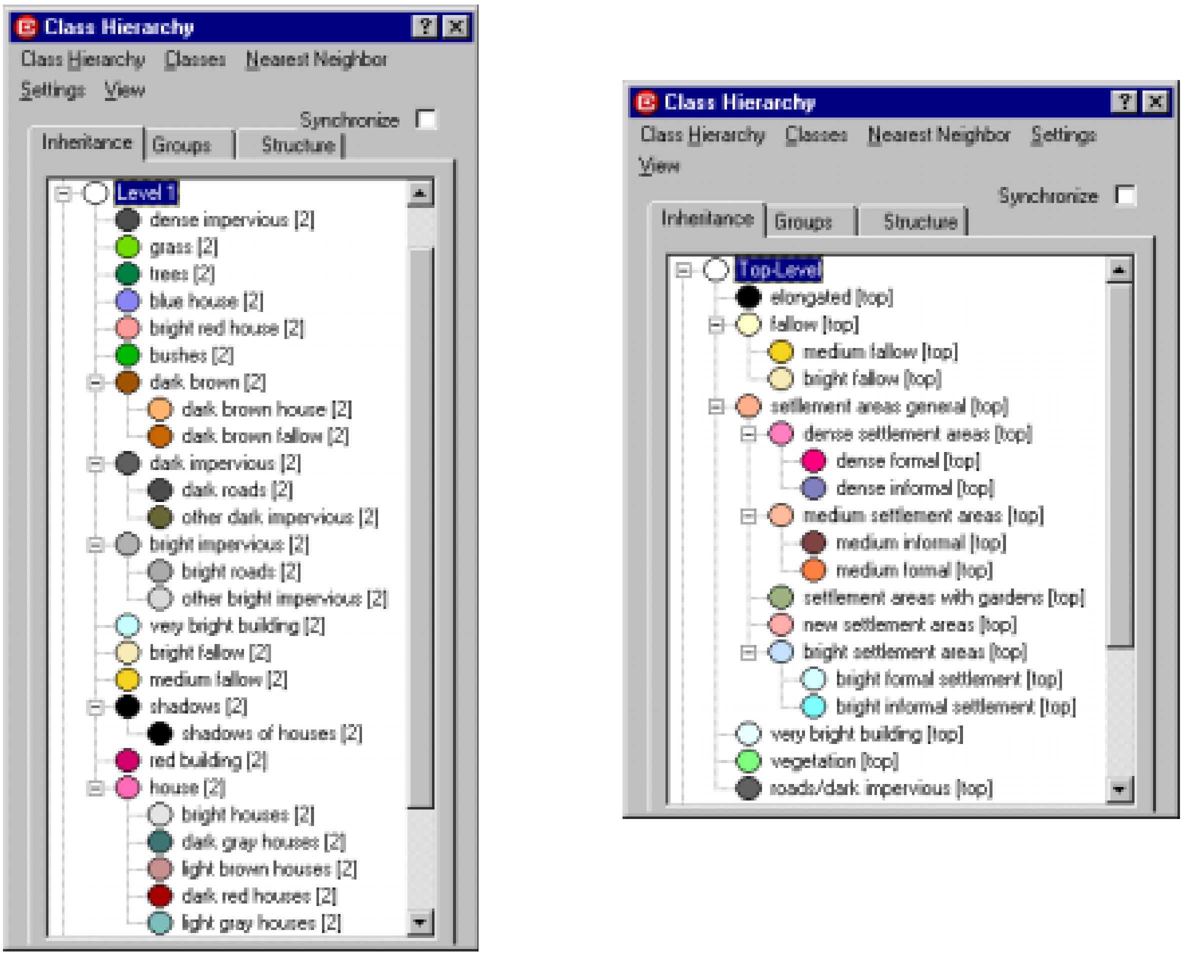

Fig. 7: Class hierarchy with Level super-classes and appropriate sub-classes.

areas. Beginning with a spectral nearest neighbour classification a stepwise refinement using inheritance mechanisms and form criteria has been applied to the fine segmented level. Although in many cases the roofs of houses are well outlined on this level a sufficient classification of single houses was not possible. The reason lies in the chosen segmentation parameters: a strong weighting of shape as homogeneity criterion leads to well fitting outlines of houses but at the same time comparably shaped segments of impervious or fallow land. Thus in many cases houses have been mis-classified as impervious or fallow and vice versa (see Fig. 9). However, having in mind that the identification and classification of settlement areas does neither need the exact numeration nor the exact shape of houses an accurate extraction of houses is not necessary. Estimating the relative area of a certain type of house (e.g. *dark red houses*) within a settlement area can be sufficient to determine the type of settlement area. In fact textural features are more suited to detect and distinguish settlement areas. In eCognition texture can be described regarding colour and shape of the objects' sub-objects. To identify and classify the different forms of settlement areas the objects of the top most level have been classified by describing on the one hand their typical colour and form and on the other hand their texture and their relations to classified sub-objects. Thereby an initial class hierarchy was generated wherein the super-classes describe the image content very roughly (Fig. 7). From these super-classes sub-classes have been generated whose details of description raise the deeper the class hierarchy is. In the example present the class *elongated* is just described by form criteria and represents objects like wider roads or other elongated non-settlement areas. The classes *fallow*, and *very bright building* have a similar function. They are described by textural features (*mean area of sub-objects*, *standard deviation in the near infrared channel*, and *average mean difference to neighbours of sub-objects in the near infrared channel*[2]). The class *roads/dark impervious* is simply described by a spectral nearest neighbour classifier whereas the class *vegetation* takes additional contextual information into account: the *relative area of sub-objects* classified as *vegetation* must not be lower than 50% (0.5). *Settlement areas general* acts as a super-class for all other settlement areas. It describes the typical texture of settlement areas by the *standard deviation in the near infrared channel* and *average mean difference to neighbours of sub-objects in the near infrared channel.* To more clearly differentiate the image the class *fallow* and *settlement areas general* have been split. Although

2 can be understood as the overall contrast of sub-objects in this channel

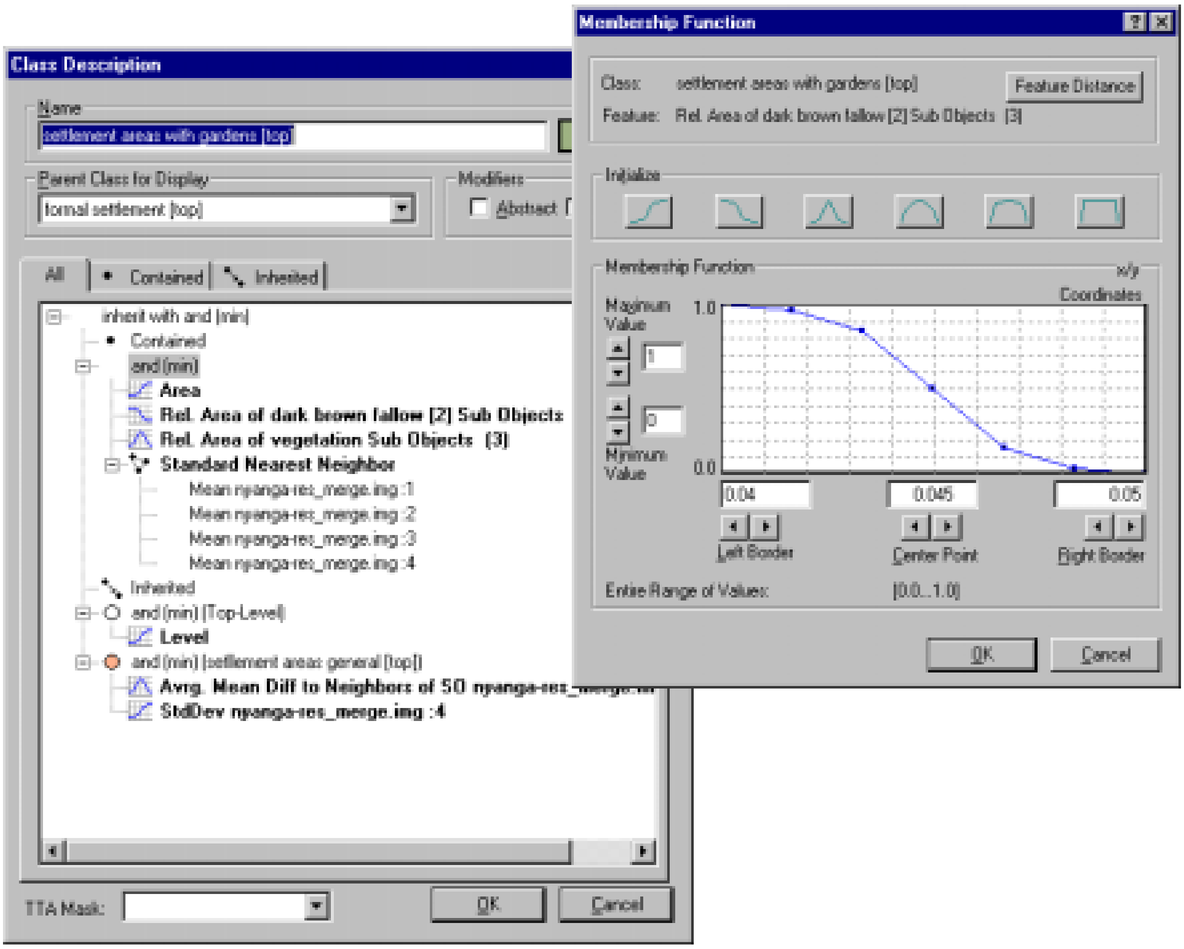

Fig. 8: Example of a class and feature description.

splitting *fallow* might not be necessary in terms of detecting informal settlement areas, to distinguish *fallow* into *bright fallow* and *medium fallow* regarding their spectral entities helps to avoid misclassifications of settlement areas. *Settlement areas general* was split according to the settlement forms described in chapter 2 into the classes:

- *dense settlement areas*
- *medium settlement areas*
- *settlement areas with gardens*
- *new settlement areas*
- *bright settlement areas*

Thereby the classes are basically described by their spectral properties. The classes *settlement areas with gardens* was refined by describing its area (not less than 10000 pixels[3]) and the relative area of *vegetation* sub-objects (between 24% and 32%) and *dark brown fallow* sub-objects (not more than 5%). *New settlement areas* has been refined by adjusting the relative area of sub-objects classified as *dark red-houses* to more than 2%. To distinguish between formal and informal settlement areas the classes *dense settlement areas, medium settlement areas and bright settlement areas* were split into according formal and informal settlement areas (Fig. 7). In the most cases the shape of sub-objects (mean value of asymmetry of sub-objects) gave a well discriminating feature. Thereby asymmetry expresses the (sub-) objects' elongation: the more close to 1 the value is, the more elongated an object is and the more close to 0 the more square an object is. For the most informal settlement areas this value is not greater than 0.59 to 0.61 except at roadside areas. In contrast in formal settlement areas a value of 0.60 to 0.70 reflects the more regular ordered texture of these areas - mostly given by the road network within these areas. *Medium informal settlements* were determined by describing their fringes using the shape index (greater than 1.8; for details how this value is calculated see DEFiNiNENS AG, 2000). Additionally the average size of sub-objects in level 1 (less than 27.5 pixels for informal settlement areas) reflects the typical texture for dense and medium informal settlements. Note: it is not the average size of shacks which is calculated. However, within *bright settlement areas* the area of sub-objects in informal settlement areas is close to that of

[3] as IKONOS has a resolution of 1m per Pixel this corresponds to 10000m^2.

formal settlement areas. This is explainable by the small bright houses surrounded by gardens with little vegetation. Thus to distinguish between formal and informal settlements in bright settlement areas is hard to perform even by human eye. In addition in some bright fallow areas the pattern of former informal settlements are still noticeable (obviously the shacks have been removed). Since the segmentation algorithm generated in these areas image objects and structures which are close to those of existing informal settlements a misclassification was unavoidable. In a final step the several forms of landuse res. settlement areas have been grouped semantically into the abstract classes: *formal settlement, informal settlement* and *no settlement*. The final classification result (grouped and ungrouped) is shown in Fig. 10.

4 DISCUSSION AND OUTLOOK

Because of its high spatial resolution IKONOS data is well suited to detect informal settlement areas although single shacks are hardly detectable. To take advantage of its spectral properties applying image enhancement methods such as the principal components method can be useful especially for application tasks dealing with large scale objects. As the image enhancement leads to a more emphasised local contrast the analysis with eCognition benefits from it in two ways:

- The image segmentation leads to a better outlining of real-world objects (mostly houses and roads).
- The difference of typical textures of the different settlement forms is more denoted. This leads to an easier describable texture of the settlement forms.

Regarding the complex and different structures of (informal) settlements, eCognition's multi-resolution segmentation is well suited to generate image objects which are easy to describe. To classify them typical textural features of (informal) settlements can be taken into account, whereas form texture (form criteria of sub-objects) acts as a well distinguishing feature. In addition contextual information can be used to enhance a classification in terms of applying semantic knowledge to analyse an image. Although in the example present a single house detection could not be performed, extracting and classifying houses and shacks could improve the classification of (informal) settlements. Thereby more quantitatively features could be determined (e.g. number of shacks or density of shacks within a certain area). However, in the example present most of the informal settlement areas have been detected satisfying. Problems mainly occurred within settlement areas where even a visual inspection could not lead to satisfying results (such as within the bright settlement areas in the south east of the image). In such cases only ground-truthing or image data of higher resolution would give evidence about the real landuse res. settlement form. In cases image objects are obviously misclassified a final correction by hand could easily be performed with eCognition. Regarding the performed workflow during classification a stepwise refinement was achieved which is very close to applying an elimination key when interpreting an image analogous (see Lillesand, T. M. & Kiefer, R. W., 1994). This usually leads to a clearly structured and comprehensible classification result.

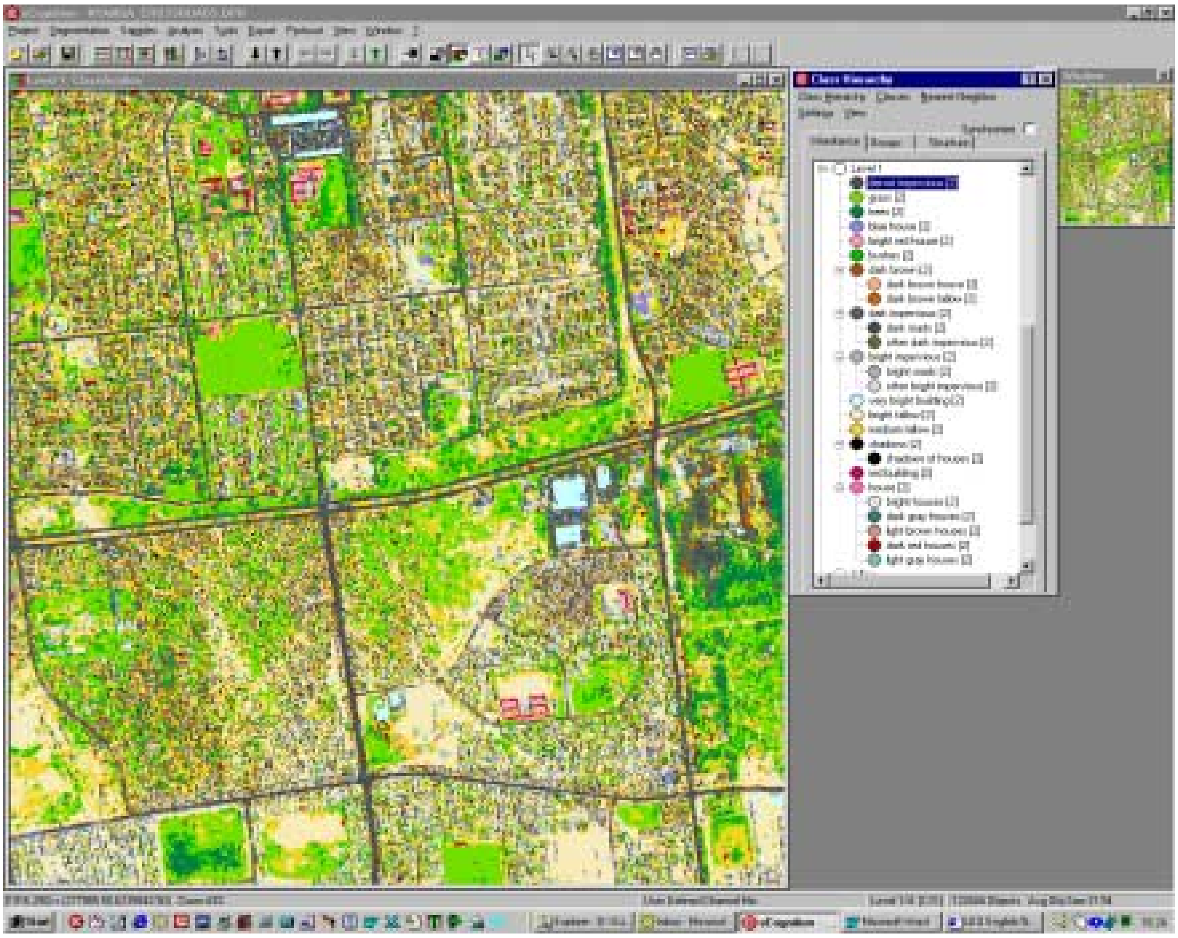

Fig. 9: Classification result of level 1. Note: The classes *grass, trees* and *bushes* have been grouped to *vegetation* for the analysis of the topmost level.

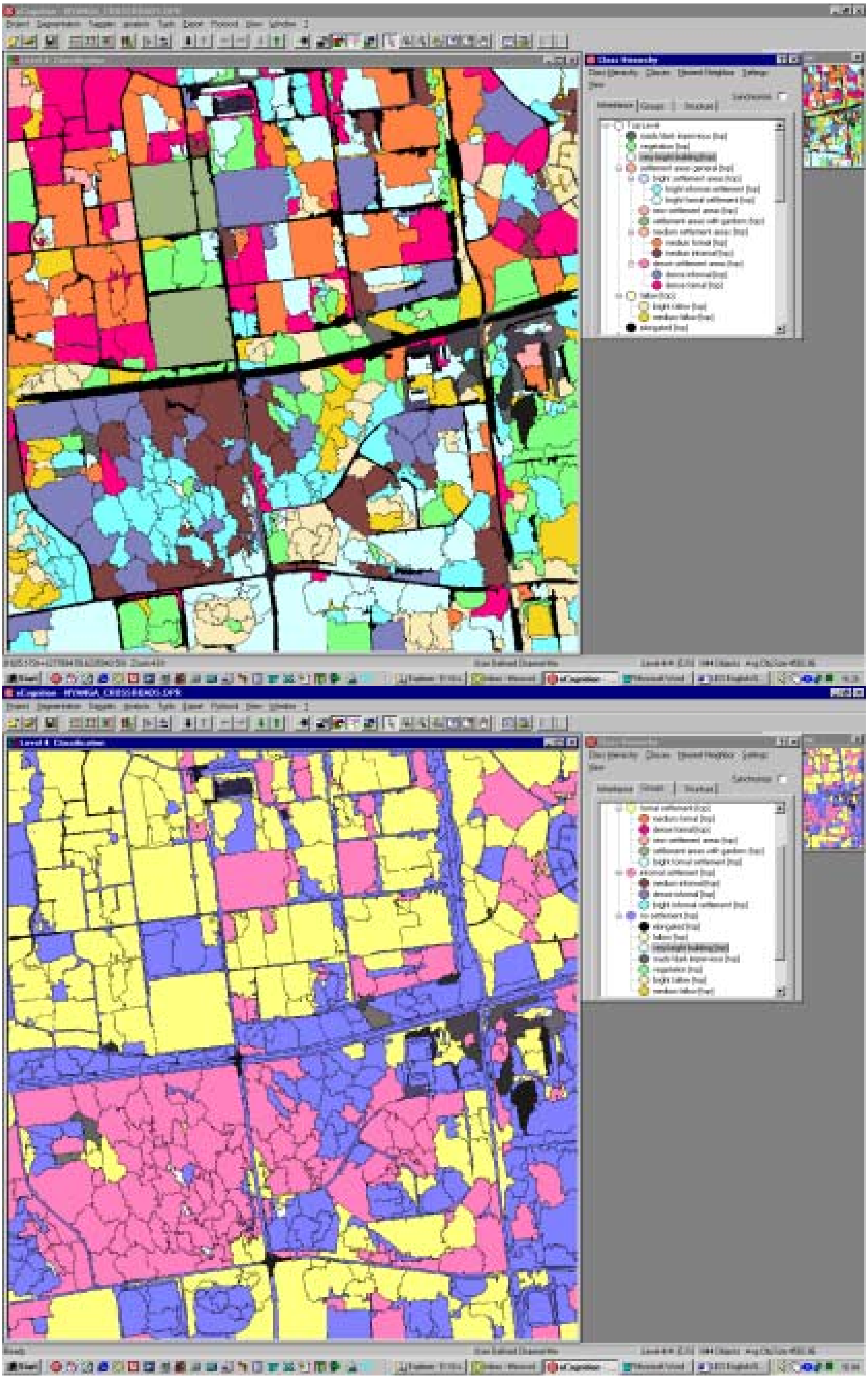

Fig. 10: Grouped (bottom) and ungrouped (top) classification result of level 4. Note: Some non-settlement segments remained unclassified.

ACKNOWLEDGEMENT
The author likes to thank Prof. Heinz Rüther (Dept. of Geomatics, Univ. of Cape Town) to provide the processed image data.

REFERENCES

BAATZ, M. & SCHÄPE, A., 1999. GIS-Aktualisierung und Image Fusion basierend auf einem neuen Verfahren zur objektorientierten und multiskalaren Bildanalyse. In: STROBL, J., BLASCHKE, T. (Hrsg): Angewandte Geographische Informationsverarbeitung XI. Beiträge zum AGIT-Symposium Salzburg 1999. Karlsruhe. Herbert Wichmann Verlag: pp. 12 - 24.

BAATZ, M. & SCHÄPE, A., 2000. Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: : STROBL, J., BLASCHKE, T. (Hrsg): Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000. Karlsruhe. Herbert Wichmann Verlag, pp. 12 - 23.

DEFiNiNENS AG, 2000. eCognition User Guide. http://www.definiens.com/register/index.html. (01. May 2001).

DEKOK, R. ET AL., 1999. Object based image analysis of high resolution data in the alpine forest area; Joint WSf ISPRS WG I/1, I/3 and IV/4: SENSORS AND MAPPINHG FROM SPACE 1999; Hannover, September 27-30.

HOFMANN, P. & REINHARDT, W., 2000. The extraction of GIS features from high resolution imagery using advanced methods based on additional contextual information – first experiences. In: IAPRS, Vol. XXXIII, Amsterdam, 2000. Pp. 51 – 58.

LILLESAND, T. M. & KIEFER, R. W., 1994. Remote Sensing and Image Interpretation. 3rd ed., New York, John Wiley, pp. 152 – 153.

MASON, O. S. & FRASER, C. S., 1998. Image Sources for Informal Settlement Management. In: Photogrammetric Record 16(92), pp. 313 – 330.