# Object-Oriented Analysis of Very High Resolution Orthophotos for Estimating the Population of Slum Areas, Case of Dar-Es-Salaam, Tanzania

Mehdi Aminipouri

March, 2009

| | |
|---|---|
| **Course Title:** | Geo-Information Science and Earth Observation<br>for Environmental Modelling and Management |
| **Level:** | Master of Science (Msc) |
| **Course Duration:** | September 2007 - March 2009 |
| **Consortium partners:** | University of Southampton (UK)<br>Lund University (Sweden)<br>University of Warsaw (Poland)<br>International Institute for Geo-Information Science<br>and Earth Observation (ITC) (The Netherlands) |
| **GEM thesis number**: | 2007-23 |

Object-Oriented Analysis of Very High Resolution Orthophotos for Estimating the Population of Slum Areas, Case of Dar-Es-Salaam, Tanzania

by

Mehdi Aminipouri

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation for Environmental Modelling and Management

Thesis Assessment Board

Chairman: Dr. Richard Sliuzas, ITC
External Examiner: Prof. Katarzyna Dabrowska, University of Warsaw
Internal Examiner: Drs. Emile Dopheide, ITC
First Supervisor: Ms. Msc. Monika Kuffer, ITC
Second Supervisor: Dr. Richard Sliuzas, ITC

ITC International Institute for Geo-Information Science and Earth Observation
Enschede, The Netherlands

بسم الله الرحمن الرحيم

***Dedicated to my dearest wife and son***
***with love and gratitude***

**Abstract**

Unplanned development of urban areas and the creation of slum settlements are responses to rapid population growth and its density in fast growing cities like Dar-Es-Salaam, the former capital of Tanzania.

Regarding the scope of the Millennium Development Goals (MDGs) established by United Nations, especially Goal 7 target 11 *“the improvement of the quality of at least 100 million slum dwellers by the year 2020”*, one of the substantial steps in meeting this goal is to find reliable procedures for detecting and monitoring the slum areas. Hence, obtaining up to date spatial information about informal settlements is of great importance for any decisions that have to be made by urban planners.

Use of Very High Resolution (VHR) airborne and satellite imagery in meter or sub-meter level has generated a new era in proceeding of the information extraction of slum areas based on object-oriented techniques and estimation methods of the number of slum inhabitants.

The main objective of this research, therefore, is to determine the feasibility of using VHR orthophotos (0.6 meter pixel size) in creating an accurate inventory of buildings to enable the estimation of slum population based on the extracted building roofs.

*e*Cognition software is used for the image segmentation and classification of the objects of interest (roofs in this study). The accuracy of the roof extraction approach and consequently the estimation of slum population are examined on three different study wards called *“Charambe”*, *“Manzese”* and *“Tandale”* based on a sample size of n=550 reference polygons for each site. Two methods of accuracy assessment are utilised:

1. Considering a roof coverage threshold (25%) to take into account the co-registration errors (positional accuracy); and
2. Ignoring the coverage threshold and calculating the classification accuracies and Kappa statistic from the confusion matrices.

Applying the first method of accuracy assessment resulted in the followings:

In total, 1504 buildings are extracted out of 1650 reference buildings in the study areas which correspond to 91.1% accuracy in the extraction rate. Across all three study sample sites, the total roof area coverage extracted from *e*Cognition software is estimated at 136720 $m^2$ and from reference data 183195 $m^2$. This amount shows that 25.7% of the reference polygons are not covered by the extracted roofs. Thereby, 74.3% of the total population of these three wards is estimated by the applied model.

The second method of accuracy assessment resulted in the followings:

The average of overall accuracy for the building class in the study sites is satisfactory at 93% which corresponds to the Kappa value of 0.82. This Kappa value indicates that the object-oriented method used for the extraction of the roofs produced quite good results. From the total population (14616 inhabitants) in the study areas, 91.7% are estimated by the implemented model to be living there which is very close to the real number of slum dwellers.

In general, the methodology developed in this research eventuated in satisfactory results which can be helpful for authorities and managers at regional and national levels.

**Acknowledgements**

All the worship and adoration to almighty ***God***, the beneficent, the merciful. Without His divine grace the successful accomplishment of this research was impossible.
I would like to express my utmost appreciation to my parents for their total support. Also my deepest gratitude to the rock of my family, my wife, ***Maedeh*** for her immense encouragement and assistance during the fulfilment of this course. I would like to dedicate my immeasurable love to my new born son ***Amir Mohammad*** who enlivened us with his birth.

I am grateful to the Erasmus Mundus Consortium for providing this opportunity and awarding the scholarship. It is a privilege to study in four countries, explore different cultures and achieve a higher level of education by pursuing the GEM course, all at the same time.

I would like to convey my acknowledgments to all my teachers and to the coordinators of this programme in University of Southampton, Lund University, University of Warsaw and ITC.

I am deeply indebted to my supervisors: Ms. Monika Kuffer and Dr. Richard Sliuzas for their guidance, suggestions and constructive criticism in all the time of research and thesis writing. My sincere thanks to Dr. Mark Zuidgeest, ITC and Dr. Stanislaw Lewinski (Institute of Geodesy and Cartography) university of Warsaw, for their cooperation and contribution to my research.

My heartfelt thanks to all the GEMs. You were wonderful classmates indeed. Eighteen months of living, studying and travelling together yielded comradeship, fellow-feeling, cordiality and unity. During this 18 months I made some unique friends whom I like to appreciate the most: Aidin and Jose, you are like my elder brothers, more than just a friend and I learned a lot from you, *Mucho Gracias*. We spent delightful days together and I will never forget that. Daphne, Kath and Fernanda, you are such an honourable friends for Maedeh and I. It is undeniable that we made our English improved by your continual help.

Last but not the least, I wish to thank my sisters, little brother and my in-law family for their moral and mental support.

**Table of contents**

**List of figures**

**List of tables**

# 1. Introduction

As the beginning of the new century places pressure on governments and international organisations to provide health care, education, safety and security, developing countries are facing a serious challenge in provision of these basic requirements. Improving the quality of life for those people who are living in poor condition and below the poverty line is one of the most important issues for developing countries. Unplanned development of urban areas and consequently the creation of slum settlements have resulted from rapid population growth and its density in fast growing cities like Dar-Es-Salaam, the former capital of Tanzania.
The report of UN-Habitat, (2003b) on *"Slums of the world"* has revealed that Africa has the world's highest urbanisation rate with an annual average of urban growth of 4%, almost two times faster than that of Latin America and Asia. Currently 37% of the total population in Africa lives in cities and by 2030 the urban population is expected to rise to 53% of the total population. For Tanzania in particular, 92.1% (14,113,000 inhabitants) of the urban population live in slum areas and on the grounds that the slum population is predicted to double in eleven years in this country (Urban-Info, 2008), the slum crisis will get more complicated.
Regarding the scope of the Millennium Development Goals (MDGs) established by United Nations, especially Goal 7 target 11 *"the improvement of the quality of at least 100 million slum dwellers by the year 2020"*, one of the substantial steps in meeting this goal is to find reliable procedures for detecting and monitoring the spatial characteristics of slum areas. Thus, obtaining spatial information about informal settlements which is up to date is vital for any actions of enhancement in terms of urban and regional planning (Hofmann *et al*., 2008).
However, a first step to be able to quantify and locate the slum population is to develop an operational definition of the term *"slum"*. An UN-Habitat expert group meeting held in 2002 agreed on the following definition: *"A slum is a contiguous settlement where the inhabitants are characterised as having inadequate housing and basic services"* (UN-Habitat, 2003a). Since a slum is often not recognised and addressed by the public authorities as an integral or equal part of the city, little survey and census data can be found on slum dwellers. It is essential, hence, to develop a more comprehensive definition at household level in order to be able to identify slum dwellers among the urban population (Turkstra and Raithelhuber,

2004). A ***"slum household"*** is defined as a group of individuals living under the same roof lacking one or more of the conditions below (UN-Habitat, 2003b):

- Access to improved water;
- Access to improved sanitation facilities;
- Sufficient-living area, not overcrowded;
- Structural quality/durability of dwellings; and
- Security of tenure.

Martínez *et al.,* (2008) have illustrated that cities in Sub-Saharan Africa are generally worse off in terms of housing durability and access to safe water (see Appendix 1). Dar-Es-Salaam in Tanzania is an example of these Sub-Saharan cities; with nearly 65 % (1,987,700 slum inhabitants) of all households estimated to be living in slum regions (see Table 1). In addition, Dar-Es-Salaam is among the world's ten fastest growing large cities. That is the reason for considerable concern because history has shown that high urban growth rates in Africa tend to translate into significant urban informal settlement and slum formation (UN-Habitat, 2008), (see also Appendix 2).

**Table 1- Population and Growth Rate of Dar-Es-Salaam (source: UN-Habitat, (2008))**

| ***Statistics*** | | | |
|---|---|---|---|
| *Total Population* | *Slum Population* | *Urban Annual Growth* | *Number of years the slum population will double* |
| 3,058,000 | 1,987,700 | 4.29 % | 15 |

Such statistics should be used to inform politicians and governors about the state of slum development and should provide information to policy makers on the effectiveness of interventions that are aimed at alleviating urban poverty through slum improvement or other such strategies (Sliuzas and Kuffer, 2008).

Hence, it is of great value to investigate the physical characteristics of slum areas, the ways which lead to the information extraction of informal settlements and estimation of population living in these unplanned regions, particularly for the case study of Dar-Es-Salaam. This investigation will also aid in monitoring the Millennium Development Goals (MDGs) of the United Nations, the improvement of the quality of at least 100 million slum dwellers by the year 2020 (UN-Habitat, 2003b).

## 1.1. Background

Use of Very High Resolution (VHR) airborne and satellite imagery in meter or sub-meter level has generated a new era in proceeding of the information extraction of slum areas based on object-oriented techniques and estimation methods of the number of slum inhabitants. In this section various articles and relevant research materials that comprised the use of object-oriented approach have been reviewed to find a reasonable and cost-effective way of information extraction. This section is divided into two parts:

- Studies linked to object-based information extraction (information related to buildings in particular); and
- Studies related to population estimation methods.

### 1.1.1. Object-Oriented Information Extraction Methods

Slum neighbourhoods are defined largely by their physical and infrastructural environment, rather than by the demographic characteristics of residents. Nonetheless, the important aspect of slum is that it is a characteristic of place, not specifically of the people residing there (Weeks *et al.*, 2007). For this reason it may be possible to identify such places from satellite imagery.

Along with the increasing availability of new high resolution satellite and airborne digital imagery, precise extraction of ground objects, instead of regions of certain land cover classes, has become increasingly important for a variety of remote sensing and GIS applications (Zhang and Maxwell, 2006). However, traditional pixel-based classification methods, such as Maximum Likelihood and some Neural Networks based approaches can hardly produce satisfactory classification results for identifying individual objects, because an object in high resolution imagery is usually composed of heterogeneous pixels with different spectral attributes (Zhang and Maxwell, 2006).

Several authors have described different methods for the analysis and classification of satellite imagery using different techniques such as pixel-based and object-oriented approaches which can be adjusted to fit the needs of this research. In a pixel-based approach conducted by Jain, (2007) image fusion technique of multispectral and panchromatic IKONOS satellite data has been used to increase the interpretation capabilities for the identification of informal settlements in a part of Dehradun, India. The research showed the potential of the methodology to be used to extract areas of poverty in regional scale within urban districts. However, on the grounds that the focus of this research is on using very high spatial resolution orthophotos, extracting information of slum buildings will not be possible from the pixels in this level of resolution, since they do not hold enough information and cannot satisfy high resolution images' classification precision to detect different phenomena (Wei *et al*., 2005). Whereas, for the detection of informal settlements textural information is essential and object-oriented methods offer new possibilities to tackle these problems (Hofmann, 2001b).

Depending on the imaging scale and the detail of spatial information to be obtained, detecting informal settlements can vary in its aim: the aim can be either detecting single shacks or just locating and outlining the informal settlement areas spatially (Hofmann, 2001a).

Giada *et al.,* (2003) have stated that the extracted information from high resolution satellite imagery for identification of refugee tents as well as their spatial extent, based on object-oriented classifiers performs best with less than 3 % errors, while the pixel-based methods yield results with errors between 10% to 15% (all with 95% confidence interval error margin).

Cleve *et al.,* (2008) have demonstrated that an object-based approach, using high resolution aerial photography provided 41.73% greater accuracy than pixel-based classification for built area category.

A case study of Dar-Es-Salaam done by Mayunga *et al.,* (2007), have explicated the potential and capability of VHR images in extracting building roofs in informal settlements by comparing two different satellite images. They explained that the information contents (buildings in particular) extracted from QuickBird imagery coincide with the investigation made by Baltsavias and Mason (2001) on information contents of IKONOS imagery.

Beside the use of object-oriented technique for extracting urban features such as roads and buildings, this method has been recently applied in a wide range of application fields and have proven their success, see examples in: urban environment (Mathieu *et al.,* (2007b); Mathieu *et al.,* (2007a)); land cover/land use mapping (Akbari *et al.,* (2003); van der Sande *et al.,* (2003); Carleer and Wolff, (2006)); landscape pattern (Zhou and Troy, 2008); biodiversity (Laliberte *et al*., 2004); ship routing and monitoring (Willhauck *et al*., 2005); monitoring mining activity (Pagot *et al*., 2008) and reviving legacy population maps (Kerle and Leeuw, 2008).

### 1.1.2. Population Estimation Methods

Apart from all the methods and techniques for detecting, monitoring and mapping informal settlements, timely and accurate estimation of the slum population is an important task for governments, decision makers and for the relief community.

As described by Lo, (2008) population estimation using remotely sensed data has a long history that can be traced back to the 1970s when aerial photographs were used to provide intercensal estimates of population or to check on the census accuracy, as illustrated by the work of Kraus *et al*., (1974) and Iisaka and Hegedus (1982).

Basically, population estimation from remotely sensed data can be carried out using four approaches (Lo, 1986):

- Counts of dwelling units;
- Measurement of areas of urbanization;
- Measurement of areas of different land-use; and
- Automated digital image analysis.

Lo's first approach (counting of dwelling units) is now becoming feasible with the general availability of orbitally acquired imagery of adequate spatial resolution (Harvey, 2002). However, dwelling unit counts and corresponding in situ information on the number of persons occupying each dwelling unit method is a more complicated but potentially more accurate method of estimating population via remote sensing (Watkins and Morrow-Jones, (1985); Bjorgo, (2000)). Thereby, the total population of an area can be derived by multiplying the total number of dwelling units with the number of persons normally living in a dwelling unit (Wu *et al*., 2005). The number of inhabitants living in a dwelling unit can be obtained from household surveys or from census data. Strictly speaking, models from this group are mainly designed to estimate an overall population count rather than population density that is relevant to population distribution (Wu *et al*., 2005).

Regardless of the methods that are used in statistical modelling for population estimation, all studies inferring population from remotely sensed data and the accuracy of their results largely rely on the size of the area. The small-area population estimation is often not as accurate as large-area estimation. It may be explained that over-estimation and under-estimation are cancelled out for large-area population estimation and thus overall accuracy is high (Lo, 1995).

### 1.2. Problem Statement

Among the several issues that organisations and institutions at local, national and international level are trying to conquer, rapid growth of slum areas in Dar-Es-Salaam during past decades, increase in number of slum dwellers and the lack of intercensal data at the appropriate spatial level are chosen as the most important research problems of this work. Therefore, it is worthwhile to search for methodologies and techniques that offer effective solutions.

Tackling the mentioned problems, acquiring information from slum areas and reflecting their spatial heterogeneity require the identification of multiple objectives. Based on the literature review, the following research objectives have been developed (see Section 1.3).

## 1.3. Research Objective

The main objective of this research is to determine the feasibility of using VHR images in creating an accurate inventory of buildings to enable the estimation of slum population based on the extracted buildings roof.

### 1.3.1. Specific Objectives

Meeting the end of the main goal will be possible by setting a number of specific objectives and breaking down the general goal into different aspects that can be formulated as follows:

- To utilise object-based image classification to extract slum buildings from the VHR orthophotos;
- To evaluate the roof extraction quality in urban fringes or low-density areas and to compare it with high-density urban slums; and
- To assess whether the extracted information on building statistics can be applied to estimate the number of slum inhabitants.

## 1.4. Research Questions and Hypotheses

In view of the fact that previous researches had not implemented the technique of estimating the slum population, using very high resolution orthophotos and object-oriented analysis, it is suggested to bridge this gap by answering the main research questions in the current work which are listed below:

a) How will be the roof extraction quality in urban fringes or low density areas compared with high density urban areas?

Hypothesis:
$H_0$: There is no difference in the overall accuracy for low and high density areas.

b) Does the occupancy rate differ from ward to ward?

Hypothesis:
$H_0$: The ratio remains constant irrespective of location of slum wards.

c) What will be the correlation between the population of slum inhabitants and roof area of the buildings like?

Hypothesis:
$H_0$: Size of the population is not proportional to the roof area of buildings in slum areas.

## 1.5. Research Approach

Figure 1 demonstrates the research approach of this study which is pursued by the author.

## 1.6. Structure of the Thesis

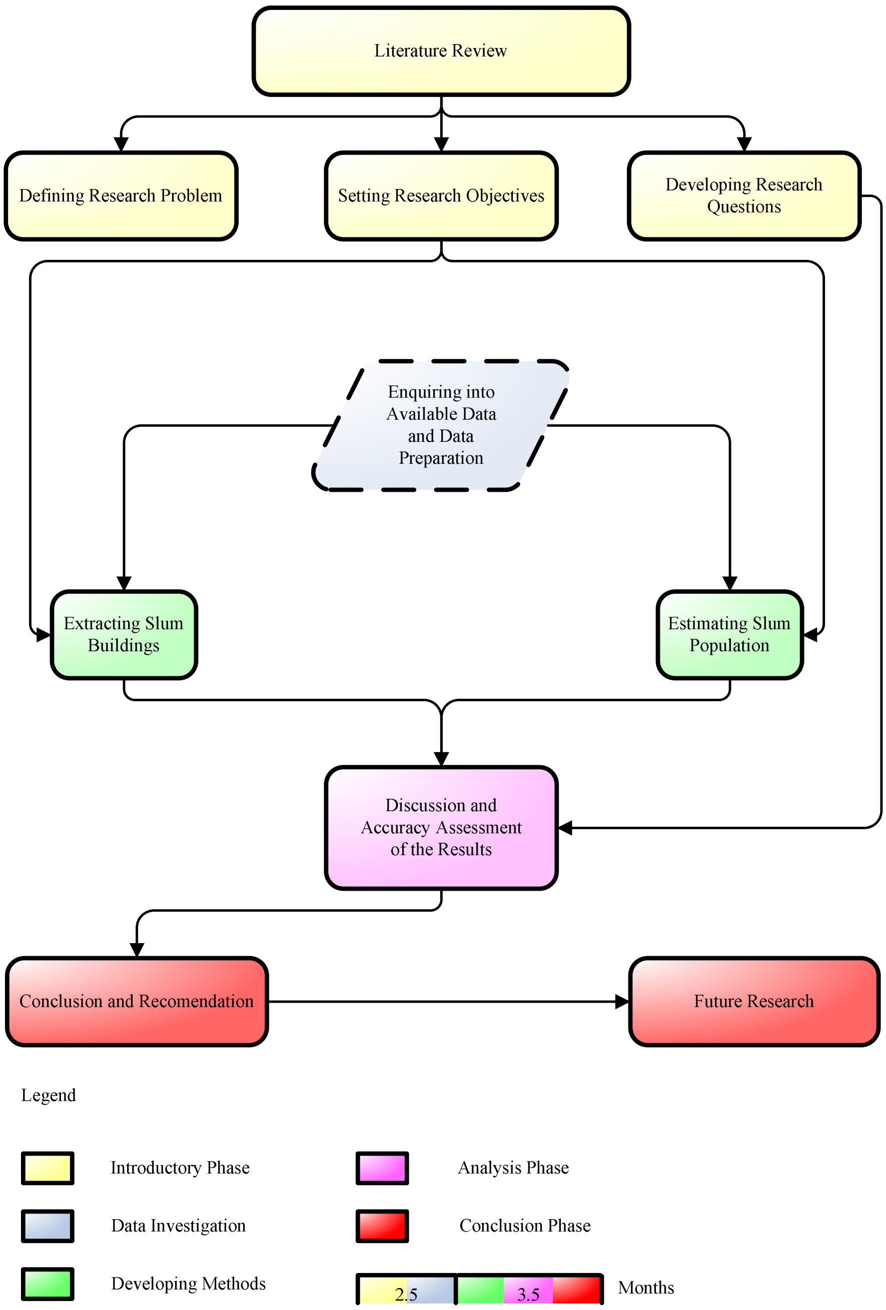
**Chapter 1** Introduction
The chapter provides an introduction to the slum situation in the World and especially in Africa, followed by a literature review in the background part. It also describes the problems and facing challenges, the objectives and the research questions and hypotheses.

**Chapter 2** Materials and Methods
The first part of this chapter gives an overview about the study area, the data and the software that are used to develop the methods. Furthermore, limitations of the data and characteristics of slum buildings in the study areas are noted. The second part explains the methods which lead to the extraction of slum building and the estimation of slum population.

**Chapter 3** Results and Discussion
Results of the object-oriented image analysis are evinced in this chapter; thereby, the outcomes are statistically analysed and their accuracy are assessed and discussed.

**Chapter 4** Conclusions and Recommendations
This chapter is arranged to make conclusions based on an overall review of the results. Further the possible applicability of the implemented method for future studies is evaluated in the recommendation part.

**Figure 1- A basic flowchart showing the entire research approach**



# 2. Materials & Methods

Section 2.1 to section 2.6 provides an overview about the study area, the data and the software that are used to develop the methods. Furthermore, limitations of the data and characteristics of slum buildings in the study areas are noted. In section 2.7 and 2.8 the methods which lead to the extraction of slum building and the estimation of slum population are explained.

## 2.1. Study Area

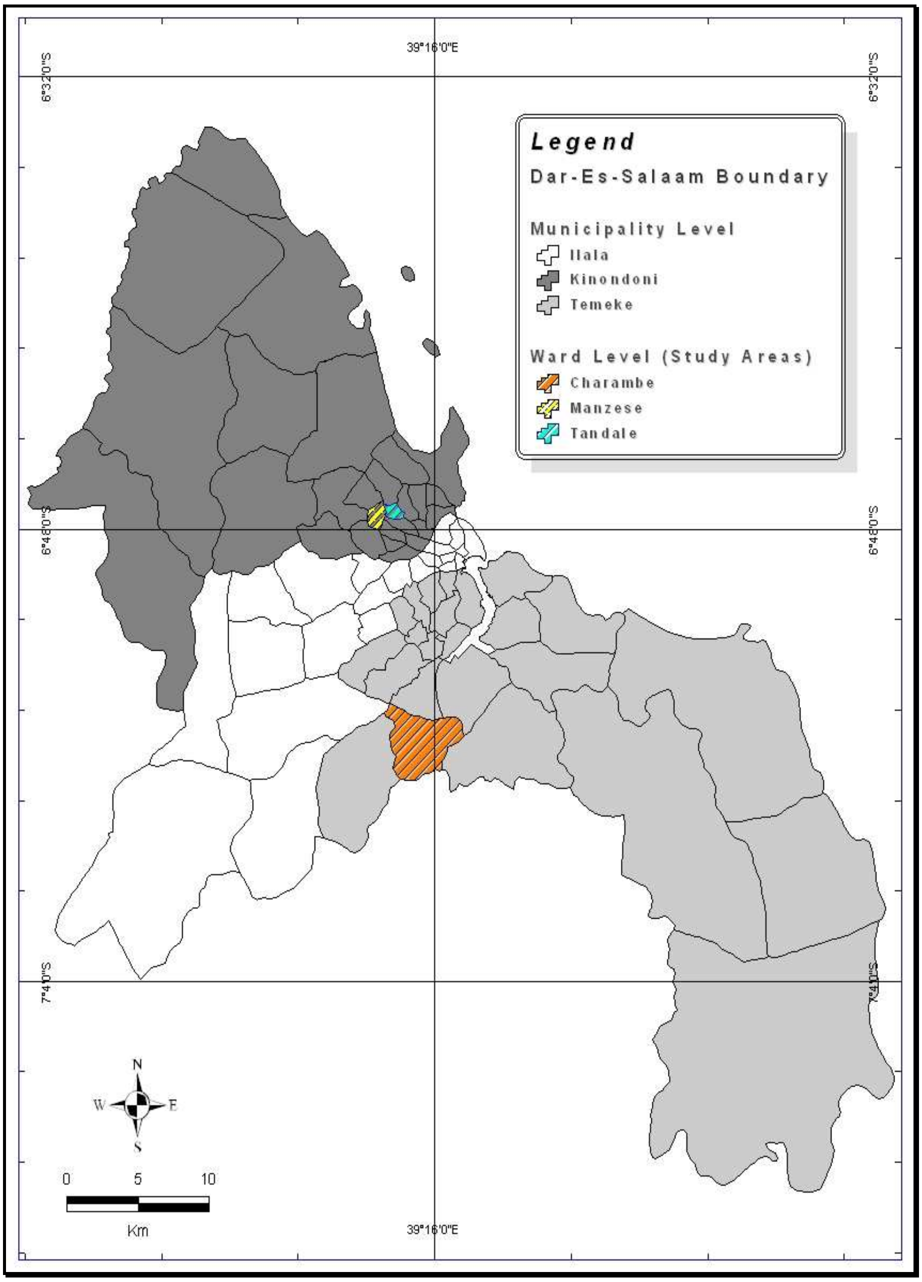
*Why Dar-Es-Salaam City?*

Most Urban centres in Tanzania have been experiencing high urbanisation trends and uncontrolled growth of informal settlements over the past four decades. However, Dar-Es-Salaam is the main centre accommodating more than a quarter of all urban informal residents. Therefore, the choice of Dar-Es-Salaam is made for two main reasons. *First*, Dar-Es-Salaam is the largest seaport, industrial, commercial and administrative centre in Tanzania. Because of the underlying socio-economic and political interests, it displays more urban infrastructure and housing problem than any other urban centre in Tanzania. *Secondly*, because of its status and role at national and international levels, it is easier to get access to data and information in land servicing in informal settlements (Kyessi, 2002). The method that is expected to extract slum houses and estimate slum population is developed for three wards called *"Manzese"*, *"Tandale"* and *"Charambe"* in two different municipalities in Dar-Es-Salaam, Tanzania (see Figure 2). The selection of these three wards is based on the rich experience and local knowledge of my supervisors.

The study areas are characterised by slum settlements of which two of them called Manzese and Tandale are recognised as high density slum areas, located in Kinondoni municipality. Charambe is the third selected study site, a medium size ward situated in Temeke municipality. While Manzese and Tandale are considered as older slum establishments in Dar-Es-Salaam (Kuffer, (2001); Sliuzas *et al*., (2004); Amer *et al*., (2007)), Charambe has relatively new settlements and it is less dense.

Nonetheless, not the entire area of the selected wards has been covered in the research given the limited time and resources.

**Figure 2- Localisation of study areas: selected wards within the municipalities of Dar-Es-Salaam**



## 2.2. Input Data

The datasets used in this research (assuming the provided dataset is accurate) are listed below:

- Very high resolution (0.6 meter pixel size) digital orthophotos of the year 2005. The image dataset has red, green and blue bands; geographically projected (by the image provider) to the UTM zone 37S (WGS-84) with WGS-84 Datum;
- Administrative boundaries at municipality, ward and sub-ward level created by Dar-Es-Salaam City Council based on aerial photographs and 1:20,000 topographic maps;
- A shapefile (polygons) dataset including the land parcels of squatters of Dar-Es-Salaam;
- A comprehensive geodatabase comprising attribute tables of land parcels, building uses with the number of occupants, sources of income etc., for the municipalities of Dar-Es-Salaam; and
- A collection of CAD files (DXF format) created based on aerial photographs and 1:2,500 topographic maps in the year 1999. In the maps at this scale individual buildings (outline or solid) can be shown clearly (Polle, 1996). The DXF files are already converted into features (polylines) by the data provider. The building centroids (points) and their roof areas for the entire city are included in the dataset.

## 2.3. Software

In order to develop the methodology of this study, three software technologies are employed as follows:

- **Definiens *e*Cognition 7:** *e*Cognition is used to implement the image segmentation and object-based classification procedures.

  *Why eCognition?*

  Segmenting an image into meaningful objects is the fundamental step in object-based image analysis; in that the total accuracy of object-oriented classification is highly dependent on the quality of image segmentation. By doing so informative attributes (shape, colour, texture, etc.) can be created. A comparison of segmentation programmes for high resolution remote sensing data (representing an urban and a rural landscape) conducted by Meinel and Neubert, (2004) has shown that *e*Cognition has a high potential in creating meaningful objects due to its multi-scale segmentation and the fuzzy logic based image classification capabilities. In the mentioned research seven segmentation programmes were compared: *e*Cognition 3.0, Data Dissection Tools, CAESAR 3.1, InfoPACK 1.0, Image Segmentation for Erdas Imaging, Minimum Entropy Approach to Adaptive Image Polygonisation and SPRING 4.0. The comparison showed that the best segmentation results have derived from *e*Cognition and InfoPACK software. However, the use of InfoPACK leads to more over-segmented results. In addition, the ability of exporting GIS-readable vector (shapefile) and raster format in *e*Cognition is considered as an advantage for this specific software.

Therefore, among the several programmes which perform the image segmentation technique, *e*Cognition software developed by Definiens Imaging (Definiens, 2007b) is chosen for the purpose of this study.

- **ESRI ArcGIS 9.3:** is used because of its capabilities and facilities in dataset management, pre-processing of raster and vector data, mapping as well as comparative analysis.

- **Microsoft Office Excel 2003:** is used for the statistical analysis part.

### 2.4. Data Preparation

The VHR orthophotos are in MrSID format (filename extension .sid) and are not compatible with Definiens *e*Cognition software. Thus, they are exported as *.IMG files (ERDAS Imagine format) to be in a compatible and readable format. Since only portions of the study areas were needed to run the image analysis part, the images were then subset to smaller extents (1000 x 800 meters) to save on computation time and also to fit in the extent of the squatter's shapefile dataset. The image subsets are assumed to be a representative region of each of the corresponded wards.

The land parcel attribute table included in the geodatabase comprises ample and useful information which two of the most important attributes are the number of inhabitants and buildings in each land parcel. These attributes would be even more valuable in conjunction with the shapfile dataset of squatters; in that, it would be possible to know exactly how many people are living in each parcel. Inasmuch as both squatters dataset and land parcel attribute table have the same and unique code for each land parcel, by using ArcGIS it is practicable to join the land parcel attribute table to the squatters dataset based on the above mentioned code.

Finally the DXF files are converted from polylines to polygons using the *"feature to polygon"* tool in ArcToolbox in order to make the roofs as polygons for future analyses. Digitisation, nevertheless, is accomplished as a supplementary process to make the roofs in those areas where the reference polygons are not available. The digitisation procedure is assumed error-free. The building roofs are then used as reference data for the accuracy assessment of the results.

### 2.5. Limitations of the data

The below statements are considered as the limitations of the data:

- It is challenging to use digital orthophotos. It would have been more functional to use high resolution multispectral (at least with NIR band) images with original DN (Digital Number) values to take advantage of the spectral signature;
- As a rule of thumb, buildings are elevated objects. The most apparent information about buildings is their elevation compared to their surroundings. Hence, lack of Digital Surface Model (DSM) is observed as a constraint. The usefulness of DSM in object-based classification has been shown in different studies (Mavrantza and Argialas, (2007); Baltsavias and Mason, (2001); Sohn and Dowman, (2007));
- Unavailability of up to date information, both raster (orthophotos, 2005) and vector data (roof coverage, 1999); although it has to be mentioned here that the roofs coverage for Manzese and Tandale wards (two of the oldest slum wards) have not been changed very much since 1999. However, this issue is problematic in new urban growth areas like Charambe ward. Moreover, some problems emerged because of the time difference in the dataset which is used for this study. The dataset of squatters of Dar-Es-Salaam is for the year 2006, while the digital orthophotos are from 2005. Thus, there are some occupied land parcels with certain number of inhabitants where there is no building in the image for that particular land parcel. Thereby, the mentioned land parcels are manually eliminated from the analysis process;
- Co-registration errors caused by the scale difference between the VHR orthophotos (0.6 m pixel size), the aerial photographs and topographic maps (1:2,500) that has been used for the creation of roof polygons; and
- Despite the fact that the whole provided datasets are sufficient for the needs of this study, although they cause some issues, not being able to collect ground-truth data in the field but to use digital data sources as ground truth causes some limitations. Gaining local knowledge and field observations can provide valuable information about the urban slums. The ideal situation

would be to combine both data collection in the field and other sources of digital data (Lemma *et al*., (2006); Kuffer, (2001)).

### 2.6. Characteristics of Slum Buildings

In order to formulate a strategy for extraction of the slum buildings from imagery, it is essential to answer this question: what characteristics of slum can be derived from remotely sensed data? It has to be considered that characteristics of slum settlements can most efficiently be derived through visual interpretation, integrated and combined with GIS analysis of the provided dataset. The process of visual interpretation of VHR orthophotos is not automated, but requires expert knowledge, systematic search and additional data to ensure effective identification of the target features (i.e., buildings). It is very time consuming and is not feasible for searching over extensive areas (Ward and Peters, 2007).

However, the answer to the question can be given by exploring the urban ecosystem aspect and physical characteristics in conjunction together. Vegetation, impervious surface and bare soil are three main components and the basic characteristics of an urban scene derived from satellite imagery (Ridd, (1995); Stow *et al*., (2007)). Generally, it is expected that slums will have a combination of impervious surface (largely from roofs) and bare soil (especially streets and walkways) that would indicate residential areas in a Sub-Saharan city, but commonly characterised by a general lack of vegetation, at least in comparison to other residential areas (Weeks *et al*., 2007).

Nonetheless, the proportional abundance of each of these types may vary from city to city and from ward to ward. High levels of impervious surface (mostly from roofs) is an indicative of dense slum wards such as Manzese and Tandale, while this is not the case for Charambe as can be seen in Figure 3 and Figure 4. Charambe ward, in contrast, is less dense, the spacing between the houses is larger as that of Manzese and Tandale wards and has more area covered by vegetation and bare soil.

Typically, the slum settlements in the selected wards are characterised by the following properties:

- Single storied and pitched-roof buildings; in some cases with near-horizontal roofs;
- Simple geometry. An investigation of building shape typology in roof polygons dataset (reference digitised data) showed that in Charambe, Manzese and Tandale wards, most of the roofs are 4-sided, (see Table 2);
- The majority of the roofs are constructed from corrugated iron and zinc sheets with variable textures. The colour of the roofs occur mostly bright grey to bright blue and in some parts bright red to dark brown; and

- Irregular patterns of buildings without any ordered orientation.

**Figure 3- Typical slum settelments in Manzese and Tandale wards**

**Figure 4- Typical slum settlements in Charambe ward**

Gaining basic knowledge of the study site is of particular importance for developing and setting up rules and methods for the image processing stage. Next section will focus on the consideration of the above mentioned physical characteristics of slum settlements to outline the procedure of object-oriented classification.

**Table 2- Summary of Physical Characteristics of Slum Buildings in the study areas**

| Ward | Roof Area ($m^2$) | | | Roof Geometry | | Roofing | | Building Use | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Min | Ave | Max [1] | 4-Sided | More | Metal Sheet | Other [2] | R | C/R | Other [3] |
| ***Charambe*** | 13 | 118 | 325 | 86% | 14% | 95% | 5% | 81% | 2% | 17% |
| ***Manzese*** | 12 | 108 | 660 | 85% | 15% | 92% | 8% | 77% | 12% | 11% |
| ***Tandale*** | 12 | 105 | 353 | 88% | 12% | 92% | 8% | 84% | 7% | 9% |

Ave: Average

R: Residential

C/R: Commercial/Residential

1- The big roof area numbers belong to few attached buildings making residential complexes.
2- Other roofing materials: Asbestos Cement, Thatch (coconut leaf), Roof Tile and so forth.
3- Other building uses (not included in this study): mosque, church, school, shed, vacant plot, play ground and factory.

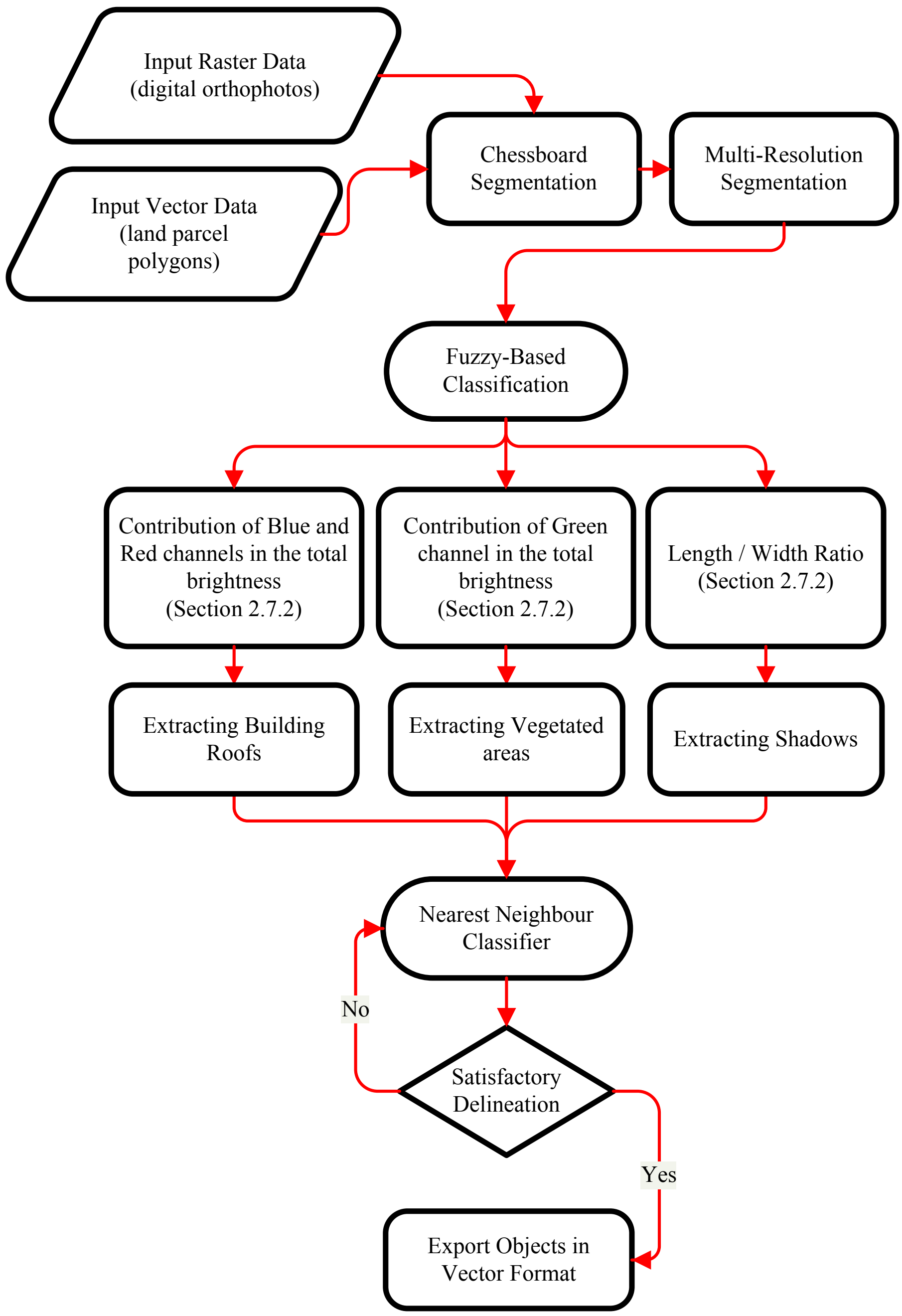
## 2.7. Object-Oriented Image Analysis

Unlike traditional pixel-based methods, an object-oriented method for image analysis treats the image as a set of meaningful objects rather than pixels. Objects are defined on the basis of the internal homogeneity of their spectral values, the contrast with neighbouring objects, their own spatial and spectral characteristics, or a combination of these three properties (Giada *et al*., 2003).

Object-oriented image analysis has the advantage of using a hierarchical network of image objects, so-called image segments. Advantages of object-oriented analysis are meaningful statistic and texture calculation, an increased uncorrelated feature space using shape (e.g. length, asymmetry, etc.) and topological features (neighbour, super-object, etc.) and the close relation between real-world objects and image objects. This relation improves the value of the final classification and cannot be fulfilled by common, pixel-based approaches (Benz *et al*., 2004).

The analysis plan for meeting the objectives of this study is a step-wise process in *e*Cognition software (Definiens, 2007b) that has been depicted as a flowchart in Figure 5. A detailed description of each step is provided in the following sections.

**Figure 5- Object-oriented image analysis workflow**

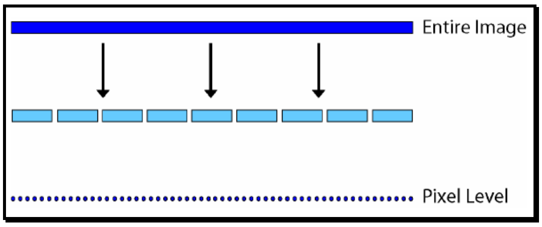


### 2.7.1. Image Segmentation Process in *e*Cognition

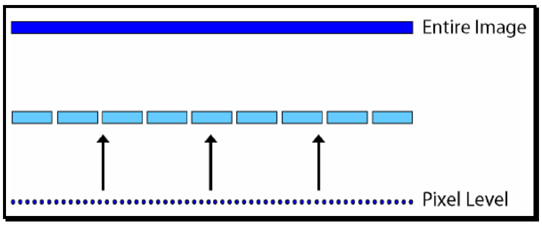
Defining meaningful objects is an immediate task for the feature extraction method. From the point view of remote sensing, most often a building object is actually visible as the building's roof (Tian and Chen, 2007). Owing to the fact that the main goal of this study is to estimate the number of slum dwellers based on object-oriented image analysis, the extraction of the slum building roofs will be examined as the preferred type of object.

In *e*Cognition software, objects are extracted from the image in a number of hierarchical segmentation levels. The term "*segmentation*" means: *"an operation that creates new image objects or alters the morphology of existing image objects according to a given criteria"* (Definiens, 2007c). The segmentation process can be done either by top-down or bottom-up strategies (see Figure 6).

**Figure 6- Two different segmentation strategies in *e*Cognition (source: Definiens, (2007c))**




a) Top-down strategy: cutting objects into smaller objects




b) Bottom-up strategy: assembling objects to create larger objects

#### 2.7.1.1. Chessboard Segmentation

*e*Cognition software provides the ability to import vector data. Specifically, attribute tables (even from single attribute or record) may be imported as complementary information. By doing so, several objects representing the land parcels are generated for the subset scenes based on an approach called *"Chessboard Segmentation"* (*CS*) which is a top-down segmentation strategy (Definiens, 2007c).

Figure 7- Delineation of land parcels in a sample study site

As an assumption, only land parcels with residential and commercial/residential usage are considered for this study. The land parcel vector file accompanying by the two mentioned usages are included in the *CS* process. Figure 7 illustrates the delineation of the residential and commercial/residential parcels in different views. The added thematic layer into the chessboard segmentation process resulted in additional splitting of image objects while enabling consistent access to its thematic information. This function drastically reduces the number of units to be handled for the upcoming image classification; consequently, it lessens the treatment time.

The most important parameter in *CS* which has to be defined by the user is the object size. In order to produce image objects based exclusively on land parcel thematic layer, the object size must be selected larger than the image size (Definiens, 2007a). In other words, the object size must be set to a number which guarantees that one tile is generated only. Since the subsets of the study areas are hardly larger than 2000 by 2000 pixels, a setting of 5000 pixels is on the safe side (see Table 3).

**Table 3- Chessboard segmentation object size parameter for the analysis**

| ***Ward*** | ***Image Resolution (m)*** | ***Size of the Image Subset (m)*** | ***Number of Pixels in the Subsets*** | ***Chessboard Object Size (pixels)*** |
|---|---|---|---|---|
| *Charambe* | 0.6 | 1000 x 800 | 600 x 480 | 5000 x 5000 |
| *Manzese* | 0.6 | 1000 x 800 | 600 x 480 | 5000 x 5000 |
| *Tandale* | 0.6 | 1000 x 800 | 600 x 480 | 5000 x 5000 |

### 2.7.1.2. Multi-Resolution Segmentation

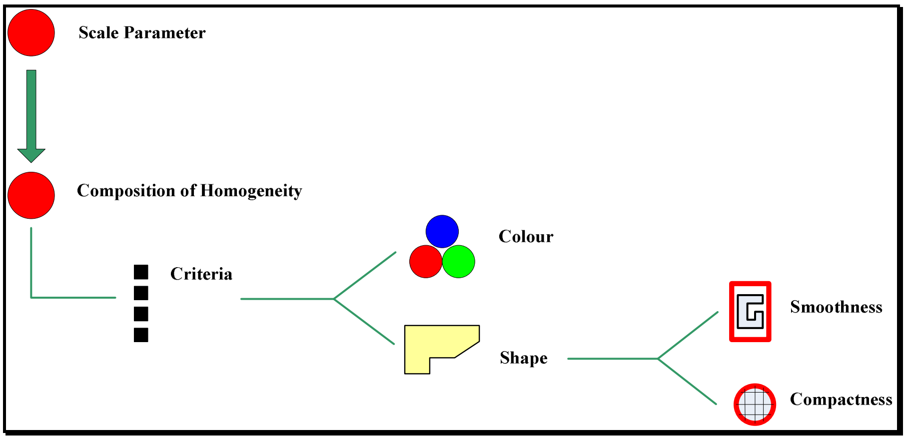
The next step after mapping out the residential and commercial/residential land parcels is to perform a multiscale image analysis method. This method is well-known as "*Multi-Resolution Segmentation*" (*MRS*) technique embedded in *e*Cognition software (Baatz and Schape, 2000). Segmenting an image by this approach is based on a set of criteria grouped into a single image parameter called "*Heterogeneity*". Three criteria are used to describe image object heterogeneities: colour criterion, smoothness criterion and compactness criterion. The smoothness and the compactness criteria are additionally summarized to the shape criterion. The colour and shape criteria can be weighted according to the type of object to be classified. Scale parameter is the most determining factor of *MRS* with regard to the size of objects (Su *et al*., 2008). Scale parameter is an abstract term which determines the average image object size; the higher the scale parameter value, the larger the image object becomes (Definiens, 2007c). Details are to be found in Baatz and Schape, (2000).

*MRS* is a bottom-up segmentation strategy, where the smallest object contains one pixel. In subsequent steps, smaller image objects are merged into larger ones based on the chosen scale, colour and shape parameters. Each parameter can be weighted from 0 to 1. Within the shape setting, smoothness and compactness can also be weighted from 0 to 1. The *MRS* approach allows for segmenting an image at different scales, which is used to construct a hierarchical network of image objects. The image objects know their horizontal neighbours (adjacent objects on the same level) as well as their vertical neighbours (objects on different hierarchical levels); the latter also termed sub-objects and super-objects (Laliberte *et al*., 2004) (see Appendix 3 for object hierarchical network in detail). The multi-resolution concept flow diagram is shown in Figure 8.

In addition to scale parameter, shape and colour criteria, the contribution of the input image channels in segmentation process is of great importance. Image channels can be weighted differently; nevertheless, when working with image data of comparable channels in size and content such as the image data of this study, each channel should be weighted equally (Hofmann, 2001b).

The *MRS* parameters for the study sites are assigned to achieve a realistic segmentation of the land parcels super-objects derived from the *CS*, such that the roofs of the smallest dwelling units are delineated. Characterisation of an image object based on its super-objects, e.g. dwelling units belonging to a super-object residential land parcel will be classified as residential house, whereas the rest of the objects will remain unclassified. The resulting *MRS* parameters are summarised in Table 4.

**Figure 8- Multi-resolution segmentation work flow diagram**




**Scale Parameter:** Defines the maximum standard deviation of the homogeneity criteria in regards to the weighted image layers for resulting image objects.
**Composition of Homogeneity:** defines the total relative homogeneity for the resulting image objects.
**Criteria:** each pair of criterion can be weighted from 0 to 1.
**Colour:** digital value (colour) of the resulting image objects (Colour = 1 – Shape)
**Shape:** defines the textural homogeneity of the resulting image objects (Shape = Smoothness + Compactness)
**Smoothness:** optimises the resulting image objects in regard to the smooth borders within the shape criterion (Smoothness = 1 – Compactness)
**Compactness:** optimises the resulting image objects in regard to the overall compactness within the shape criterion
(source: Definiens,(2007a))

All parameters of *MRS* are assigned through a trial-and-error but systematic experimentation. First of all, the segmentation parameters are defined for Charambe study area where separable buildings are available and thus the visual investigation through the entire image subset is straightforward. The tested scale parameter is ranged from 50 to 10. Eventually, the scale parameter of 15 gave the most visually pleasing result for Charambe. Since the residential houses in this region are relatively new, they look more homogeneous in terms of roof's colour compared to the building roofs in Manzese and Tandale. Thus, the colour parameters representing 90%, 20% and 20% are defined for Charambe, Manzese and Tandale, respectively. The same trial-and-error but systematic strategy is implemented for assigning the *MRS* parameters of Manzese and Tandale.
It has to be noticed that the changes in different *MRS* parameters have taken place simultaneously. In other words, several combinations of the parameters are tested and one optimum set of parameters is selected for each study area based on the

visual inspection of the resulting objects, as also described by Im *et al*., (2008) and Nobrega *et al*., (2005).

**Table 4- Multi-resolution segmentation parameters for the analysis**

| ***Ward*** | ***Scale parameter (no unit)*** | ***Shape [1] (%)*** | ***Colour (%)*** | ***Image layer weight*** | | |
|---|---|---|---|---|---|---|
| | | | | ***R*** | ***G*** | ***B*** |
| ***Charambe*** | 15 | 10 | 90 | 1 | 1 | 1 |
| ***Manzese*** | 10 | 80 | 20 | 1 | 1 | 1 |
| ***Tandale*** | 10 | 80 | 20 | 1 | 1 | 1 |

1- Smoothness and Compactness factors are weighted equally within the shape parameter.

As the results of the image segmentation strongly depend on the image data and the assessment of the segmentation results depends on the classification task, it is almost impossible to suggest well-suited segmentation parameters in general (Hofmann, 2001b).

### 2.7.2. Extraction and Classification of Image Objects

The real strength of object-oriented analysis lies in a combination of multi-scale segmentation with subsequent contextual analysis, whereby the spatial, spectral, and contextual properties of extracted segments at different spatial scales are used in conjunction with spatial rules in a subsequent classification (Sliuzas *et al*., 2008).

The object features provided in *e*Cognition software supplies a large number of shape, spectral and textural attributes that are used as sources of information to define the inclusion or exclusion parameters used to classify image objects. Object features are obtained by evaluating image objects themselves as well as their embedding in the image object hierarchy. In a group of six, object features are as follows (Definiens, 2007c):

1. User-customised features;
2. Layer values: this feature describes image objects with information derives from object spectral properties;
3. Shape feature: the basic shape feature are calculated based on hierarchical structure;
4. Texture: this feature performs based on an analysis of sub-objects;
5. Hierarchy: This feature provides information about the embedding of the image object in the image object hierarchy; and

6. Thematic attributes: the object's thematic properties may be evaluated by this function.

Each object feature group comprises of several sub-groups (e.g., mean layer, brightness, asymmetry, width, existence of super or sub-object). Image objects may be extracted and classified by combining one or more of these features. The classification of the objects in *e*Cognition software can be done by two methods: fuzzy membership functions and nearest neighbour classifier. While the fuzzy membership functions describe intervals of feature characteristics wherein the objects do belong to a certain class or not, the nearest neighbour classifier describes the classes to detect by sample objects for each class which the user has to determine (Hofmann, 2001b).

Firstly, the fuzzy membership thresholds are implemented for the object features of group 2 and 3 to describe and classify different classes. The ratio of the bands (see Equation 1) is used in the subsequent building roof extraction process based on a machine-learning decision tree. Moreover, length to width ratio function (see Equation 2) is used just for the separation of the shadows casted by the buildings. Not all the object features provided by *e*Cognition are used for the classification of the image objects. All features concerning texture (group 4) are based on sub-object analysis (Definiens, 2007a) that is not applicable for this research, since all the analyses in this study are on the basis of land parcels super objects. Object features from group 5 and 6 are already embedded in the development of the classification process, but not as separate object features.

For most of the roofs in the study areas, the ratio of blue channel is comparatively high, see also Yuan and Bauer, (2006). However, the ratio of red channel is also used to include the roofs in red colour in the extraction procedure. Although the contribution of vegetation and green areas in slum urban environment is low, they must be masked out as a different class in the process. In some cases the trees are misclassified as buildings and this is because of its reflectance also in blue channel which has been applied for the extraction of the rooftops. Hence, the ratio of the green channel is used exclusively for the separation of the vegetated areas inside each land parcel.

Referring to section 2.6 most of the buildings in the study wards are single storied. Thus, in terms of geometric characteristics, shadows casted by the buildings are narrow in width. The length to width ratio, therefore, would be an appropriate class descriptor for the separation of the shadows, although the ratio varies with the sun azimuth and elevation angles.

Fuzzy membership functions are incorporated at the first level of object class categorisation process. Building rooftops, vegetated areas and shadows are filtered out corresponding to the fuzzy thresholds that have been set for each class.

**Equation 1- The "Ratio" object feature in *e*Cognition**

**Ratio:**
The ratio of layer K reflects the amount that layer K contributes to the total brightness.

Parameters:
$W_K^B$: Brightness weight of layer K
$\overline{C}_K(v)$: mean intensity of layer K of an object v
$\overline{C}(v)$: brightness

**Equation 1:**

If $W_K^B = 1$ and $\overline{C}(v) \neq 0$ then Ration = $\dfrac{\overline{C}_K(v)}{\overline{C}(v)}$

If $W_K^B = 0$ or $\overline{C}(v) = 0$ then ratios is equal to 0.
Feature value range: $[0,1]$

**Equation 2- The "Length / Width" ratio function in *e*Cognition**

**Length/Width:**
Length to width ratio is calculated by comparing the results of two models and taking the smaller number as the feature value.

**Equation 2:**

$$\gamma(v) = \min \gamma_v^{EV} \; and \; \min \gamma_v^{BB}$$

Where:
$\gamma_v$ : is length/width ration of image object v
$\gamma_v^{EV}$ : ratio length of v of the eigenvalues
$\gamma_v^{BB}$ : ratio length of v of the bounding box

For more details please refer to Definiens, (2007a).

Table 5, Table 6 and Table 7 represent the classes, feature objects and fuzzy membership thresholds that are used during the extraction and classification procedure for Charambe, Manzese and Tandale wards, respectively. The fuzzy thresholds are derived from a series of tests across the study areas.

**Table 5- Classes, object features and their fuzzy thresholds defined to extract and classify objects in Charambe ward**

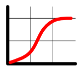
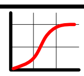
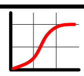
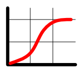
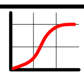
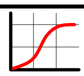
| ***Ward*** | ***Class*** | | ***Object feature*** | ***Membership function*** | ***Thresholds*** |
|---|---|---|---|---|---|
| ***Charambe*** | Building roofs | Common roofs | Blue channel ratio |  | 0.305 – 0.33 |
| | | Red roofs | Red channel ratio | | 0.381 – 0.465 |
| | Vegetated areas | | Green channel ratio |  | 0.35 – 0.645 |
| | Shadows | | Length / Width |  | 4.47 - 26 |

**Table 6- Classes, object features and their fuzzy thresholds defined to extract and classify objects in Manzese ward**

| ***Ward*** | ***Class*** | | ***Object feature*** | ***Membership function*** | ***Thresholds*** |
|---|---|---|---|---|---|
| ***Manzese*** | Building roofs | Common roofs | Blue channel ratio |  | 0.3 – 0.37 |
| | | Red roofs | Red channel ratio | | 0.35 – 0.42 |
| | Vegetated areas | | Green channel ratio |  | 0.356 – 0.478 |
| | Shadows | | Length / Width |  | 3.51 - 22 |

High spectral heterogeneity within the roof and bare ground objects in the study wards caused misclassifications and decreased the level of object separability (see examples in Figure 9). The purpose now is to extract part of the roofs which are ignored or misclassified by the fuzzy membership classifier. Hereafter, “*Nearest Neighbour*” (*NN*) classifier is used to examine whether it is possible to refine the classification results derived from the fuzzy membership classifier.

**Table 7- Classes, object features and their fuzzy thresholds defined to extract and classify objects in Tandale ward**

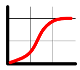
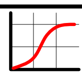
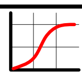
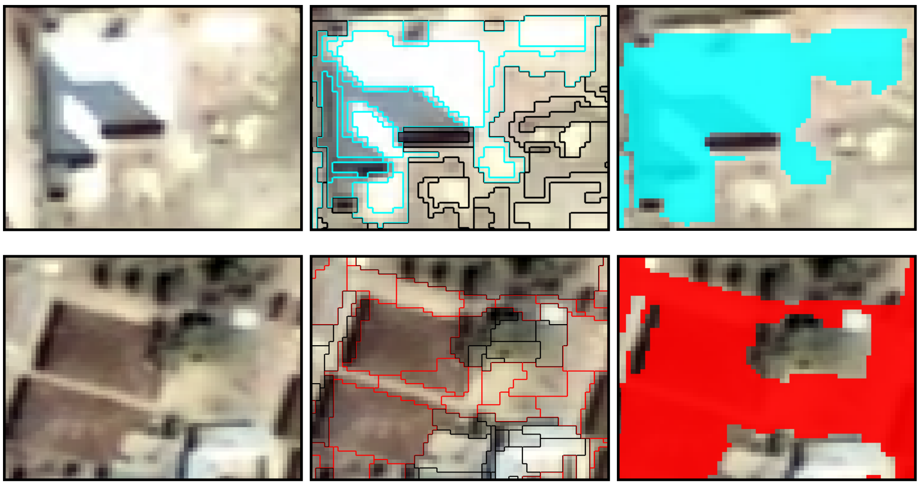
| *Ward* | *Class* | | *Object feature* | *Membership function* | *Thresholds* |
|---|---|---|---|---|---|
| *Tandale* | Building roofs | Common roofs | Blue channel ratio |  | 0.315 – 0.378 |
| | | Red roofs | Red channel ratio | | 0.356 – 0.413 |
| | Vegetated areas | | Green channel ratio |  | 0.356 – 0.412 |
| | Shadows | | Length / Width |  | 5 - 19 |

**Figure 9- Common misclassification examples of roof and bare ground**

**Original view** **Segmentation view** **Classification view**



Upper row: typical settlement of Charambe ward
Lower row: typical settlements of Manzese and Tandale wards

### 2.7.2.1. Classification Refinement Based on *NN* Classifier

Classification with membership functions are based on user-defined thresholds of object features. In contrast, nearest neighbour classification uses a set of samples for different classes in order to assign membership values. After a representative set of sample image objects has been defined for each class, the algorithm searches for the closest sample image object in the feature space for each image object. In fact, this algorithm computes the Euclidean distance from the object to be classified to the nearest training object and assigns it to the class of the training object (Im *et al*., (2008); Definiens, (2007c)). The principle of the *NN* classifier is illustrated in Figure 10.

Based on the image object's feature space distance to its nearest neighbour, a membership value between 0 and 1 is assigned by the classifier. The closer the potential degree of membership to 1 is, the higher the probability of belonging to the same class as the user-defined samples. The membership function diagram created by *NN* classifier is shown in Appendix 4.

The *"Feature Space Optimisation"* (*FSO*) function in *e*Cognition software offers a mathematical method to calculate the best combination of features in the feature space. In order to classify image objects by *NN* classifier, implementing the below steps are indispensable in *FSO* function:

- Selecting the favoured classes that are going to be classified;
- Training the classifier by selecting a small number of samples for each class;
- Determining an initial set of feature objects, corresponding to the purpose;
- Defining the maximum number of features within each combination (*FSO* function will choose the best combination of features from the initial set);
- Calculating the best separation distance between the samples (separation distance value is the minimum overall class combinations, because the overall separation is only as good as the separation of the closest pair of classes); and
- Applying the nearest neighbour classification to the classes based on the optimised feature space.

In light of the above explanations, the *NN* classifier is applied for individual study areas. A sample size of n=10, assumed to be a significant representative of a certain class and feature, is selected for each class. The *NN* classifier classifies a part of the scene correctly even when selecting a small sample size for each class (Definiens, 2007c). Image objects which are classified incorrectly are likely to be at the border between the feature spaces of the classes. By selecting a bigger sample size along the border between the classes, the borders could be defined more accurately. However, increasing the number of sample size in the already well defined areas, will not add information (Definiens, 2007c).

Among the object feature groups (refer to section 2.7.2), *Mean Layer* and *Pixel-Based Ratio* features (8 features in total, see Definiens, (2007a)) belonging to group 2 are inserted into the *FSO* function to see how the *NN* classifier can refine the classification and decrease the number of misclassified objects.

Determining an initial set of feature objects for the *FSO* function is just based on the purpose of the classification and the objects to be classified. A maximum number of three object features (out of eight) are defined in *FSO* function, in order to compute the separation distance. The *FSO* results have shown that the most effective combination of feature for the separation of the classes is: the ratio and mean of red channel and the ratio of blue channel, representing the best separation distance of 1.7, 1.92 and 1.53 for Charambe, Manzese and Tandale, respectively.

It is important to notice that the best selection of sample size and object features for running the *NN* classifier is achieved through an iterative procedure.

Ultimately, the classification algorithm in *e*Cognition software uses the class descriptor (here is the nearest neighbour descriptor) to assign a class. It evaluates the class description and determines on this basis whether an image object can be a member of this class. According to the aim of this study, the objects of interest which have to be extracted and classified are building roofs. By the time that the entire classification procedure is ended, the rest of the objects (e.g., shadow, vegetated area, bare ground) are considered as unclassified objects category. Figure 11 shows how the classification results change by *NN* classifier.

**Figure 10- Principle of the nearest neighbour classification (source: Definiens, (2007c))**

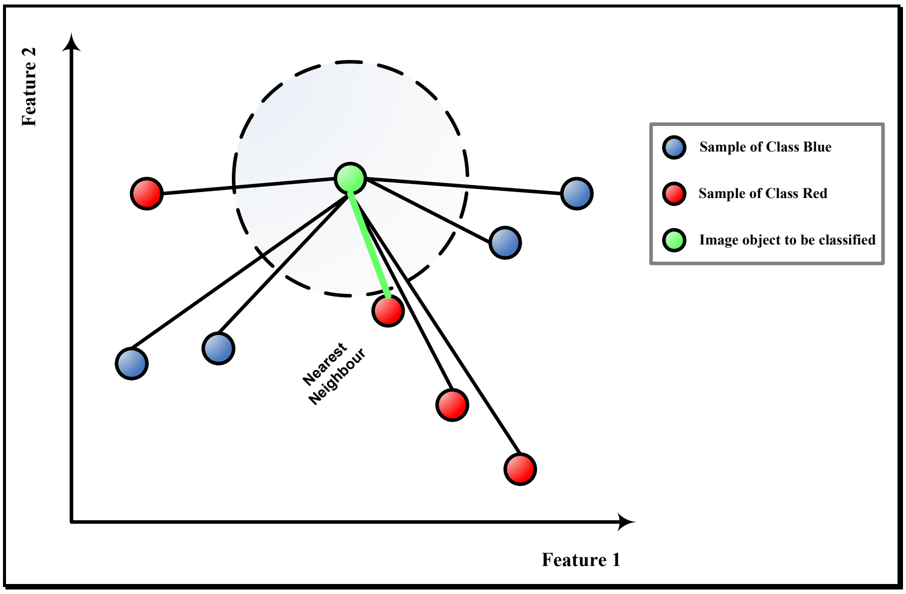
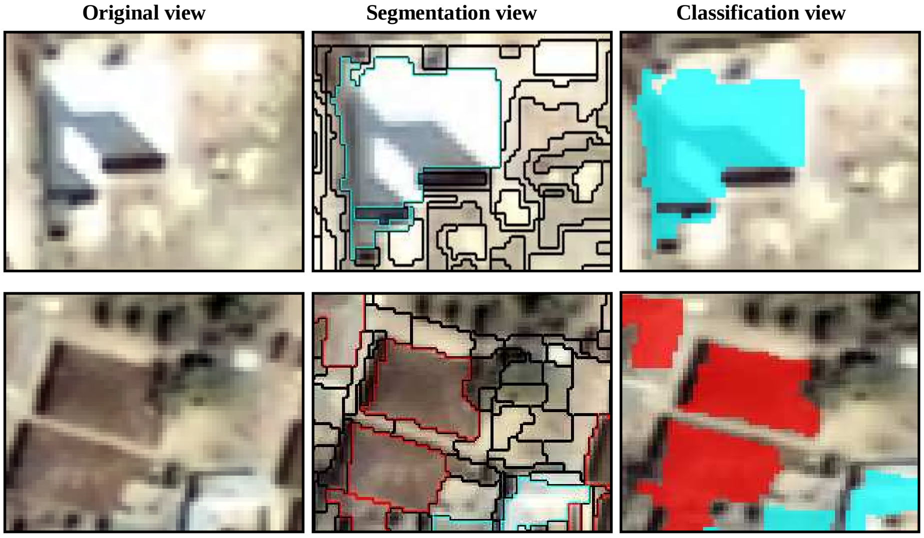


**Figure 11- Classification refinement based on *NN* classifier**




Upper row: typical settlement of Charambe ward
Lower row: typical settlements of Manzese and Tandale wards

### 2.7.3. Export of Classification Results

Derived objects from the sequential classification process have to be exported, so that detailed analysis and further investigations can be done in the downstream GIS tool that uses the data (Robson *et al*., 2006).

However, the first step is to reduce the number of classified segments resulted from the *MRS* method. *"Merge Region"* function is used to merge all the neighbouring image objects of a specific class to one large object without changing the classification. Thereby, the image segments on top of each roof are merged into a single object making the building roofs.

The second step is to export the roof objects as polygons together with their attributes. A large range of feature values may also be exported as the attributes of the roof polygons. Nevertheless, the only feature value that is of great importance for the next phase of the analysis (estimating the number of slum inhabitants) is the "*Area*". Accordingly, all the merged roofs of the three study wards are exported as smoothed polygons with the mentioned object feature.

### 2.7.4. Accuracy Assessment of the Roof Extraction Process

As the application of object-oriented image analysis is creative for the objective of this study, it is important to test and assess the accuracy of the objects being produced. The digitised roof polygons are used as a reference data to evaluate the quality of the roof extraction. Under assumption of being a representative sample, a sample size of n=550 for each study area is determined. The size of the sample is calculated by analysing the sampling error (probability level of p=0.05, allowable error of 10%). All the formulas for calculating the sampling and allowable error is shown in Appendix 5 (more details to be found in Thompson, (2002)).

Using both extracted roofs and reference polygons in vector format, different parameters are assessed and the results are reported (see Chapter 3): the number of detected buildings, the total extracted roof area and the ratio of roof area coverage (intersection area between extracted and reference buildings). Such parameters are important for slum urban monitoring, planning and management.

#### 2.7.4.1. Extraction rate

The number of detected buildings derived from semi-automatic extraction process, plays a significant role in calculating the extraction rate. Mathematically, extraction rate can be expressed as the blow equation (Avrahami *et al*., (2004); Mayunga *et al*., (2007); Stassopoulou *et al*., (2000)):

$$BER = \frac{BCE}{RBP} \times 100 \qquad \text{(Equation 3)}$$

Where *BER* is the "*Building Extraction Rate*", *BCE* is the number of "*Building Correctly Extracted*" and *RBP* is the number of "*Reference Building Polygons*". The only assumption defined for calculating the *BER* is that if a reference building is covered by more than 25% of an extracted building (see Equation 4), then it is considered as a correctly identified building. This arbitrary threshold is fixed to take into account the co-registration errors and the fact that semi-automated extraction process may not delineate entire buildings as manual photo-interpretation does (Durieux *et al*., 2008).

#### 2.7.4.2. Roof Area Coverage

In the roof area coverage metric, the intersected area between reference polygons and extracted buildings is computed in percentage. The "*Roof Area Coverage*" *(RAC)* is characterised as the roof area coverage of the extracted buildings with respect to the reference polygons, as shown by Equation 4:

$$RAC = \frac{EBA}{RPA} \times 100 \qquad \text{(Equation 4)}$$

Where *RAC* is the *"Roof Area Coverage"*, *EBA* is the total *"Extracted Building Areas"* and *RPA* is the total *"Reference Polygon Areas"*.

The accuracy of the results is also assessed irrespective of the assumed coverage threshold (25%) and co-registration error. Overall accuracy, user and producer accuracy and Kappa statistic are computed from the confusion matrix.

## 2.8. Estimation of Slum Population

The statistical modelling approach for estimating the population of slum inhabitants is interested in inferring the relationship between population and other variables such as the urban area, land use and number of dwelling units (see Section 1.1.2). However, the model designed for the purpose of this study is called *"Roof Area" (RA)* which uses a combination of different factors. In *RA* model, population has a direct relation with roof area and also with roof area per person. The constitutive components of this model are shown as a mathematical function in Equation 5:

$$TEP = EBA \times RApP \qquad \text{(Equation 5)}$$

And

$$RApP = \frac{RPA}{TRP} \qquad \text{(Equation 6)}$$

Where in Equation 5, *TEP* is the *"Total Estimated Population"*, *EBA* is the total *"Extracted Building Areas"* and *RApP* is the *"Roof Area per Person"*. In Equation 6, *RPA* is the total *"Reference Polygon Areas"* and *TRP* is the *"Total Reference Population"*.

The roof area of dwelling units is the independent variable, whereas the total population is dependent on the particular study area (Purusottam, 1998).

Two main assumptions for estimation of slum population are as follows:

- All buildings within the study areas are of residential or commercial/ residential use and of one storied height; and
- All buildings within the study areas are occupied.

### 2.8.1. Accuracy Assessment of the Slum Population Results

Accuracy assessment of the estimation of slum population process is of great importance, since it is the main objective of this research. In addition to $R^2$ which is often used to evaluate the performance of a model based on the modelling dataset, *"Relative Error"* (*RE*) is used to assess the model performance based on the validation dataset. The relative or proportional error of population estimation is defined as:

$$RE = \frac{(TEP - TRP)}{TRP} \times 100 \qquad \text{(Equation 7)}$$

Where *TEP* and *TRP* are the *"Total Estimated Population"* and the *"Total Reference Population"*, respectively. Since the error of estimation for individual study areas may be positive or negative, an average of the absolute values of relative error is further calculated (see Equation 8) which is an indicator of overall accuracy:

$$MRE = \frac{\sum_{k=1}^{n} RE_k}{n} \qquad \text{(Equation 8)}$$

Where *MRE* is the *"Mean Relative Error"* and *k* indexes the number of study areas that is *n* (Lu *et al*., (2006); Harvey, (2002)).

The accuracy of the population estimation results is also evaluated regardless of the assumed coverage threshold (25%). User accuracy of the class "Building" plays an important role in the estimation of slum population, in that it has a direct relation with the amount of roof area which is extracted (both correctly and incorrectly) from the applied methodology.

# 3. Results & Discussion

The results of the applied methodology for roof extraction process and estimation of slum population are evinced in this chapter; thereby, the outcomes are statistically analysed and their accuracy are assessed and discussed.

## 3.1. Results of the Roof Extraction Process

The purpose here is to evaluate the capacity of the roof extraction methodology to estimate the slum population. It is decided, therefore, to first concentrate the accuracy assessment on the estimation of the roof extraction quality in different study wards. The resulted maps of the roof extraction process for three sample study areas in comparison with the reference roof polygons are shown in Appendix 6. Visual inspections overlaid with the reference data revealed that the shape and the orientation of the extracted buildings are different and that there are some small positional errors. It is important to mention that the resulted roof objects from our methodology differ from the reference polygons. In that there can be more than one reference building in one extracted roof with consecutive impacts on comparative area (see Figure 12). Furthermore, the reference polygons are formed by straight lines and they delineate the entire buildings. Whereas, our method detects only part of the roofs that match the extraction criteria which is established in *e*Cognition software (see Table 5, 6 and 7) with not very smoothed and orthogonal borders. Thus, only the number of correctly detected buildings, the total roof area and the intersected area between extracted and reference polygons are considered for the accuracy assessment.

Extraction accuracy is computed on three different study wards, 550 references polygons each. Referring to Equation 3 and 4, the number of correctly classified buildings and also the total roof area extracted by the developed method is calculated. Considering the 25% coverage threshold, 1504 buildings are extracted out of 1650 reference buildings in the study areas which correspond to 91.1% accuracy in extraction rate (see Equation 3).

The methodology developed in this research eventuated in satisfactory results; considering all the constraints. Across all three study sample sites, the total roof area coverage (see Equation 4) extracted from *e*Cognition software is estimated at 136720 $m^2$ and from reference data 183195 $m^2$. This amount shows that 25.7% of the reference polygons are not covered by the extracted roofs. This last result must be interpreted with caution, since it could express the fact that the missing parts of the buildings are not extracted because they are covered by trees or by different roof materials even in one dwelling unit. Nevertheless, this error could be counterbalanced by the merged adjacent objects like cars or bright bare ground (Durieux *et al*., 2008). Owing to the fact that each study area has its specific physical characteristics, they must be studied separately. Although it is predicted to have close results for Manzese and Tandale which are two dense slum neighbours with similar physical characteristics. The *BER* and *RAC* for each sample study site are explained in more details in Table 8.

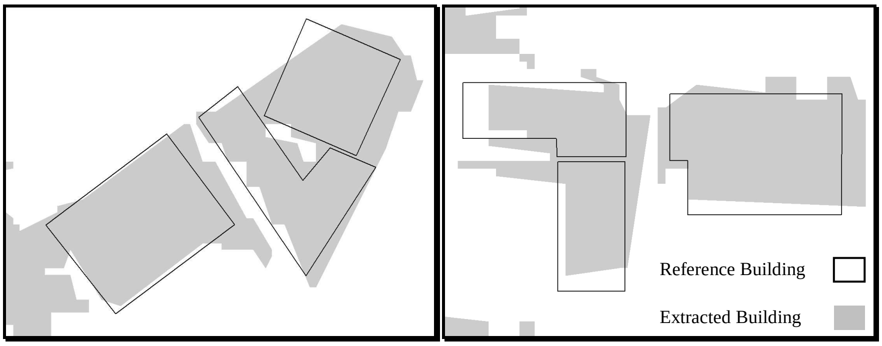
**Table 8- The resulted extraction rate and roof area coverage for the study areas**

| ***Ward*** | ***BCE*** [1] | ***BER*** | ***EBA/RPA (m2)*** | ***RAC*** |
|---|---|---|---|---|
| *Charambe* | 534 | **97.1%** | 53852 / 65518 | **82.2%** |
| *Manzese* | 496 | **90.2%** | 43025 / 59332 | **72.5%** |
| *Tandale* | 474 | **86.2%** | 39843 / 58345 | **68.3%** |
| ***Average*** | | **91.1%** | | **74.3%** |

1- The *BCE* is calculated based on the 25% coverage threshold.

In Tandale study area, a total of 474 roofs (82.6%) are correctly extracted that is highly comparable with the results of Manzese with 496 (90.2%) roofs detected. In addition, the *RAC* represents 68.3% and 72.5% for Tandale and Manzese, respectively. The contribution of *RAC* in the total roof area of these two dense slum wards is not very high. This could be as a result of discrepancies in dense slum areas that different buildings with variety of roof materials are available. However, the 4.2 % difference in the *RAC* for these two neighbour wards can only be explained by the difference in the parameters used for the image segmentation process. In Comparison with dense slum areas, it is anticipated that the quality of the results in Charambe that is composed of individual housing areas would be better (refer to Table 8). In Charambe test site, 97.1% of the reference buildings are identified and extracted correctly which is the highest rate between all the study areas. Besides, the *RAC* with 82.2% indicates that the total area of the extracted roofs is the closest one to the reference data. The frequency of the percentage of intersected area between reference polygons and extracted results for three study areas are illustrated as a graph in Figure 13.

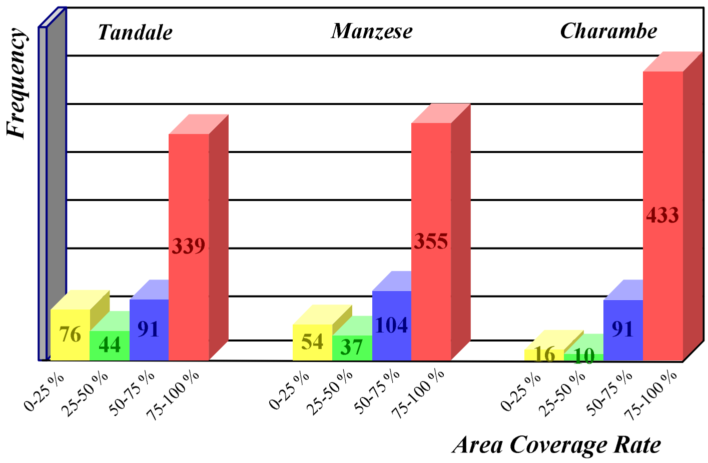
**Figure 12- Examples of extracted and reference polygons, showing the positional error**




Visual interpretation of the results has shown that very big or relatively small buildings do not yield reasonable coverage rates by the applied methodology. The coverage rate, especially for the big buildings has more influence on the overall accuracy, consequently on the forthcoming *TEP* estimation. Hence, In order to see whether the results of the *RAC* change with the size of the reference buildings, a correlation analysis is done for each study site, as it can be seen in Figure 14.

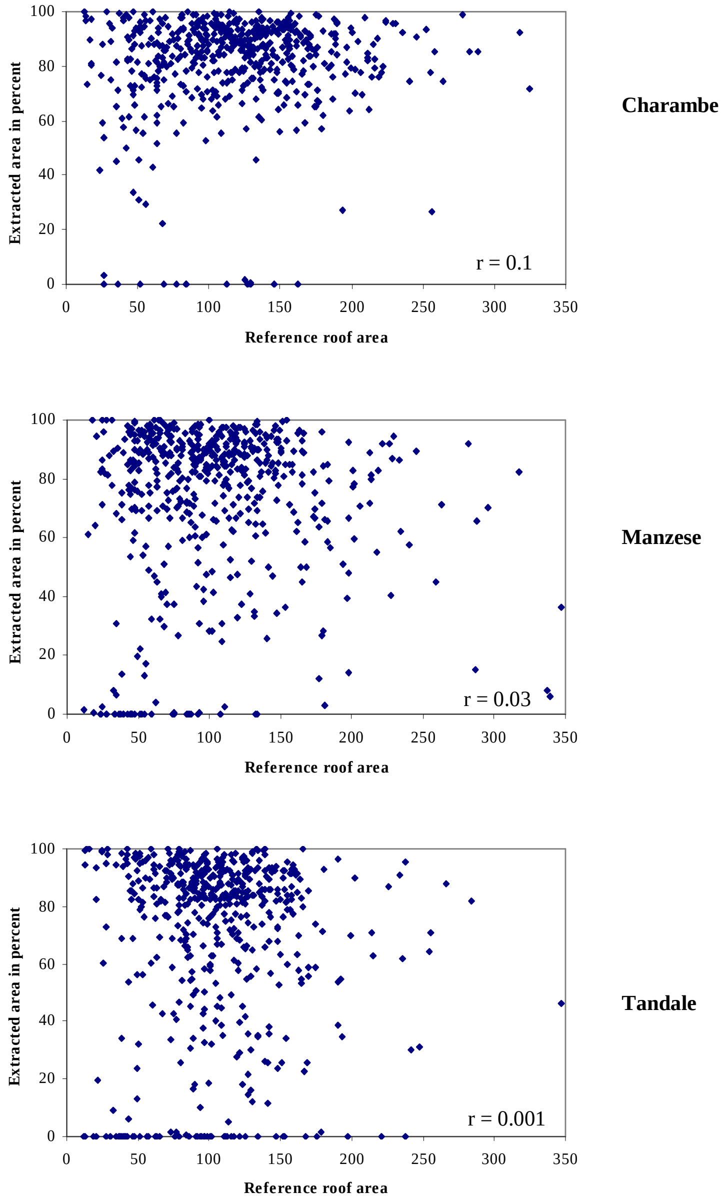
The correlation analysis shows that increase or decrease in the actual roof area of the buildings has no effect on the *RAC*. This theory is corroborated by the correlation value which is very low for all the three study sites. The correlation value of 0.1 for Charambe, 0.03 for Manzese and 0.001 for Tandale, strongly confirms that there is no relation between the actual size of the buildings and the resulted roof area from the object-oriented approach.

**Figure 13- Frequency of the percentage of intersected area between reference and extracted polygons**



Reviewing the evidences demonstrate that most of the slum buildings are detected and extracted satisfying. Problems mainly occurred within slum areas where even a visual inspection of the VHR images could not lead to pleasing detection of houses.

**Figure 14- Correlation analysis of *RAC* and reference roofs areas**



Implementing the coverage threshold (25%) in the model causes the ignorance of Type I error (existence of the building in reality but not detected by the method) and Type II error (no building on the ground but detected by the applied method). In order to modify the accuracy assessment of the roof extraction process and obtain a more realistic approximation, overall accuracy, producer accuracy that corresponds to the error of omission (i.e., Type II error) and user accuracy that corresponds to the error of commission (i.e., Type I error) are calculated by using the confusion matrix (see Appendix 7 for the calculations). The results are reported in Table 9. Moreover, the Kappa statistic is computed from the error matrices. While the overall accuracy assesses the quality of the map and the number of correctly classified objects, Kappa statistic evaluates the quality of the method used to produce the map (Congalton and Green, 1999).

The average of overall accuracy for the building class in the study sites is satisfactory at 93% (Table 9). The user accuracy is ranged from 79% (Manzese) to 83% (Charambe) and the producer accuracy is ranged from 90% (Tandale) to 98% (Charambe).

**Table 9- Accuracy assessment of the roof extraction process for the study areas**

| ***Ward*** | ***Overall Accuracy*** | ***User Accuracy*** | ***Producer Accuracy*** | ***Kappa*** |
|---|---|---|---|---|
| *Charambe* | 94% | 83% | 98% | 0.85 |
| *Manzese* | 93% | 79% | 96% | 0.81 |
| *Tandale* | 92% | 82% | 90% | 0.80 |
| ***Average*** | **93%** | **81%** | **95%** | **0.82** |

The average Kappa value of 0.82 indicates that the object-oriented method used for the extraction of the roofs produced quite good results. The Kappa values representing 0.85, 0.81 and 0.80 for Charambe, Manzese and Tandale, respectively. The reported results of *RAC* (refer to Table 8) are based on the ignorance of omission and commission errors. However, the user accuracy (building class) at 83% for Charambe and 79% for Manzese are very close to that of *RAC* (refer to Table 8) for these two wards. For Tandale ward, in contrast, there is a big difference in the user accuracy and *RAC* representing 82% and 68.3%, respectively. Although the results vary by considering and ignoring the 25% coverage rate, the *BER* which is an indication of extraction accuracy cannot be calculated by ignoring the coverage threshold.

### 3.2. Results of Slum Population Estimation

There are two main purposes that are going to be pursued in this section: to evaluate how close to reality the total *EBA* derived from the object-oriented method can estimate the number of slum inhabitants and to find out what is the relation between roof area and slum population.

By using the provided dataset about the number of occupants in the squatters of Dar-Es-Salaam (see section 2.2) and also the total roof area of each sample study site, the calculation of occupancy rate is feasible. The term *RApP* (*Roof Area per Person*) which is used in the mathematical formula is an equivalent term of the occupancy rate. The population estimation results for each study site and all the fundamental components are summarised in Table 10.

**Table 10- Total estimated slum population for each test area**

| ***Ward*** | ***RPA ($m^2$)*** | ***TRP*** | ***RApP (occupancy rate)*** | ***EBA ($m^2$)*** | *TEP* | *Percentage of TRP* |
|---|---|---|---|---|---|---|
| *Charambe* | 65518 | 3189 | **20.54** | 53852 | **2622** | **82.2%** |
| *Manzese* | 59332 | 5813 | **10.20** | 43025 | **4218** | **72.5%** |
| *Tandale* | 58345 | 5614 | **10.39** | 39843 | **3834** | **68.3%** |
| ***Average*** | | | | | | **74.3%** |

The occupancy rates for these three wards are showing that there is more living space in Charambe by 20.54 $m^2$, in Manzese by 10.2 $m^2$ and in Tandale by 10.39 $m^2$ for each individual. As it can be seen from Table 10, the occupancy rate drastically declined from 20 $m^2$ to 10 $m^2$, approximately. However, this difference is understandable since the number of inhabitants for either Manzese or Tandale is double that of Charambe.

The overall *TEP* calculated for the three test sites shows that 74.3% of the population is estimated by the *RA* model. Nevertheless, the estimation of population varies in different wards representing 82.2%, 72.5% and 68.3% for Charambe, Manzese and Tandale, respectively. Statistically, the model performance is assessed using the *Relative Error* (*RE*) function (see Equation 7). The *RE* results ranged from -17.8% for charambe, -27.5% for Manzese and -31.7% for Tandale. The negative values indicate that the population for all the three study sites are under-estimated and below the reference population data. Since the error of estimation for individual study areas may be positive or negative, an average of the absolute values of relative error (*MRE*) would be an indicator of overall accuracy which is calculated and stated in Table 11.

**Table 11- *Relative Error* and *Mean Relative Error* of population estimation for each study site**

| ***Ward*** | ***TRP*** | *TEP* | ***RE*** |
|---|---|---|---|
| *Charambe* | 3189 | 2622 | **-17.8%** |
| *Manzese* | 5813 | 4218 | **-27.5%** |
| *Tandale* | 5614 | 3834 | **-31.7%** |
| ***MRE*** | | | **25.7%** |

The derived *MRE* value determines that the estimation of population in general, is 25.7% less than the real number of slum inhabitants. Referring to Table 10, the overall *TEP* (74.3%) conforms to the *MRE* value (25.7%). Both values can be used to describe the overall accuracy and the error (in this case under-estimation) of the slum population based on the *RA* model.
In order to avoid the effect of ignoring Type I and Type II errors in the estimation of slum population, the number of slum inhabitants is approximated without considering the coverage threshold (refer to section 2.8.1). The results are showed in Table 12.

**Table 12- Total estimated slum population without considering the coverage threshold**

| ***Ward*** | ***RPA ($m^2$)*** | ***TRP*** | ***RApP (occupancy rate)*** | ***CIC [1] ($m^2$)*** | *TEP* | *Percentage of TRP* |
|---|---|---|---|---|---|---|
| *Charambe* | 65518 | 3189 | **20.54** | 65285 | **3178** | **99.6%** |
| *Manzese* | 59332 | 5813 | **10.20** | 54645 | **5357** | **92.1%** |
| *Tandale* | 58345 | 5614 | **10.39** | 48723 | **4689** | **83.5%** |
| ***Average*** | | | | | | **91.7%** |

1- *CIC* stands for *"Correctly and InCorrectly"* extracted roof areas by the applied methodology. This number is used in the confusion matrix for the calculation of user accuracy.

Even with the calculation of the incorrectly extracted roof areas in the model, the population is not over-estimated. From the total population (14616 inhabitants) in the study areas, 91.7% are estimated by the *RA* model to be living in the region. The ignorance and consideration of coverage threshold is resulted in 17.4% difference in the average percentage of *TRP* reported in Table 10 and Table 12.

The *RE* is calculated for Charambe at -0.4%, for Manzese at -7.9% and for Tandale at -16.5%. Since the error of estimation for individual study areas may be positive or negative, an average of the absolute values of relative error is calculated (refer to section 2.8.1). The *MRE* value (8.3%) shows that the estimation of slum population in the study areas is quite close to the real number of dwellers.

### 3.2.1. Regression analysis

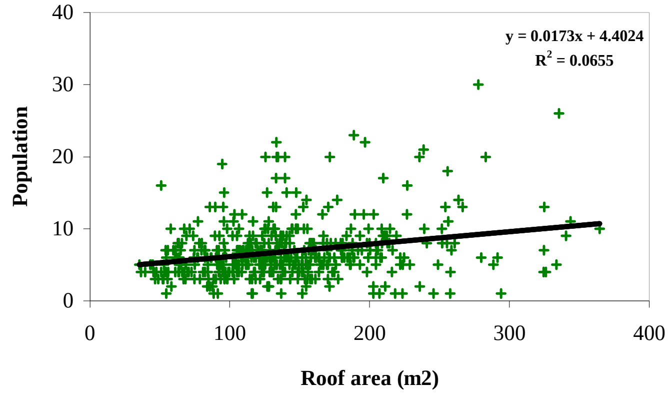
The relation between two variables roof area and slum population is assessed based on the regression analysis. Figure 15 shows the relation between the two mentioned variables and the $R^2$ value, separately for each study area. The points in Figure 15 show that with increase of the roof area, the population also increases indicating a positive relation of roof area and number of inhabitants. In spite of the fact that the $R^2$ value is very low for each study site, linearity in relation of these two variables must be tested statistically (see Table 13).

A regression test is run on the slope coefficient of the regression line to see whether the relation between roof area and population does exist, even with very low regression values. The *null hypothesis* now is that the slope coefficient for the regression lines is equal to zero at the confidence level of 95% ($\alpha$=0.05). The null hypothesis then will be rejected if the *P-value* is less than $\alpha$=0.05. Table 13 shows that all the slope coefficients for the three regression lines are highly significant with a *P-value* less than $\alpha$=0.05. This indicates that all the *null hypotheses* (no linear relation between the variables) are rejected.

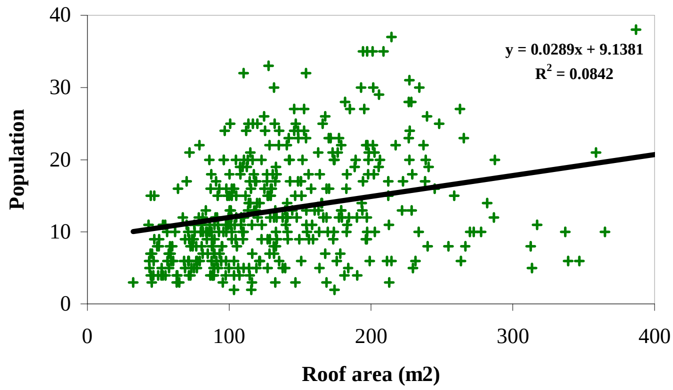
**Table 13- Regression analysis for testing the linearity in relation between roof area and population**

| ***Ward*** | $R^2$ | *Slope coefficient* | ***P-value*** | ***Confidence level ($\alpha$)*** | ***Test $H_0$*** |
|---|---|---|---|---|---|
| *Charambe* | 0.065 | 0.017 | 2.00533E-08 | 0.05 | **Rejected** |
| *Manzese* | 0.084 | 0.028 | 2.22891E-09 | 0.05 | **Rejected** |
| *Tandale* | 0.081 | 0.036 | 1.81607E-10 | 0.05 | **Rejected** |

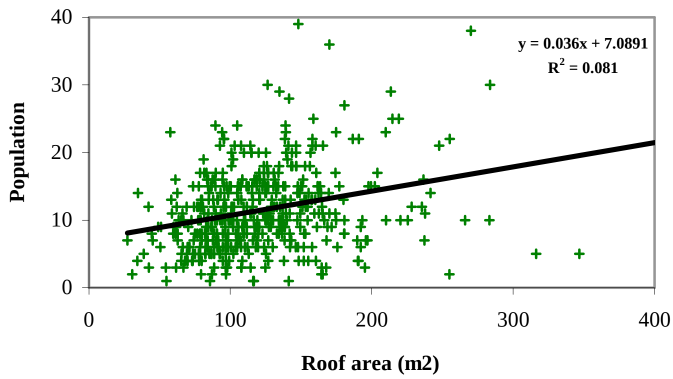
**Figure 15- Relation between roof area and slum population for the three study sites**




Charambe




Manzese




Tandale

# 4. Conclusions & Recommendations

An overall review of the results and further the possible applicability of the implemented method for future studies are stated in the recommendation part.

## 4.1. Conclusions

This work has explored an object-oriented method to count the slum buildings from which the number of inhabitants can be estimated. The results presented in this study proved to be satisfactory both for the roof extraction and estimation of slum population. Simplicity is an advantage of the methodology outlined in this study. The proposed approach is simple, effective and can be applied by researchers and non-professional users. The objectives of this study are addressed. Conclusions to specific research questions are summarised as follows:

***Research Question 1***

The comparison of the above (refer to section 3.1) results have revealed that the roof extraction quality in low density slum areas differs from that of high density slum areas (rejection of the *null hypothesis*). It can be concluded that the rule-set used in *e*Cognition for the extraction of buildings in one ward is not adaptable to the investigation of other study areas.

***Research Question 2***

A conclusion can be made that the *RApP* (occupancy rate) changes with respect to its location (refer to Table 10). Considering the equal size of the study area for all of them, the ratio showed that there is more living space in Charambe compared to Manzese and Tandale (rejection of the *null hypothesis*). However, the difference in the amount of living space is understandable since the number of inhabitants for either Manzese or Tandale is double that of Charambe.

***Research Question 3***

In Figure 15 roof area was plotted against the population. Although very slightly, the graphs showed that with the increase of the size of the roof area, the number of inhabitants increases. Thus, in general it can be summarised that in slum areas population increases proportionally with the increase of roof area of buildings which rejects the *null hypothesis*.

### 4.2. Recommendations

- Finding a well suitable method for the building extraction process is highly depending on the user-defined scale parameter, shape and colour criteria. Therefore, it would be worthwhile to examine the effect of assigning different factors on the classification accuracy results; and

- Future works should investigate the incorporation of multispectral satellite images (at least with the inclusion of NIR band) and ancillary data (e.g. DSM) in object-based methods to support the identification and extraction of slum buildings.

## 5. References


Akbari, H., Shea Rose, L. & Taha, H. (2003) Analyzing the land cover of an urban environment using high-resolution orthophotos. *Landscape and Urban Planning,* 63**,** 1-14.

Amer, S., Ottens, H. F. L. P. & De Jong, T. P. (2007) Towards spatial justice in urban health services planning : a spatial - analytic GIS - based approach using Dar es Salaam, Tanzania as a case study. *ITC Dissertation;140.* Utrecht, Utrecht University.

Avrahami, Y., Raizman, Y. & Doytsher, Y. (2004) Semiautomatic 3D of building from medium scale (1:40000) aerial photographs. *International Society for Photogrammetry and Remote Sensing Commission III* Istanbul, Turkey.

Baatz, M. & Schape, A. (2000) Multiresolution segmentation-an optimization approach for high quality multi-scale image segmentation. IN Strobl, J., Blaschke, T. & Griesebner, G. (Eds.) *Angewandte Geographische Informations-Verarbeitung XII.* Karlsruhe, Germany.

Baltsavias, E. & Mason, S. (2001) Image-based reconstruction of informal settlements. *Automatic Extarction of Man Made Objects from Aerial and Space Images II.* Berlin, Germany.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I. & Heynen, M. (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing,* 58**,** 239-258.

Bjorgo, E. (2000) Using very high spatial resolution multispectral satellite sensor imagery to monitor refugee camps. *International Journal of Remote Sensing,* 21**,** 611 - 616.

Carleer, A. P. & Wolff, E. (2006) Urban land cover multi-level region-based classification of VHR data by selecting relevant features. *International Journal of Remote Sensing,* 27**,** 1035 - 1051.

Cleve, C., Kelly, M., Kearns, F. R. & Moritz, M. (2008) Classification of the wildland-urban interface: A comparison of pixel- and object-based classifications using high-resolution aerial photography. *Computers, Environment and Urban Systems,* In Press, Corrected Proof.

Congalton, R. G. & Green, K. (1999) *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices,* New York, USA, Lewis Publishers.

De Gier., A. (2003) Sampling Basics. *Handout notes.* Enschede, ITC.

Definiens (2007a) Definiens Imaging Developer 7. *eCognition Reference Book*. Definiens AG.

Definiens (2007b) Definiens Imaging Developer 7. *eCognition Software*. *EII Earth.* Munich, Germany.

Definiens (2007c) Definiens Imaging Developer 7. *eCognition User Guide*. Definiens AG.

Durieux, L., Lagabrielle, E. & Nelson, A. (2008) A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. *ISPRS Journal of Photogrammetry and Remote Sensing,* 63**,** 399-408.

Giada, S., De Groeve, T., Ehrlich, D. & Soille, P. (2003) Information extraction from very high resolution satellite imagery over Lukole refugee camp, Tanzania. *International Journal of Remote Sensing,* 24**,** 4251 - 4266.

Harvey, J. T. (2002) Estimating census district populations from satellite imagery: some approaches and limitations. *International Journal of Remote Sensing,* 23**,** 2071 - 2095.

Hofmann, P. (2001a) 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 / Fernerkundung in urbanen Räumen.* Regensburg, Germany.

Hofmann, P. (2001b) Detecting urban features from IKONOS data using an object-oriented approach. *First Annual Conference of the Remote Sensing & Photogrammetry Society*. Munich, Germany.

Hofmann, P., Strobl, J., Blaschke, T. & Kux, H. (2008) Detecting informal settlements from QuickBird data in Rio de Janeiro using an object based approach. *Object-Based Image Analysis.*

Iisaka, J. & Hegedus, E. (1982) Population estimation from Landsat imagery. *Remote Sensing of Environment,* 12**,** 259-272.

Im, J., Jensen, J. R. & Tullis, J. A. (2008) Object-based change detection using correlation image analysis and image segmentation. *International Journal of Remote Sensing,* 29**,** 399 - 423.

Jain, S. (2007) Use of IKONOS satellite data to identify informal settlements in Dehradun, India. *International Journal of Remote Sensing,* 28**,** 3227 - 3233.

Kerle, N. & Leeuw, J. D. (2008) Reviving legacy population maps with object-oriented image processing techniques. *In Press on IEEE Transaction on Geoscience and Remote Sensing*.

Kraus, S. P., Senger, L. W. & Ryerson, J. M. (1974) Estimating population from photographically determined residential land use types. *Remote Sensing of Environment,* 3**,** 35-42.

Kuffer, M. (2001) *Monitoring density changes in informal settlements in Dar es Salaam by remote sensing : exploring the use of SPOT, ERS and small format aerial photography,* Enschede, ITC.

Kyessi, A. G. (2002) *Community participation in urban infrastructure provision: servicing informal settlements in Dar es Salaam, Tanzania,* Dortmund, SPRING.

Laliberte, A. S., Rango, A., Havstad, K. M., Paris, J. F., Beck, R. F., Mcneely, R. & Gonzalez, A. L. (2004) Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. *Remote Sensing of Environment,* 93**,** 198-210.

Lemma, T., Sliuzas, R. V. & Kuffer, M. (2006) Participatory approach to monitoring slum conditions : an example from Ethiopia. *In: Participatory learning and action (2006)54. Special issue on: Mapping for change : practice, technologies and communication : proceedings of the international conference on participatory spatial information management and communication PGIS '05 - KCCT,* *Nairobi, Kenya, 7-10 Sept 2005. / ed. by Giacomo Rambaldi; Jon Corbett; Mike McCall; ...[et al.] London : IEED, 2005. pp. 54-58*.

Lo, C. (2008) Population estimation using geographically weighted regression. *GIScience & Remote Sensing,* 45**,** 131-148.

Lo, C. P. (1986) *Applied remote sensing,* Harlow, Longman Scientific & Technical.

Lo, C. P. (1995) Automated population and dwelling unit estimation from high-resolution satellite images: a GIS approach. *International Journal of Remote Sensing,* 16**,** 17 - 34.

Lu, D., Weng, Q. & Li, G. (2006) Residential population estimation using a remote sensing derived impervious surface approach. *International Journal of Remote Sensing,* 27**,** 3553 - 3570.

Martínez, J., Mboup, G., Sliuzas, R. & Stein, A. (2008) Trends in urban and slum indicators across developing world cities, 1990-2003. *Habitat International,* 32**,** 86-108.

Mathieu, R., Aryal, J. & Chong, A. K. (2007a) Object-based classification of IKONOS imagery for mapping large-scale vegetation communities in urban areas. *Sensors,* 7**,** 2860-2880.

Mathieu, R., Freeman, C. & Aryal, J. (2007b) Mapping private gardens in urban areas using object-oriented techniques and very high-resolution satellite imagery. *Landscape and Urban Planning,* 81**,** 179-192.

Mavrantza, O. D. & Argialas, D. P. (2007) Identification of urban features using object-oriented image analysis. *Photogrammetric Image Analysis.* Munich, Germany.

Mayunga, S. D., Coleman, D. J. & Zhang, Y. (2007) A semi-automated approach for extracting buildings from QuickBird imagery applied to informal settlement mapping. *International Journal of Remote Sensing,* 28**,** 2343 - 2357.

Meinel, G. & Neubert, M. (2004) A comparison of segmentation programs for high resolution remote sensing data. *Geo-Imagery Bridging Continents XXth ISPRS Congress.* Istanbul, Turkey.

Nobrega, R. A. A., O'hara, C. G. & Quintanilha, J. A. (2005) Detecting roads in informal settlements surrounding Sao Paulo city by using object-based classification. *OBAI'06 - 1st International Conference on Object-Based Image Analysis.* Slazburg, Austria.

Pagot, E., Pesaresi, M., Buda, D. & Ehrlich, D. (2008) Development of an object-oriented classification model using very high resolution satellite imagery for monitoring diamond mining activity. *International Journal of Remote Sensing,* 29**,** 499 - 512.

Polle, V. F. L. (1996) feasibility of embedding of remote sensing and GIS for improved urban management. *In: Geoinformatics '96, : international symposium on the occasion of the 40th anniversary of Wuhan technical University of Surveying and Mapping WTUSM: Vol. 2 / by Daosheng Du [et al...]***,** 131-139.

Purusottam, D. (1998) Estimating population of unplanned residential areas: in Dar es Salaam. Enschede, ITC.

Ridd, M. K. (1995) Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for citiesâ€ *International Journal of Remote Sensing,* 16**,** 2165 - 2185.

Robson, M., Vachon, P. W. & Seeker, J. (2006) Evaluation of eCognition for Assisted Target Detection and Recognition in SAR Imagery. *2006 IEEE International Symposium on Geoscience and Remote Sensing***,** 145-148.

Sliuzas, R. V., Kerle, N. & Kuffer, M. (2008) Object - oriented mapping of urban poverty and deprivation. *Proceedings of the 4th EARSeL (European Association of Remote Sensing Laboratories) workshop on remote sensing for developing countries in conjunction with GISDECO 8.* Istanbul, Turkey.

Sliuzas, R. V. & Kuffer, M. (2008) Analysing the spatial heterogeneity of poverty using remote sensing : typology of poverty areas using selected RS based indicators. IN Jurgens, C. (Ed.) *Remote sensing : new challenges of high resolution, EARSeL, joint workshop.* Bochum, Germany.

Sliuzas, R. V., Ottens, H. P. & Kreibich, V. P. (2004) Managing informal settlements : a study using geo - information in Dar es Salaam, Tanzania. *ITC Dissertation;112.* Enschede, ITC.

Sohn, G. & Dowman, I. (2007) Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing,* 62**,** 43-63.

Stassopoulou, A., Caelli, T. & Ramirez, R. (2000) Automatic extraction of building statistics from digital orthophotos. *International Journal of Geographical Information Science,* 14**,** 795 - 814.

Stow, D., Lopez, A., Lippitt, C., Hinton, S. & Weeks, J. (2007) Object-based classification of residential land use within Accra, Ghana based on QuickBird satellite data. *International Journal of Remote Sensing,* 28**,** 5167 - 5173.

Su, W., Li, J., Chen, Y., Liu, Z., Zhang, J., Low, T. M., Suppiah, I. & Hashim, S. A. M. (2008) Textural and local spatial statistics for the object-oriented classification of urban areas using high resolution imagery. *International Journal of Remote Sensing,* 29**,** 3105 - 3117.

Thompson, S. K. (2002) *Sampling,* Chichester etc., Wiley & Sons.

Tian, J. & Chen, D. M. (2007) Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition. *International Journal of Remote Sensing,* 28**,** 4625 - 4644.

Turkstra, J. & Raithelhuber, M. (2004) Urban slum monitoring. IN Dangermond, J. (Ed.) *GIS the language of geography : 24th ESRI international user conference proceedings.* San Diego, California.

Un-Habitat (2003a) Guide to monitoring Target 11: Improving the lives of 100 million slum dwellers.

Un-Habitat (2003b) Slums of the World: The face of urban poverty in the new millennium? World-wide Slum Dweller Estimation.

Un-Habitat (2008) The state of the African cities: A framework for addressing urban challenges in Africa.

Urban-Info (2008) United Nations Human Settlements Programme. *UrbanInfo v1.0 v5 Database.* UN-HABITAT.

Van Der Sande, C. J., De Jong, S. M. & De Roo, A. P. J. (2003) A segmentation and classification approach of IKONOS-2 imagery for land cover mapping to assist flood risk and flood damage assessment. *International Journal of Applied Earth Observation and Geoinformation,* 4**,** 217-229.

Ward, P. M. & Peters, P. A. (2007) Self-help housing and informal homesteading in peri-urban America: Settlement identification using digital imagery and GIS. *Habitat International,* 31**,** 205-218.

Watkins, J. F. & Morrow-Jones, H. A. (1985) Small area population estimates using aerial photography. *Photogrammetric Engineering and Remote Sensing* 51**,** 1933-1935.

Weeks, J., Hill, A., Stow, D., Getis, A. & Fugate, D. (2007) Can we spot a neighborhood from the air? Defining neighborhood structure in Accra, Ghana. *GeoJournal,* 69**,** 9-22.

Wei, W., Chen, X. & Ma, A. (2005) Object-oriented Information Extraction and Application in High-resolution Remote Sensing Image. *IGARSS Symposium.* Seoul, Korea.

Willhauck, G., Caliz, J. J., Hoffmann, C., Lingenfelder, I. & Heynen, M. (2005) Object-oriented ship detection from VHR satellite images. *Semana Geomática.* Barcelona, Spain.

Wu, S.-S., Qiu, X. & Wang, L. (2005) Population estimation methods in GIS and Remote Sensing: A Review. *GIScience & Remote Sensing,* 42**,** 80-96.

Yuan, F. & Bauer, M. E. (2006) Mapping impervious surface area using high resolution imagery: A comparison of object-based and per pixel classification. *ASPARS 2006, Annual Conference.* Reno, Nevada.

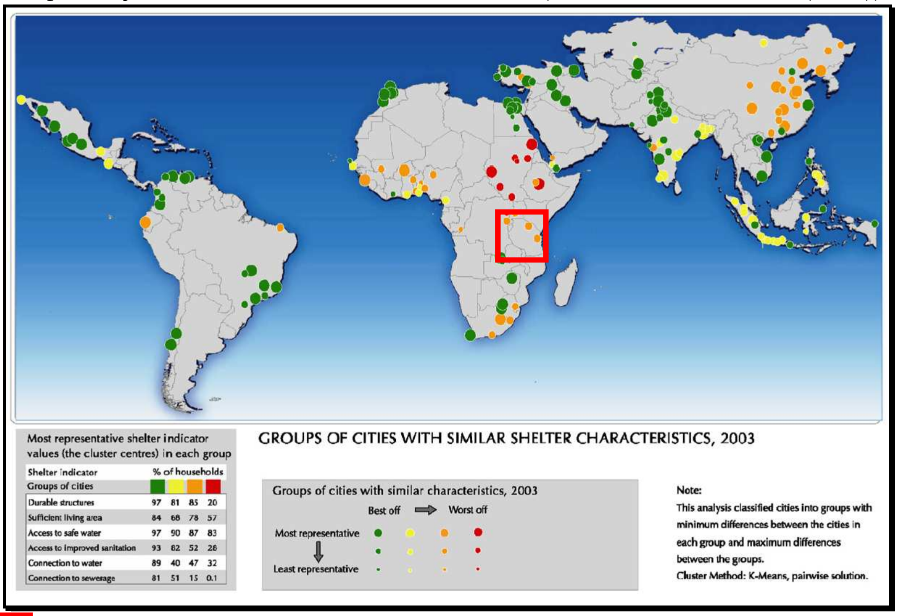
Zhang, Y. & Maxwell, T. (2006) A fuzzy logic approach to supervised segmentation for object-oriented classification. *ASPRS Annual Conference.* Reno, Nevada.

Zhou, W. & Troy, A. (2008) An object-oriented approach for analysing and characterizing urban landscape at the parcel level. *International Journal of Remote Sensing,* 29**,** 3119 - 3135.

## 6. Appendices

### Appendix 1:

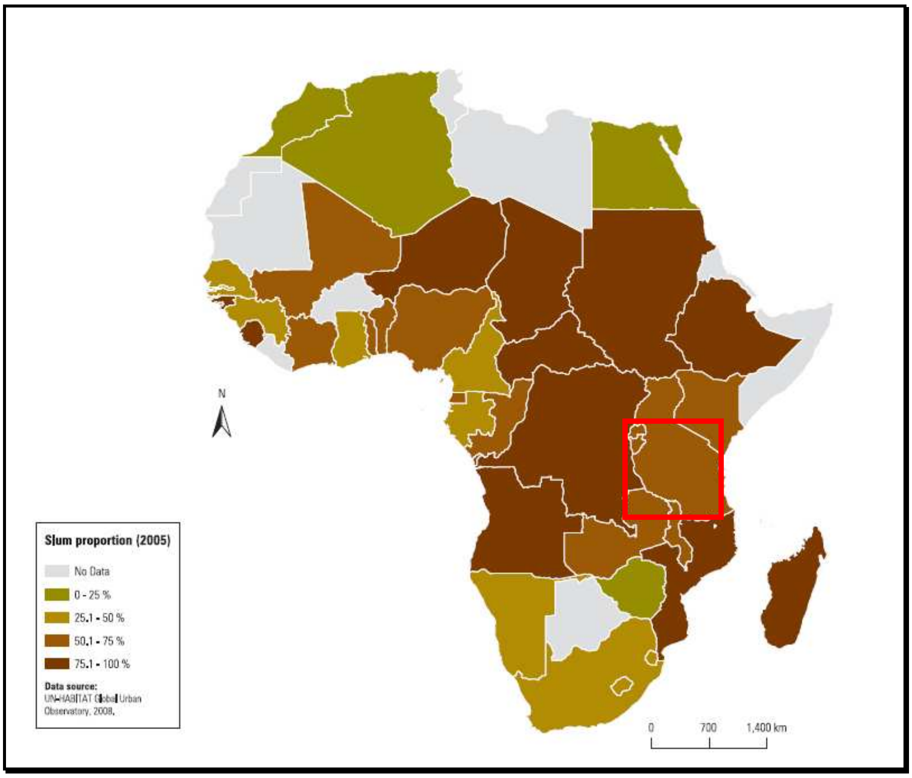
Map of city clusters based on 6 shelter indicators (source: Martínez *et al*., (2008))




Tanzania

**Appendix 2:**

Slum prevalence in Africa (source: UN-Habitat, (2008))




Tanzania

**Appendix 3:**

Image Object Hierarchy

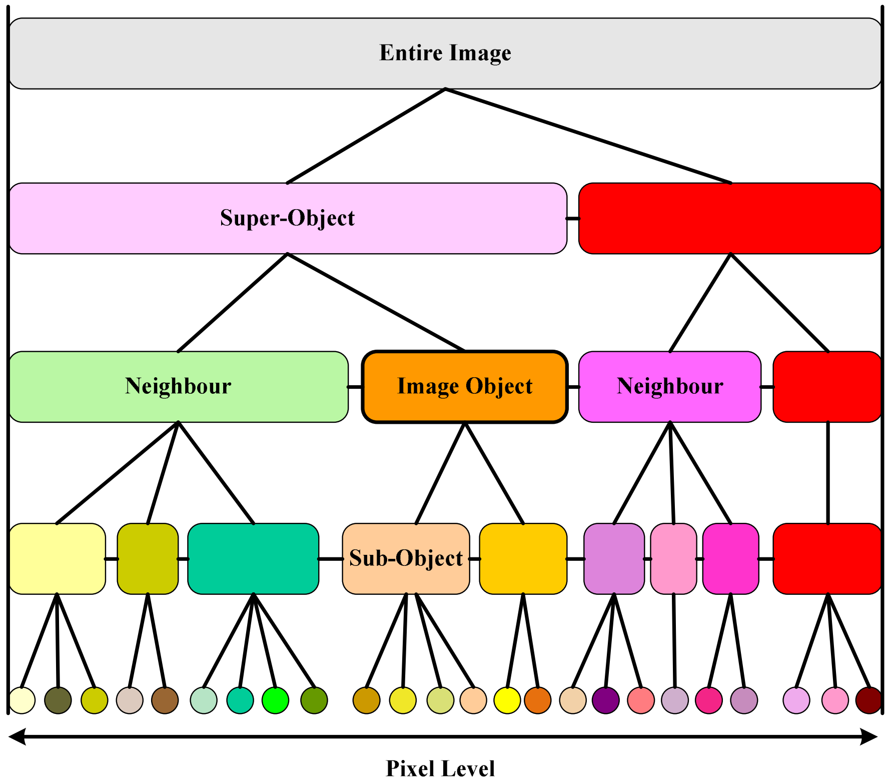
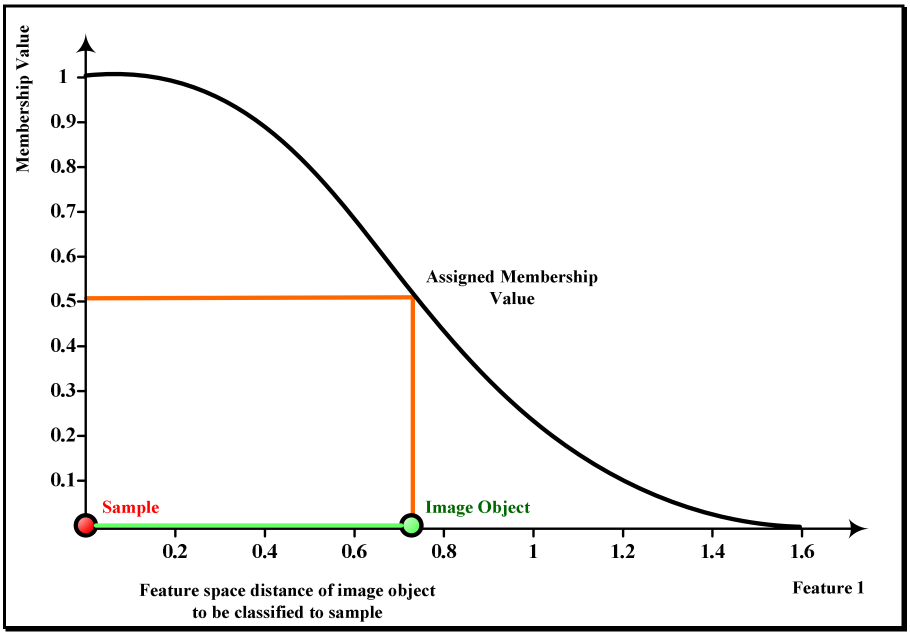



Image objects within an image object level are linked horizontally. Similarly, image objects are linked vertically in the image object hierarchy. The image objects are networked in a manner that each image object knows its context, that are its neighbours, its super-object on a higher image object level and its sub-objects on a lower image object level (Definiens, 2007c).

**Appendix 4:**

Membership function created by *NN* classifier




The Nearest Neighbour classifier returns a membership value between 0 and 1 based on the image object's feature space distance to its nearest neighbour. The membership value is 1 if the image object is identical to a sample. If the image object differs from the sample, the feature space distance has a fuzzy dependency on the feature space distance to the nearest sample of a class. You can select the features to be considered for the feature space (Definiens, 2007c).

**Appendix 5:**

Sampling error = $\dfrac{(t \times s_{\bar{x}})}{\bar{x}} \times 100\%$

Where *t* is the tabular data of the *t*-distribution, $\bar{x}$ is the sample mean and $s_{\bar{x}}$ is the standard error of the mean which can be calculated as follow:

$$s_{\bar{x}} = \frac{s}{n}$$

Where *S* is the standard deviation of the sample and n is the sample size. The required sample size can be computed by the below formula:

$$n_{req} = \frac{t^2 \times CV\%^2}{AE\%^2}$$

Where $n_{req}$ is the required sample size, *AE* is the *"Allowable Error"* of the sampling that is defined by the user and *CV* is the *"Coefficient of Variation"* which is defined as:

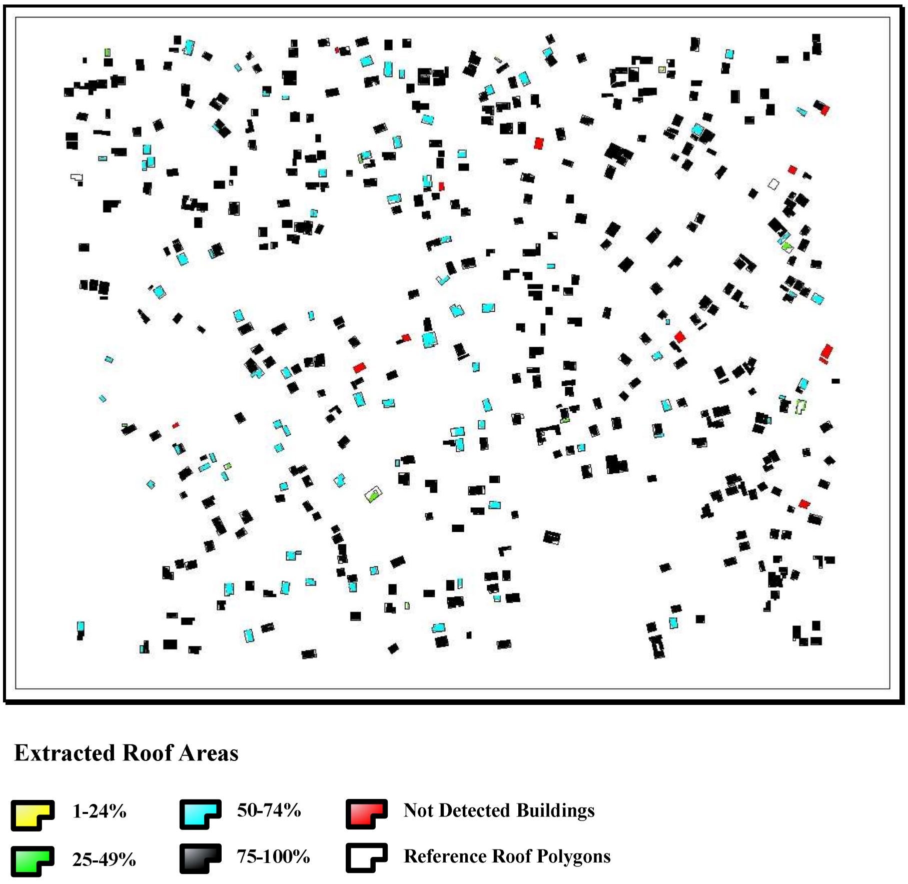
$$CV = \frac{s}{\bar{x}} \times 100$$

Defining the required or adequate sample size that does not exceed the *AE* could be achieved through an iterative process (source: De Gier., (2003); Thompson, (2002)).

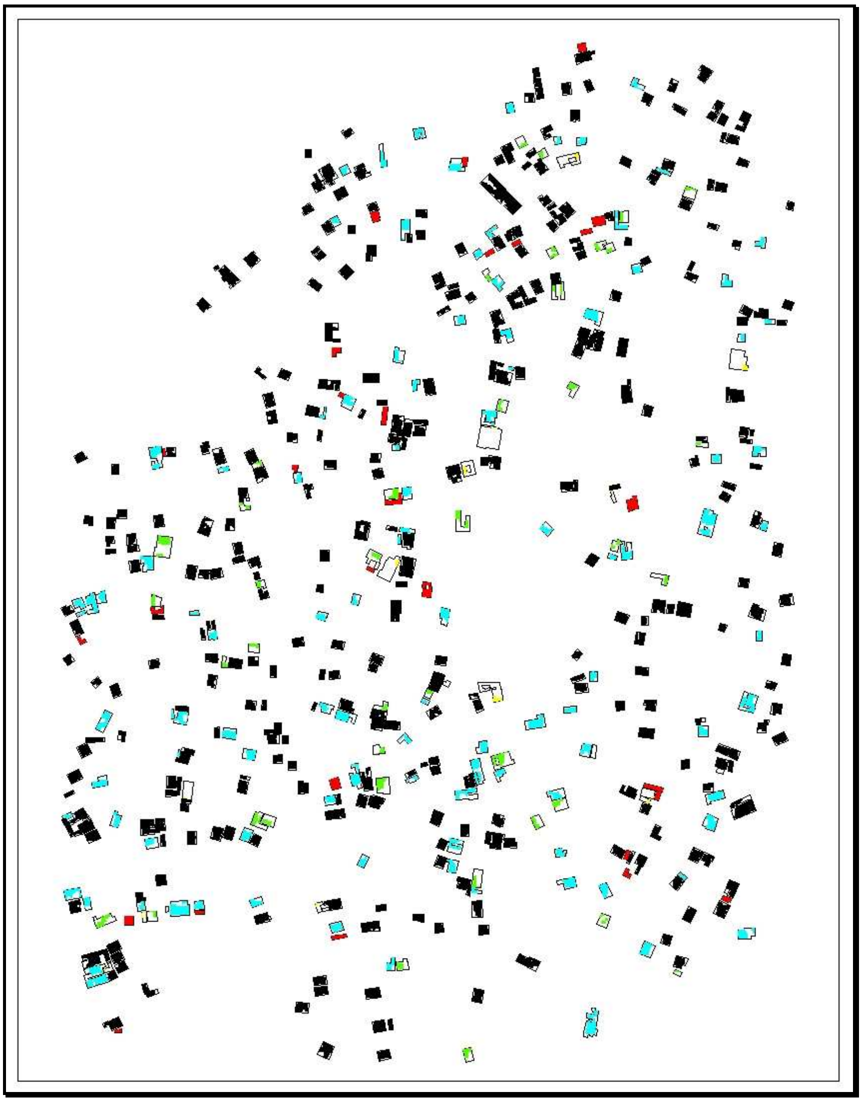
**Appendix 6:**

The extracted roof polygons compared to the reference polygons:

a) Charambe study area

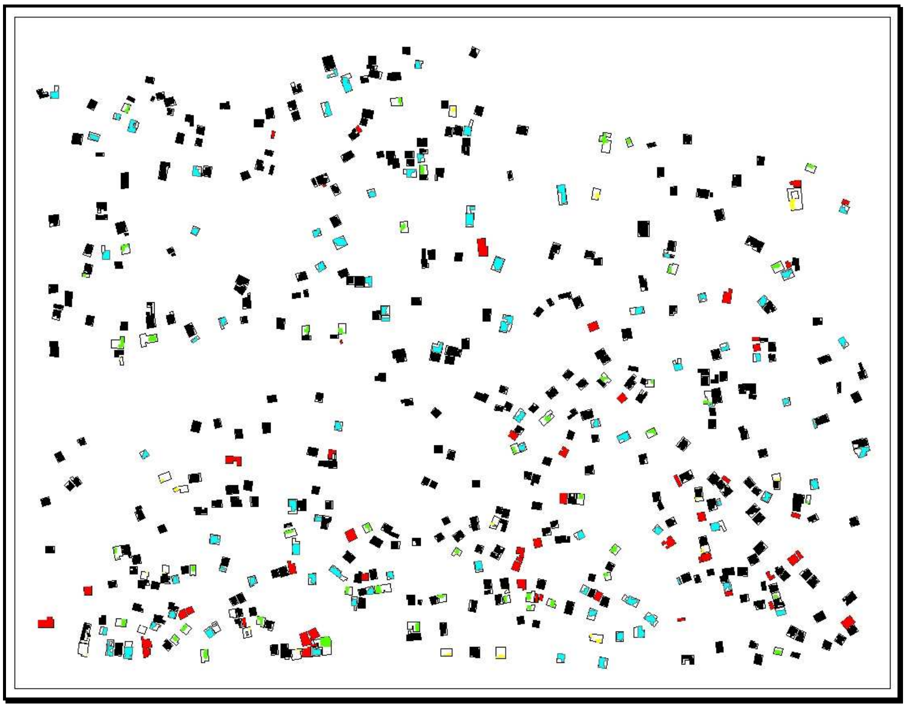


b) Manzese study area:



**Extracted Roof Areas**

c) Tandale study area:



**Extracted Roof Areas**

| | | |
|---|---|---|
| 1-24% | 50-74% | Not Detected Buildings |
| 25-49% | 75-100% | Reference Roof Polygons |

## Appendix 7:

Confusion Matrix for Charambe

| ***Charambe*** | ***Reference Data*** | | | | |
|---|---|---|---|---|---|
| ***eCognition Data*** | | Building | No Building | Row Total | User Acc % |
| | Building | 53852 | 11433 | 65285 | 83 % |
| | No Building | 1104 | 138828 | 139932 | |
| | Column Total | 54956 | 150261 | 205217 | |
| | Producer Acc % | 98% | | | |

Overall Accuracy: 94%, Kappa: 0.85

Confusion Matrix for Manzese

| ***Manzese*** | ***Reference Data*** | | | | |
|---|---|---|---|---|---|
| ***eCognition Data*** | | Building | No Building | Row Total | User Acc % |
| | Building | 43025 | 11620 | 54645 | 79 % |
| | No Building | 1785 | 126268 | 128053 | |
| | Column Total | 44810 | 137888 | 182698 | |
| | Producer Acc % | 96% | | | |

Overall Accuracy: 93%, Kappa: 0.81

Confusion Matrix for Tandale

| ***Tandale*** | ***Reference Data*** | | | | |
|---|---|---|---|---|---|
| ***eCognition Data*** | | Building | No Building | Row Total | User Acc % |
| | Building | 39843 | 8880 | 48723 | 82% |
| | No Building | 4614 | 121778 | 126392 | |
| | Column Total | 44457 | 130658 | 175115 | |
| | Producer Acc % | 90% | | | |

Overall Accuracy: 92%, Kappa: 0.80