CHANGE DETECTION OF INFORMAL SETTLEMENTS USING MULTI-TEMPORAL AERIAL PHOTOGRAPHS – THE CASE OF VOI, SE-KENYA

Hurskainen, P. & P. Pellikka
Department of Geography, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland
e-mail: pekka.hurskainen@helsinki.fi

KEY WORDS: urban growth, informal settlements, building extraction, change detection, image segmentation, visual interpretation.

ABSTRACT:

Black and white and true-colour aerial photography from 2004, 1993 and 1985 are used for studies of growth and change of informal settlements in Voi, SE-Kenya. The digital data is orthorectified, corrected for brightness variations caused by light falloff and bi-directional effects and mosaicked using *EnsoMOSAIC* and *Erdas Imagine* software. The constructed mosaics are segmented with *eCognition* software and classified to built-up and non-built areas, in order to produce a mask of structures, which will be extracted into vector format. A post-classification comparison change detection in building-level will be done with *ArcView* software between all three masks. The work is still under way, but some preliminary results are given.

1. INTRODUCTION

1.1 Urban Growth in the Developing World

One of the biggest obstacles for development in the developing countries is rapid population growth. This, together with continuing poverty and a lack of basic needs for an acceptable life (e.g. food, clean water, shelter, basic health care, security of tenure) imposes a great challenge for sustainable development. Most of the population growth in the world during the next 15 years will be urban growth, and the vast majority of it will take place in developing countries. The population growth rate for years 2000-2020 is estimated at 1.3 percent in developing regions (2.4 percent in Sub-Saharan Africa) compared to only 0.1 percent in developed regions (HABITAT, 2003A: Table A.1). There remains a huge potential for urban growth for the countries in Africa, since in the year 2000 only 37.1 percent of their total populations were urban – this is estimated to grow up to 47.8 percent by the year 2020 (ibid.).

In Kenya, the level of urbanization in 2000 was 33.4 percent and it is estimated to go up to 50.8 percent by 2020, indicating an annual growth rate of 3.76 percent in the next 15 years (HABITAT, 2003A: Table B.2). Of all urban population in Kenya, 70.7 percent lived in informal settlements, a total of 7.6 million people (HABITAT, 2003B: Table 1), which is alarmingly high when compared to the average of developing regions, 43.0 percent (HABITAT, 2003A: Table 1.3). Thus, a major focus in sustainable development and issues related to it in Kenya should be directed towards the urban environment and problems caused by rapid urban growth.

Informal, or *spontaneous, settlements* are settlements whereby persons, or *squatters*, assert land rights to or occupy for exploitation land which is not registered in their names, or government land, or land legally owned by other individuals (Kibwana, 2000: 110). Squatters are people who occupy land or buildings without the explicit permission of the owner (HABITAT, 2003A: 82). In Kenya, there exists no official definition of slums or informal settlements, and these terms are used interchangeably (HABITAT, 2003A: 219). City authorities, however, usually view lack of basic services and infrastructure as characteristics of slums, an aspect that slum dwellers do not emphasize (ibid). For these reasons, the term '*informal settlement*' will be preferred to '*slum*' within this study.

1.2 Aim of the Study

Given the facts raised above in the urban areas of the developing world, there is an urgent need for fast decision-making and planning in order for the government officials and urban planners to maintain at least some control of the city growth. This is not possible without up-to-date information about various aspects in the urban areas, e.g. cadastral and socio-economic data.

Unfortunately for the developing countries, this kind of data is unreliable, obsolete or just simply non-existent (Baudot, 2001: 225-226; Balzerek, 2001). Usually the only ready-to-use data available for third world cities are limited to outdated topographic maps and National Census population data. Their accuracy and suitability vary greatly, however the key problem is how to cost-effectively update the datasets to meet the needs of the changing situation and a fast growing population, given the limited budgets, equipment and personnel in the local administration.

Detecting change in urban areas is not only of academic interest, because it serves as the major data source for strategic planning and analysis in urban areas (Donnay et al., 2001). In addition, to get information of the current situation, the need to move from 2.5D or 3D representations of urban areas into 4D – that is, including time – is clear (Bähr, 2001: 95). In order to understand the dynamics of an urban area, it is very important to take into consideration the historic aspects as well.

Finding patterns in the formation of informal settlements might give valuable clues for the planners to learn from the mistakes made in the past. When no other data is available for this kind of analysis, one solution is to use remote sensing imagery as the primary data source, and GIS (geographic information systems) plus secondary data sources in the key role of providing a framework for spatial analysis of remote sensing data (Bähr, 2001: 95; Donnay, et al. 2001: 10-11). However, due to the microstructure and instability of shape of the informal settlements, the detection is substantially more difficult than in formal settlements (Hofmann, 2001). According to Mason & Baltsavias (1997), possible applications for spatial models of informal settlements include monitoring informal settlement growth from the regional level to the management of individual settlements, shack counting for electoral boundary determination, generation of GIS/CAD databases of the infrastructure for service upgrading, soil and ground water quality evaluation for environmental quality assessment and settlement upgrade scenario modelling. Requirements for these models emphasize less the need for positional and object modelling accuracy and more completeness of records in the face of highly dynamic environments (Mason & Baltsavias, 1997).

The principal aim of this study is to detect changes that have taken place in the built-up environment of the informal settlements in Voi during the timeframe of investigation; 1985-2004. From each temporal aerial photograph mosaic, a structure mask is extracted using two methods: (1) object-oriented image segmentation and (2) visual interpretation. The classification itself is very simple, consisting of built-up and non-built-up classes only. These masks are then compared with post-classification comparison change detection, in order to highlight the changes in built-up environment (either loss or gain of structures) in time. This method was chosen, because it is not dependent on pixel spectral values, which is essential when the data consists of both black-and-white and true-colour images.

The scale of the study must be understood – the study is made at a regional level to reveal general patterns of growth in informal settlements. The structure masks are not accurate representations of the boundaries of individual buildings, due to the limitations of data quality, accuracy and suitability for segmentation, so the presented method is not meant to be a tool for extracting buildings e.g. to create or update cadastral GIS datasets. Although the data is gathered at a structure (micro) level, the accuracy is more suitable for larger scale studies, such as the growth of settlements at a regional level.

It is hoped that this study has some significance, at least in creating awareness of the history and current situation of informal settlements in Voi, or it could maybe even be used locally for detecting the growth of these settlements. In the Taita Taveta District Development Plan 2002-2008, it is stated that “lack of updated land information” and “increasing number of squatters” are the major constraints of land administration in the district, and the strategies to overcome these would be “computerization of land information, adequate staff and alienation of public land for settlement schemes” (Republic of Kenya 2002: 32).

1.3 Theoretical Framework

In 1989, the Department of Geography of the University of Helsinki carried out a minor field study in Taita Hills. Dr. Sakari Tuhkanen led more than 20 students from the University of Helsinki and the University of Nairobi to Taita Hills in order to study land use. The study resulted in several M.Sc. theses and a few research reports (Pellikka et al., 2004). In 2004, a second field expedition was organised to Taita Hills led by professor Petri Pellikka. The expedition was linked to a TAITA project, the objective being to develop a cost-effective and practical land use change detection methodology and to create a GIS database of land use in the area. This paper presents M.Sc. student Pekka Hurskainen’s case study within the project and is directly connected to TAITA project and its aims and objectives. The project is presented in more detail in Pellikka et al. (2004).

As mentioned above, the methodology used in this study is a mixture of remote sensing (RS) and GIS methods. The initial selection of the methods was guided by the following principles:

- The methods have to apply in heterogenic urban informal settlement context, taking into consideration their special needs.
- Methods suitable for high-resolution aerial image data, such as segmentation, have to be considered (Longley et al., 2001, Mason & Baltsavias, 1997). Thus the usage of SPOT XS satellite images was rejected, although they would have been available in TAITA project.
- The focus will be in mapping the changes of informal settlements on a regional level, not to produce accurate data of individual buildings.
- Integration in a desktop GIS environment to enable further processing and analysis (Mason & Baltsavias, 1997), namely ESRI ArcGIS environment.
- The methods have to be cost-effective and simple, so they could be replicable locally in the target country.

The simplest method to extract meaningful information from RS data is visual interpretation. Visual interpretation can be defined as “the science and art of observing images with the objective of identifying different objects and judging their significance” (Philipson, 1977, cit. Jensen, 2000: 119). Although as old a method as remote sensing itself, visual interpretation has still kept its significance for large-scale urban RS studies, at least as a method available when all other methods fail or are not available for some reason. It is also used for doing visual checking and evaluation. While being a relatively easy and accurate method for an experienced interpreter, the downsides include slowness, low cost/efficiency ratio and lack of objectiveness.

Pixel-based classification algorithms, like nearest neighbour (NN), are widely used in land use studies, but the limitations are clear and widely accepted. The idea of a pixel belonging to a certain class is that is has to be close to the spectral feature space of that class (Burnett & Blaschke, 2003: 239). In the case of classifying complex environments, such as urban areas, with very high-resolution RS data, the pixel-based approach is not justifiable any more. Firstly, pixels do not sample the urban environment at the spatial scale of the features to be mapped, and buildings are represented by groups of pixels which should be treated as individual objects instead (Smith & Hoffmann, 2001: 100). Secondly, a building produces a wide range of spectral signatures as the pixels will represent different facets of the roof (ibid.). This is especially true in the context of informal settlements, where roofs are constructed from diverse materials with variable texture and colour (Mason & Baltsavias, 1997). Thirdly, many features in the urban environment appear similar spectrally (e.g. concrete roofs and tarmac roads) and can be discriminated only by some external context information (Smith & Hoffmann, 2001: 100).

Pixel-based approaches take only one scale, one pixel at a time, into investigation ignoring the concepts of hierarchy, neighbourhood and scale (Burnett & Blaschke, 2003: 239). Thus, the basic principle is to move from dependency on individual pixel DN values into a way of incorporating shape, texture and contextual information for image classification, which is only possible by creating meaningful objects and their mutual relationships (Darwish et al., 2003). In this study, the objective is to handle structures as single objects.

One possible object-oriented approach is multi-scale segmentation/object relationship modelling technology (Burnett & Blaschke, 2003: 240). The concept of segmentation was initially presented by Haralick et al. (1973), where textural features based on grey-tone spatial dependencies where used to classify images "on a block of contiguous resolution cells". The key problem in segmentation is how to define a set of meaningful features or objects from the image scene (Haralick et al., 1973: 610, Burnett & Blaschke, 2003: 240). The solution is to look for changes in image object heterogeneity and homogeneity (Burnett & Blaschke, 2003: 240). Common segmentation methods can be grouped into three broad categories, namely point-based (e.g. grey-level thresholding), edge-based (e.g. edge detection techniques) and region-based (e.g. split and merge) (Jähne, 1993, cit. Darwish et al., 2003).

The problem of extracting individual buildings from RS data has got some recent attention, and various approaches for different applications have been made. Mason & Baltsavias (1997) presented a method for automatic shack detection by fusing shadow data with the 2.5D blobs derived from segmentation of a normalized DSM, i.e. Digital Surface Model (DSM) minus Digital Terrain Model (DTM). The DSM generation requires, however, nadir stereo imagery with strong shadows and high overlap (Mason & Baltsavias, 1997).

Jung (2004) used a two-step algorithm to detect changes in buildings for updating a GIS database. Firstly a large part of the scene was eliminated without losing any changes by comparing a DEM for the two dates, and secondly the resulting "regions of interest" (ROI) were classified as "building" and "non-building" based on a combination of several decision trees induced from the training data (Jung, 2004). Knudsen (2003a; 2003b) had the same objective of updating an existing GIS database, but he used an algorithm that is based purely on spectral recognition of roofs; after the input image was segmented and the existing database projected onto it, using these clusters as a training set for the classification (Knudsen, 2003a). An edge detection and image segmentation program (EDISON) was used, in addition to MATLAB (Knudsen, 2003a; 2003b). Knudsen et al. (2002) compared the neuro-fuzzy network clustering method called Weighted Incremental Neural Networks (WINN) with a method based on ISODATA unsupervised classification algorithm to detect buildings from colour-infrared aerial photos. The results showed that the traditional ISODATA method has an advantage of splitting background from foreground and the neural network method results in a much more uniform segmentation of the input images (Knudsen et al., 2002).

The work by Hofmann (2001) was quite influential for this study. Hofmann used the *eCognition* software to detect and discriminate informal settlements, using an IKONOS scene, from other land-use types by describing typical characteristics of colour, texture, shape and context. The method used was object-oriented multi-scale image segmentation approach, which creates segments at various scales. These segments act as image objects whose physical and contextual characteristics can be described and classified by means of fuzzy logic and the traditional nearest neighbour algorithm (Hofmann, 2001).

In conclusion, none of these methods was directly applicable to the study problem. The idea of using elevation data to extract buildings was not possible, because the necessary data for DSM generation was not available for the study. Methods generated mostly to update cadastral GIS databases were not designed for detecting buildings in informal settlements with their respective special requirements. However, the author had the possibility to evaluate and use the *eCognition* software for a few weeks, so the image segmentation, classification and post-classification comparison change detection methods became justifiable for the purposes of this study, with a few modifications.

2. STUDY AREA AND DATA

The study area is the township of Voi, located in South-Eastern Kenya, Coast Province, Taita Taveta District, 327 km South-East of Nairobi and 159 km North-West of Mombasa (lat. 3°25´, long. 38°20´). It is situated at an altitude of approximately 580 m above sea level. It borders in the north and east the Tsavo East National Park, in the south Sagala Hills, and in the west Voi Sisal Estates.

The newest population census from 1999 puts the population of Voi Municipality at 33,077 and that of the Voi Township at 24,404 residents (Republic of Kenya, 2001). Voi Municipality (the same as Voi location) is much more bigger in area, but also contains rural areas. The borders of Voi Township, however, were drawn back in 1932, the centre point being the Voi railway station in Voi, from where a full circle with 1-mile radius was drawn (see Hurskainen, 2004: figure 1). It is clear that these administrative boundaries no longer serve their original purpose, because Voi has grown well outside the circle. Statistics based on administrative boundaries should thus be used with caution.

The growth of Voi Township has been fast and steady since independence (Table 1), although the growth rate is now slightly smaller than 10 years ago. The study area is well suited for a change detection study, because it is of intermediate size, it has steady population growth, it lies on relatively flat terrain and it has little vegetation cover. More details about the development and structure of Voi are given in Hurskainen (2004).

Table 1. Census data for Voi Township, 1969-2003 (Republic of Kenya, 1981; 1994; 2001, VMC 1995).

Year	Population	Households
1969	5,313	-
1979	7,397	2,182
1989	16,273	4,507
1999	24,404	6,818
2003 (projected)	26,161	-

The RS data used in this study consists of three aerial photograph datasets (Pellikka et al., 2004). The first dataset was acquired in March 2003 using the University of Helsinki's custom Nikon D1X true colour digital camera system with navigation unit and software, mounted on a single-engine aeroplane. Unfortunately, these images had serious problems with shadows cast by clouds, and the images didn't have enough side-lap for proper mosaic production. The 2003 mosaic was, however, used in the fieldwork as a reference.

Another data acquisition was successfully made in January 2004 with the same equipment. 110 images were selected to produce the mosaic, with one-meter spatial resolution (pixel size). The second and third datasets consist of black-and-white aerial photographs from years 1985 and 1993 acquired by PhotoMap Nairobi. The 1993 images are already pre-processed, ortho-rectified and mosaicked with a pixel size of 0.5 m. The 1985 photographs over Voi are not processed yet.

Other data sources include the Survey of Kenya 1:50 000 scale topographic map of Voi (1991), vector features digitised from the map, and a DEM created from the digitised contour lines (Broberg & Keskinen, 2004; Pellikka et al., 2004). The census population data also provides some background to the study.

The fieldwork session in Voi was conducted during January-February 2004. The mosaic prints from the year 2003 were used as "base maps" for field notes. In these maps, different settlement areas, land-use types, plot boundaries, place names, important buildings and other points of interest were marked. Special attention was given to informal settlements. In addition, some 20 interviews were made. For more information on the fieldwork, see Hurskainen (2004).

3. METHODOLOGY

3.1 Data Pre-Processing

Before mosaicking, each image was corrected for light falloff effect in *Erdas Imagine* with a modified correction method by Pellikka (1998). After this, processing of the 2004 images continued in *EnsoMOSAIC* software, developed by Stora Enso Forest Consulting Ltd. and the Technical Research Centre of Finland (VTT) (Holm et al., 1999; StoraEnso, 2002). GPS data was collected for each individual image during the flight, providing three-dimensional coordinates for the centre of each image. Neighbouring image pairs were manually linked together with three common points between each pair. After this, an automatic tie point search was run. More tie points were added manually to some areas where the automatic search had failed to find points. In addition, tie points with large residuals had to be removed manually.

After the number of tie points was adequate between all image pairs, the bundle block adjustment (BBA) was run. BBA essentially is an "iterative mathematical process to solve the orientation of the images and the location of the perspective centres simultaneously for a large image block" (StoraEnso, 2002: 2). After each iteration, an estimation of the global accuracy of the image rectification, called adjustment error, is given (StoraEnso 2002: 35). Adjustment error is the mean error of unit weight, a function of all the residuals and all the weights (StoraEnso 2002: 33). After each round of the BBA, erroneous tie points with large residuals (over 5.0) had to be removed, and the BBA was run again. This continued until the largest residuals were removed and the accuracy of the mosaic was considered adequate, that is when the total adjustment error was reduced to 1.92.

After the BBA, a Digital Terrain Model (DTM) with 5-meter ground resolution was made in *EnsoMOSAIC* for ortho-correction of the mosaic. The DTM is created based on the elevation values generated for each tie point during the BBA (StoraEnso 2002: 38).

A correction for bi-directional effects (BRDF - bi-directional reflectance distribution function) introducing brightness variations (Pellikka, 1998) is implemented within *EnsoMosaic*, and it is calculated on the fly when the mosaic is created. The user needs to input a correction factor f to take into account the land surface reflectance properties and atmospheric conditions in each of the three channels (Pellikka, 1998: 46-49). The factors should be chosen to decrease the correlation between scattering angle and the pixel DN value (Mikkola & Pellikka, 2002). In this case, factors were determined empirically. Some brightness variations can still be seen between flight lines, but the result is the best that can be achieved with the current correction algorithm implemented in the software, which doesn't allow the user to input different correction values for different land use types. The study area has rather heterogeneous soil, vegetation, and land cover types which results in unequal anisotropic scattering to the recording instrument (Pellikka, 1998: 13-16). An optimum approach would have been, of course, to correct the BRDF for each image individually. This would, however, have required extensive work, taking into consideration the large number of individual images and the fact that the correction parameters would have had to be determined for each image individually.

Finally, the 2004 mosaic was created in *EnsoMOSAIC* using histogram matching between three images and bilinear interpolation methods. The final 1-meter ground resolution mosaic gave satisfactory results. The mosaic was then imported into *Erdas Imagine*, where it was re-sampled to the same co-ordinate system (TM/Clarke 1880) as the topographic maps and the vector data.

The pre-processing for black-and-white aerial photographs from 1993 was not possible in *EnsoMOSAIC*, because there was no GPS data available. Therefore, the whole processing chain had to be done in *Erdas Imagine*. Only the light falloff correction was made, using the same method described above. Visual inspection of the images revealed that BRDF correction was not necessary.

The ortho-rectification was done using the camera parameters provided by *Photomap* and the DEM created from the digitised contour lines (Pellikka et al. 2004). The images were ortho-rectified using the 2004 mosaic as a reference, in order to maximize the positional accuracy between the two mosaics. A total of seven images were ortho-rectified and then put into a mosaic. The final ground resolution for the 1993 mosaic is 0.5 meters.

3.2 Image Segmentation and Building Extraction

There are various methods created for image segmentation (see, for example, Burnett & Blaschke, 2003: 240), but the one chosen for this study is called multi-resolution segmentation, implemented in the *eCognition* software. It is a bottom-up region-merging technique, whereby each pixel is considered initially as a separate object, and subsequently pairs of image objects are merged to form bigger segments (Darwish et al., 2003). The principle idea is that meaningful information is not represented in individual pixels, but in meaningful image objects and their mutual relationships (Defininens, 2004). Image objects are further classified using fuzzy logic allowing the integration of a broad spectrum of different object features, which can be grouped into three categories: intrinsic features such as colour, texture and form of the objects; topological features describing the geographical relationships between the objects or the whole scene; and context features, which describe the objects semantic relationships (ibid.). A hierarchical object-oriented approach means basically that each object has a sub- or super-class, which enables inheritance of class features, making the classification procedure much more powerful and versatile compared to traditional pixel-based methods.

Neubert and Meinel (2003) undertook an extensive evaluation of six segmentation programs using an IKONOS scene. Results were evaluated visually using reference areas, and *eCognition* software ranked first in overall quality. Neubert and Meinel (2003) argued that "the segments are sometimes irregular and in areas of low contrast the occurrence of faulty segmentations is possible". The multi-scale segmentation and fuzzy logic features attributed to the high potential of the software.

One of the strengths that *eCognition* has is the fact that in can be used either as a "traditional" segmentation tool, whereby the segments can be exported to another software for further processing, or it can be utilized further, conducting detailed multi-scale classifications or building complex classification protocols for fully automated image classifications. In this study, however, *eCognition* was used only for the segmentation of the 1993 and 2004 mosaics, and for classification of the 2004 mosaic. The sole reason for not utilizing the advanced functions of the software was that the author had only a test licence for a few weeks, which was not enough to experiment fully with the fuzzy classifier and multi-scale segmentation possibilities.

Multiresolution segmentation is a "heuristic optimisation procedure, which locally minimizes the average heterogeneity of image objects for a given resolution over the whole scene" (Defininens, 2004). In *eCognition*, the user can control the segmentation process with various parameters, but the most important ones are the scale parameter and composition of homogeneity criterion. The scale parameter is "an abstract term which describes the maximum allowed heterogeneity for the resulting image objects" i.e. the bigger the scale value, the bigger the segments (ibid.). Heterogeneity criterion consists of two parameter pairs: colour/shape and smoothness/compactness, which are given weights from 0.1 to 0.9. The former controls the influence of colour (spectral values) vs. shape homogeneity on the object generation, and the latter controls the smoothness vs. compactness of the objects. There are no clear suggestions for the usage of these parameters, as they should be determined on an *ad hoc* basis.

After experimentation, the best results for this study were obtained when parameters of scale, shape and smoothness were given values of 5, 0.9 and 0.8, respectively (Figure 2). In addition, thematic layers were incorporated to improve the classification accuracy. Three vector datasets, digitised from topographic 1:50 000 scale maps and updated from the 2004 mosaic, were already available from the TAITA project (Broberg & Keskinen 2004). The vectors represented the main tarmac roads, the railway line and major dirt roads within Voi town. Since this information was already known, it could be excluded from the classification, because roads and roofs have nearly similar spectral values, causing misclassifications. Multi-scale segmentation was not used, however, because single-scale segmentation was enough for the purpose of extracting roofs or parts of roofs from the background.

The roof materials in the study area consist of roughly 2/3 of iron sheets and 1/3 of other materials, such as tiles, concrete, asbestos, tin, grass and *makuti* (Republic of Kenya, 2001: 43). Iron sheet roofs were segmented and classified with 95 percent accuracy, whereas other roof materials had considerably lower accuracies (see Figures 1 and 2).

The difference between *land use* and *land cover* must be stressed. Land use is understood as a function between humans and nature, and it can be economical, social or political (Baudot, 2001: 226). Land cover, however, is the physical material that covers the earth surface (ibid.). In segmentation and classification, a problem arises when two land use types (e.g. built-up and bare soil) consist of the same land cover type (e.g. red latosoil). Figure 1 illustrates a typical example from Voi. In these cases, contrast between the object and the background is low, making the segmentation and classification task very hard.

Sometimes informal settlements are so dense that delineating individual buildings becomes hard, but for the purposes of this study it does not matter so much, since we are mainly interested in general patterns. Yet another limitation is vegetation: when there is a big tree near a building, it can cover partially or completely the building, leading to misclassifications. Usually visual interpretation is the only way to deal with such cases.

Figure 1. Land use vs. land cover problem (left), variety of roofs (right).

3.3 Class Hierarchy and Image Classification

When the segmentation is done and the results are satisfactory when inspected visually, the next task is to build the class hierarchy. In *eCognition*, class hierarchy can be understood as a rule base wherein the user determines physical and semantic properties typical for the objects of a certain class (Hofmann, 2001: 112). Two classifiers are implemented in the software: the first is a traditional nearest neighbour (NN) classifier, where the user interactively collects samples from the image to train the classifier, the second is based on fuzzy membership functions, which describe the intervals of feature characteristics wherein the objects do belong to a certain class or not by a certain degree (ibid.). In this study, however, the NN classifier was mostly used; the fuzzy functions were only used to classify roads based on the thematic layers. Lack of auxiliary data, such as elevation, and the fact that the informal settlements have very few common characteristics due to their irregularity and spontaneous nature, made the task of creating fuzzy membership functions very tedious. For most cases spectral values was enough to distinguish buildings from the background with satisfactory classification accuracy, but for the problematic areas visual interpretation was used instead.

The class hierarchy is very simple (Figure 2), consisting of four super-classes: no data, open, roofs and trees, with sub-classes for open and roofs. The sub-classes inherit any class properties that their super-classes have. The "open" super-class has sub-classes for dirt roads, tarmac roads, railroads and "other open" which basically is anything else than roads. The roofs super-class has sub-classes for roofs that have different spectral values (e.g. red, brown and blue roofs).

Sample data was collected throughout the image subset, trying to get as representative samples as possible for each class. Instead of using pixels, the samples were basically segments with their respective spectral values. After some samples were collected, the nearest neighbour classifier was run for the first time. The results were then evaluated and more samples were collected, where needed, and the classifier was run again. This continued until the classification accuracy was satisfactory.

Because the classification was done solely with spectral values, the black-and-white image mosaic of 1993 could not be classified within *eCognition*. Grey tones for roofs were not spectrally distinct enough to separate them from the background. Using texture as a fuzzy membership function was tried as well, with disappointing results. Although the segmentation itself was accurate enough to separate buildings from the background due to enough contrast between the two, it was decided that the classification should be done in another way. Image objects were exported as polygons into a shapefile, which was then opened in *ArcView* software. Classification of the 1993 image objects into built-up and background was thus done visually using the mosaic and the exported image objects. Classified objects from the true-colour mosaic of 2004 were exported into shapefile as well. Visual inspection of the classification was done in *ArcView* and corrections made, where needed.

3.4 Post-Classification Comparison Change Detection

After the visual evaluation of the object shapes, polygons classified as "built-up" were exported into a new shapefile. Because one building could consist of several objects, or polygons, the shapefile was dissolved using the geoprocessing tools of *ArcView*. The result was one big multi-part polygon, which was subsequently exploded into individual polygons using an *ArcView* extension called EditTools (http://www.ian-ko.com). Now each polygon represented one building, and the built-up mask was ready to be used in modified post-classification comparison change detection. For an overview of the method, see, for example, Jensen (1996: 269-270).

The procedure used for change detection is again very simple. The two building masks, one for 1993 and one for 2004, were both opened in *ArcView*. In order to distinguish new buildings from the 2004 mask, a polygon overlay function was used. Every polygon from the 2004 mask, that intersected a polygon from the 1993 mask, was selected using the "select by theme" function. These selected polygons now represented the buildings that were already present in 1993. To get the new buildings, we simply reverse the selection and export it into a new shapefile. Similarly, to get buildings that have been demolished between 1993 and 2004, every polygon form the 1993 mask that intersects the 2004 mask was selected, the selection reversed and then exported into a new shapefile.

Some of the changes are not representative of the real situation, though. For example, there are cases of a building that existed in 1993 that was later demolished, and a new building was erected on the very same plot. Some buildings have been extended after erection, maybe with a few rooms more. The post-classification comparison change detection method cannot detect these kind of changes, which are anyway insignificant for the purpose of this study.

4. RESULTS AND DISCUSSION

The work is still under way, but some preliminary results can be given. Figure 3 shows the change detection results of a test site in a very typical informal settlement in Voi, called Mwakingali. As can be seen from the 1993 image, the area used to be quite sparsely populated with 392 structures, dominated by a patchwork of homesteads and agricultural fields (*shambas*). By 2004, the area has changed completely, 1914 new buildings have been erected haphazardly and spontaneously, *shambas* have disappeared completely to make more space for new settlers. A total of 111 structures have been demolished.

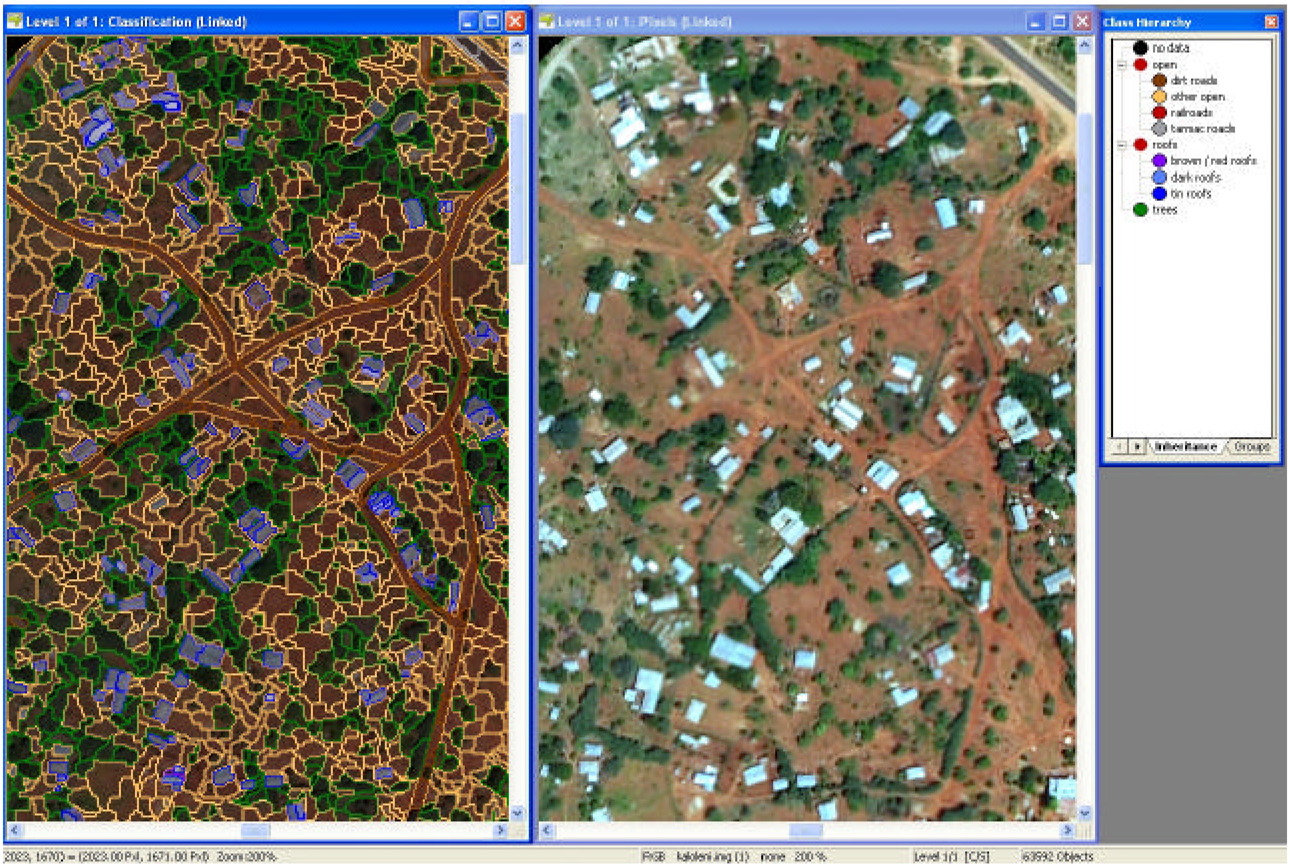

Figure 2. Segmentation and classification of Kaloleni settlement (left), original image (middle) and class hierarchy (right).

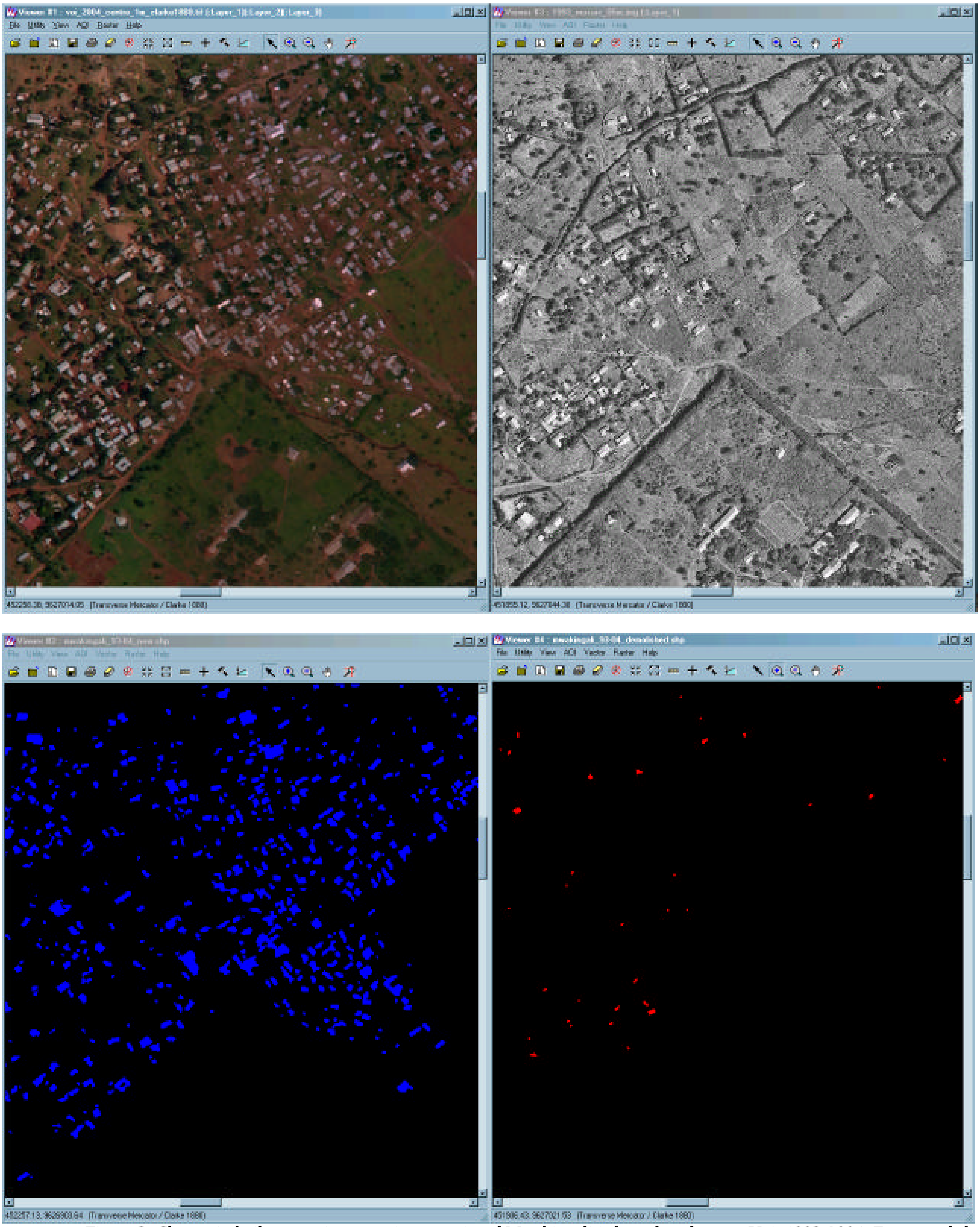

Figure 3. Change in built-up environment in a test site of Mwakingali informal settlement, Voi, 1993-2004. From top-left: 2004 mosaic, 1993 mosaic, new buildings since 1993 (blue), demolished buildings since 1993 (red).

The classification accuracy (producer's accuracy) for the 2004 segments was 95.7 percent (4.3 omission error) for built-up and 98.7 percent (1.3 omission error) for non-built-up segments, when comparing the *eCognition* classification against the visually checked classification. However, the sample site consisted almost completely of iron sheet roofs, so the accuracy of other areas with other roof materials will be somewhat smaller.

The method presented above for change detection of informal settlements has its strengths in simplicity, straightforwardness, cost-effectiveness and relative accuracy.

Drawbacks include limitations caused by the source data (e.g. difficulties in classifying automatically the black-and-white images), dependency on many software and lack of automation. However, when considering the conditions in developing countries, where labour is cheap and manual work is not considered as a thing to be avoided, labour-intensive methods can be favoured to expensive high-end solutions.

Future work includes the preprocessing, mosaicking, and building extraction of the 1985 aerial photographs.

Change from 1985 to 2004 will give a time-span of almost 20 years, revealing even more patterns of how the informal settlements have grown and changed. According to the persons interviewed in Voi, most of the growth took place in the 1980s and 1990s - this hypothesis will be tested with the data. Importance must also be given to understand *why*, *where* and *when* these changes have taken place. Secondary data sources, like interviews and literature, are taken into consideration when trying to answer to those questions.

ACKNOWLEDGMENTS

The authors would like to thank, first of all, all the friendly and co-operative people met during the fieldwork in Voi, namely field assistant, Donald Maganga, as well as the whole staff of *Taita Taveta Wildlife Forum*, headed by Donald Mombo, who all made this whole study possible by giving their valuable time, efforts and advice. Dr. Charles Burnett, University of Victoria, Canada, is acknowledged for the valuable help on *eCognition* and segmentation concepts. The whole staff of *Novosat* Ltd. is appreciated for the practical knowledge, experience and training on GIS and remote sensing given for Pekka Hurskainen. The Department of Geography of the University of Helsinki provided GIS laboratory and the *GeoInformatics Research Group*, GIRG, provided personal expertise. Barnaby Clark is appreciated for checking the language and comments on the paper. This study was funded by the Academy of Finland. The research permission from the Republic of Kenya is greatly appreciated. *Asanteni sana*!

REFERENCES

Balzerek, H., 2001. Applicability of IKONOS-Satellite Scenes Monitoring, Classification and Evaluation of Urbanisation Processes in Africa. Case Study of Gombe / Nigeria. In: Jürgens, C. (ed.), *Remote Sensing of Urban Areas*, Regensburger Geographische Schriften, Heft 35, Regensburg, pp.25-35.

Baudot, Y., 2001. Geographical Analysis of the Population of Fast-Growing Cities in the Third World. In: Donnay, J-P., M. J. Barnsley & P. A. Longley (eds.), *Remote Sensing and Urban Analysis*, Taylor & Francis, London & New York., pp. 225-241.

Broberg, A. & A. Keskinen, 2004. Geodatabase over Taita Hills, Kenya In: Pellikka, P., J. Ylhäisi & B. Clark (eds.): Taita Hills and Kenya, 2004. *Expedition Reports of Department of Geography, University of Helsinki 40*, Department of Geography, University of Helsinki. In print.

Burnett, C. & T. Blaschke, 2003. A Multi-Scale Segmentation / Object Relationship Modelling Methodology for Landscape Analysis. *Ecological Modelling* 168, pp. 233-249.

Bähr, H-P., 2001. Image Segmentation for Change Detection in Urban Environments. In: Donnay, J-P., M. J. Barnsley & P. A. Longley (eds.), *Remote Sensing and Urban Analysis*, Taylor & Francis, London & New York., pp. 95-113.

Darwish, A., K. Leukert & W. Reinhardt, 2003. Image Segmentation for the Purpose of Object-Based Classification. http://www.defiens-imaging.com/documents/publications/FRO7_1420.pdf (accessed 1 May 2004)

Donnay, J-P., M.J. Barnsley & P.A. Longley, 2001. Remote Sensing and Urban Analysis. In: Donnay, J-P., M. J. Barnsley & P. A. Longley (eds.), *Remote Sensing and Urban Analysis*, Taylor & Francis, London & New York., pp. 3-18.

Defininens AG, 2004. eCognition User Guide v4.0.

HABITAT, 2003A. *The Challenge of Slums. Global Report on Human Settlements 2003*. United Nations Human Settlements Programme, Earthscan Publications Ltd, London and Sterling.

HABITAT, 2003B. *Slums of the World: The Face of Urban Poverty in the New Millenium*? Working Paper. UN-HABITAT, Nairobi.

Haralick, R.M., K. Shanmugam & I. Dinstein, 1973. Textural Features for Image Classification. *IEE Transactions on Systems, Man, and Cybernetics* 3(6), pp. 610-621.

Hofmann, P., 2001. Detecting Informal Settlements from IKONOS Image Data Using Methods of Object Oriented Image Analysis – An Example from Cape Town, South Africa. In: Jürgens, C. (ed.), *Remote Sensing of Urban Areas*, Regensburger Geographische Schriften, Heft 35, Regensburg, pp. 107-118.

Holm, M., A. Lohi, M. Rantasuo, S. Väätäinen, T. Höyhtyä, J, Puumalainen, J. Sarkeala & F. Sedano, 1999. Creation of large mosaics of airborne digital camera imagery. Proc. of 4th Airborne Remote Sensing Conf. and Exhibition, 21-24 June 1999, Ottawa, Canada.

Hurskainen, P., 2004. The informal settlements of Voi. In: Pellikka, P., J. Ylhäisi & B. Clark (eds.): Taita Hills and Kenya, 2004. *Expedition Reports of Department of Geography, University of Helsinki 40*, Department of Geography, University of Helsinki. In print.

Jensen, J.R., 1996. *Introductory Digital Image Processing. A Remote Sensing Perspective*. Prentice Hall, New Jersey.

Jensen, J.R., 2000. *Remote Sensing of the Environment. An Earth Resource Perspective*. Prentice Hall, New Jersey.

Jung, F., 2004. Detecting Building Changes from Multi-temporal Aerial Stereopairs. *ISPRS Journal of Photogrammetry & Remote Sensing* 58, pp. 187-201.

Kibwana, K. 2000. Land Tenure, Spontaneous Settlement and Environmental Management in Kenya. In: Wanjala, S. C. (ed.), *Essays on Land Law. The Reform Debate in Kenya*, Faculty of Law, University of Nairobi, pp. 105-136.

Knudsen, T., 2003a. Detection of Buildings in Aerial Photos. http://research.kms.dk/~thk/pubs/nmc-2003.pdf (accessed 19 March 2004)

Knudsen, T., 2003b. Simplified Detection of Buildings in Aerial Photos Using Synergistic Low Level Feature Extraction. http://research.kms.dk/~thk/pubs/dsagm-2003.pdf (accessed 19 March 2004)

Knudsen, T., H.H. Muhammed, & B.P. Olsen, 2002. A comparison of Neuro-Fuzzy and Traditional Image Segmentation Methods for Automated Detection of Buildings in Aerial Photos.

http://research.kms.dk/~bpo/graz2002/graz2002segmentpaper.pdf (accessed 19 March 2004)

Longley, P.A., M.J. Barnsley & J-P. Donnay, 2001. Remote Sensing and Urban Analysis: A Research Agenda. In: Donnay, J-P., M. J. Barnsley & P. A. Longley (eds.), *Remote Sensing and Urban Analysis*, Taylor & Francis, London & New York., pp. 245-258.

Mason, S. & E. Baltsavias, 1997. Image-Based Reconstruction of Informal Settlements. http://e-collection.ethbib.ethz.ch/ecol-pool/inkonf/inkonf_80.pdf (accessed 30 March 2004)

Mikkola, J. & P. Pellikka, 2002. Normalization of bi-directional effects in aerial photographs to improve land cover classification accuracy of boreal and sub-arctic vegetation for pollen-landscape calibration. *International Journal of Remote Sensing* 23, pp. 4719-4742.

Neubert, M. & G. Meinel, 2003. Evaluation of Segmentation Programs for High Resolution Remote Sensing Applications. http://www.ipi.uni-hannover.de/html/publikationen/2003/workshop/neubert.pdf (accessed 8 April 2004)

Pellikka, P., 1998. Development of Correction Chain for Multispectral Airborne Video Camera Data for Natural Resource Assessment. *Fennia* 176:1, pp. 1-110.

Pellikka, P., B. Clark, P. Hurskainen, A. Keskinen, M. Lanne, K. Masalin, P. Nyman-Ghezelbash & T. Sirviö, 2004. Land use change monitoring applying geographic information systems in the Taita Hills. *Proceedings of the 5th AARSE conference* (African Association of Remote Sensing of the Environment), 18-21 October, 2004, Nairobi, Kenya. In print.

Republic of Kenya, 1981. *Kenya Population Census 1979 Vol. 1.* Central Bureau of Statistics & Ministry of Economic Planning and Development.

Republic of Kenya, 1994. *Kenya Population Census 1989 Vol. 1.* Central Bureau of Statistics, Office of the Vice President & Ministry of Planning and National Development 1994.

Republic of Kenya, 2001. *1999 Population and Housing Census Vol. 1 & II.* Central Bureau of Statistics & Ministry of Finance and Planning 2001.

Republic of Kenya, 2002. *Taita Taveta District Development Plan 2002-2008. Vol. 22: Effective management for sustainable economic growth and poverty reduction.* Ministry of Finance and Planning.

Smith, G. M. & A. Hoffmann, 2001. An Object Based Approach to Urban Feature Mapping. . *In* Jürgens, Carsten (ed.) *Remote Sensing of Urban Areas.* Regensburger Geographische Schriften, Heft 35, Regensburg, pp. 100-106.

StoraEnso, 2002. *EnsoMOSAIC Mosaicking User's Guide*. Version 4.00. Stora Enso Forest Consulting Ltd.

Voi Municipal Council (= VMC), 1995. Local Authority Development Programme 1991-1995. Unpublished report, GTZ Nairobi.