Verification of tsunami reconstruction projects by object-oriented building extraction from high resolution satellite imagery

Ye Du

March, 2008

Verification of tsunami reconstruction projects by object-oriented building extraction from high resolution satellite imagery

by

Ye Du

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, Specialisation: Geoinformatics

Thesis Assessment Board

Chairman: Dr. V. A. Tolpekin
External examiner: Dr. S. Zlatanova
Supervisor: Ms. Dr. Ir. W. Bijker
Second supervisor: Prof. Dr. A. Stein

INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
ENSCHEDE, THE NETHERLANDS

Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Abstract

A powerful earthquake (Richter magnitude 9.2) occurred near to the northwest shore of Sumatra, on 26 December, 2004. It triggered a giant tsunami that devastated Banda Aceh, Indonesia. Many donors provided recovery aid, without coordination or proper auditing. This may have led to waste, fraud and corruption. This study investigated an application of remote sensing to enhance financial accountability and transparency in managing reconstruction projects following natural disasters, by automatically identifying buildings constructed as a result of the disaster response, using Banda Aceh as a test area. The increasing availability of high-resolution satellite images, such as the KOMPSAT-2 used in this study, together with aerial orthophotos, makes such a procedure potentially a practical part of a disaster recovery audit.

The segmentation algorithm of eCognition was used to generate image segments. These segments were then classified as “building” and “background” by using a rule-base decision tree based on ancillary information: texture, contextual and semantic properties of objects. Building footprints were extracted to a GIS.

Accuracy was assessed by four methods. First was the traditional approach of generating random points and computing an error matrix. This gave accuracies of 98.6% (user’s) and 63.4% (producer’s) for the “building” class. The second method was based on the overlaying of geometric centres of extracted and manual-identified buildings, with a threshold based on building size. This method gave accuracies at the optimal threshold of 81.0% (correctness) and 84.7% (completeness). The third method applied a bounding box to the extracted and reference data, to take both shape and size into account. The ratio of length to width was defined as the shape condition, and the ratio of areas as the size condition; these were then averaged, giving accuracies of 82.3% (correctness) and 82.5% (completeness). The fourth method combined the second and the third methods, giving the highest accuracy. None of the object-based assessments accounted for “one to many” and “many to one” relationships between extracted and reference data.

New buildings were separated from old by overlaying the extracted footprints with a pre-reconstruction building map, taking those with common areas less than 50% as new buildings. These are then ready for audit.

Building footprints were successfully extracted from high-resolution images by object-oriented classification. Remaining problems include identification of multi-faceted roofs and connected buildings, and correction for these.

Key words: Building footprint extraction; object-oriented classification; object-based accuracy assessment; tsunami; KOMPSAT-2 imagery

Acknowledgements

First, I would like to give my sincere thanks to Spatial Information and Mapping (SIM-Centre) of Bureau reconstruction and rehabilitation NAD (BRR) who provided data for my research. This data was collected in the framework of the Indonesian Supreme Audit Institute (BPK) / International organization of Supreme Audit Institutions (IntoSAI) / Strengthening Accountability for the Audit of Disaster Related Aid (SAADRA) program.

I would like to give my sincere thanks to Mr. Egbert Jongsma, of the Court of Audit of the Netherlands and project manager of IntoSAI tsunami task force, for initiating this research.

I would like to give my sincere thanks to my supervisors Ms. Dr. Ir. W. (Wietske) Bijker and Mr. Prof. Dr. Ir. A. (Alfred) Stein. I feel very happy to work with Ms. Dr. Bijker. She is so responsible and nice to her students. She gave me much help and supported me during my research period. Thanks to my second supervisor Mr. Prof. Stein. His advices and comments were always constructive and lead me to think more to solve problems from different points of view.

Now I would like to give my deepest thanks to my parents. Thank you my dad and mom for giving me this chance to study here. Without your support and trust, I would not have this wonderful experience in my life. When I lost myself sometimes during this one and half years, you have always encouraged me and that let me never give up. I will never forget about what you gave to me. Thank you very much!

Thanks to my friend Zhang Qiuju. I am so happy for sharing this wonderful time with you – in the same class, had trips together and in the same cluster during the last six months and helped each other. I really learned something from you. Thank you!

Thanks to my friend Juan Pablo Ardila Lopez. You are one of my best friends in ITC. You have never stopped giving me help and advice and sharing your knowledge with me since at the beginning. You taught me a lot.

Thanks to all 2006 GFM.2 students. It is a big honour to meet you in ITC and study together.

Thanks to all Chinese in ITC. Thanks for all help from you.

Table of contents

List of figures

List of tables

1 Introduction

1.1 Background

A powerful earthquake with magnitude 9.2 on the Richter scale occurred in the morning of local time at the northwest part of Sumatra, Indonesia on 26 December, 2004. The epicenter of the earthquake was on the Australia-Asia tectonic plate around 150 km south of Meulaboh and 250 km from Banda Aceh [1]. It triggered a giant Indian Ocean tsunami which affected the coastal regions of Indonesia, India, Thailand, Malaysia, Bangladesh, Sri Lanka, and even some east Africa countries. Tsunami caused extensive damage to buildings, roads and facilities. The number of death including missing is approximately 280000 according to United Nation and governments [2]. The tsunami had two international features [1]. First, it was the first global natural disaster, covering countries in two continents, Asia and Africa, on the edge of the Indian Ocean. Second, the response to the disaster had also been global as private sector, non-governmental organizations and international institutions provided helping immediately.

In Indonesia, 20 minutes after the earthquake, the tsunami crashed into the northwest coast of Sumatra and northern part of Aceh province. The worst affected areas were Banda Aceh, Meulaboh and Calang and other towns and villages along the coast of Aceh province. The tsunami destroyed houses, factories, roads, bridges, telecommunications, water systems, electricity networks, forests and agriculture areas [1], and also the databases of inhabitants and cadastre.

The response to the Indian Ocean Tsunami has been a world wide activity, the tsunami affected countries received donations such as cash, food and goods, and technical assistance from countries, official and international organizations around the world. Post-disaster reconstruction planning was started as soon as possible, and houses reconstruction seemed to be the central focus of relief projects. At the beginning, numbers of temporary houses were built for emergency as well as constructing permanent houses later on. However, there were lots of issue about whether those temporary houses would waste resources. Reconstruction projects were slowed down because of several limitations such as lack of suitable building land, environmental problems and lack of construction resources. Governments hoped that planning for reconstruction of settlements could be completed by the end of 2006 [3].

1.2 Motivation and problem statement

Within first two months after tsunami, many countries and international organizations provided huge financial and material support to reconstruct affected areas [3]. In Indonesia, an International Multidonor Trust Fund for Aceh and Nias was established to support the rehabilitation and reconstruction of the disaster area in 2005-2009. The aid department was built up and based on different complexities among which the existence of multiple donors and recipients on a national and international level. The donors mixed and split up aid flows with lack of coordination, cooperation and harmonization. These complexities may have led to a waste, fraud and corruption of the aid fund [4]. Therefore, to minimize those problems, Supreme Audit Institutions (SAIs) from tsunami affected countries and major donor countries supposed to enhance financial accountability and transparency in managing funds related to the tsunami reconstruction projects [4]. The main purposes were to ensure that the funds were distributed efficiently, effectively and economically to the projects [1]. The traditional method for auditing is recording related information and check with reality manually; it is time and labour consuming, and not efficient for updating. SAIs want to find better approaches to audit the money assigned to the relief projects; in addition, they want to find out whether this audit method can be applied to other disaster relief projects and where the difficulties are. This research will contribute to solving above problems.

1.3 Research identification

Nowadays, remotely sensed data which are obtained from both airborne and spaceborne sensors provide huge and valuable information of the earth's surface for many applications such as mapping, analysis, monitoring and management. Therefore, an alternative solution which is the focus of this research is the integration of remote sensing and GIS techniques. It is proposed that these techniques can be used for analyzing the ongoing activities, and indicate where the risks for waste, competition, fraud and corruption are highest.

1.3.1 Research objectives

The general objective of this research was to apply remote sensing techniques with high resolution satellite imagery on object-oriented classification for building extraction and change detection to verify tsunami relief projects. The reconstruction projects include buildings, roads, harbours and other infrastructures. Object-oriented classification technique was used to extract building footprint from high resolution image. There are some essential reasons for studying building extraction: houses are the most important objects in the reconstruction projects, people cannot live without houses; meanwhile from high resolution remote sensing image, houses are the most clearly objects can be interpreted. The aim of change detection was to compare the situation of relief projects of pre- and post-reconstruction from remotely sensed data, and then link the detection result with ancillary data using GIS tools to verify projects.

There are several challenges during the research. Firstly, it can be observed from images that because of the variety of building roof materials, using only spectral characteristics for classification are not enough; therefore, spectral characteristics must be combined other information and rules to identify buildings. Secondly, it is a challenge to integrate different types of ancillary data from project and building detection result into a GIS system for project verification.

The aims of this research were:

(1) To detect and extract buildings from high resolution satellite imagery.

(2) To develop rules for object-oriented classification.

(3) To detect new buildings for assisting the verification of tsunami reconstruction projects.

1.3.2 Research questions

To achieve the objectives, several research questions are posed:

(1) What kind of rules could be the optimal choice for classification in this study?

(2) How successful is the extraction of buildings?

(3) What kind of method is the optimal choice to evaluate the building extraction result by using reference data?

(4) How to detect new buildings from the extraction result?

(5) Which level of detail of reconstruction project can be verified?

1.3.3 Innovation aimed at

Integration of remote sensing and GIS techniques were used for building extraction and detection of new buildings. Object-oriented classification and building extraction were based on roof colours. Different approaches were applied for evaluating the building extraction result. And lessons learned from determining the classification rules and improving the method can be used in new areas.

1.4 Thesis structure

The thesis is divided into seven chapters. Chapter 2 is the literature review which contains image segmentation techniques, image classification approaches, the knowledge of object extraction, accuracy assessment methods and related works of building extraction. Chapter 3 is the description of study area and the datasets which were used in the research. Chapter 4 is about data processing methods which were used in the research. Chapter 5 is the data processing and the results. Chapter 6 is the discussion of the results. Then chapter 7 is conclusions and recommendations.

2 Literature review

Remote sensing imagery provides huge amount of data about earth surface for analysis, monitoring and management. With the increasing availability of high resolution satellite image and aero-photographs, the extraction of ground objects has become more and more important for remote sensing and geographic information system (GIS) application. One solution for extraction of objects and change analysis from high resolution images is classification. The traditional pixel based classification approaches didn’t provide satisfactory result because of the heterogeneous spectral reflectance of pixels [5].

In recent years, object-oriented classification method has become a main way for analysis of ground objects [6, 7]. The advantage of this approach is that the digital image is not considered as a grid of pixels but as a group of objects. Using these image objects, one can handle the problem of classification by applying local conditions for classification. Furthermore, the contextual information can be applied through spatial relationships between objects to improve classification result.

Generally speaking, object-oriented image analysis contains two steps: segmentation and classification. Segmentation involves grouping pixels into homogeneous segments. As long as the image objects are generated, the second step is classification of these image objects based on spectral, texture, contextual and semantic information.

2.1 Image segmentation

Image segmentation is described as the process that divides the image into segments. It is a critical process in image analysis because the segmentation result will influence the following image process and analysis. The main aim of image segmentation is to distinguish homogeneous regions within an image and to split the image into regions which are homogeneous in terms of pixel values [8]. There are three segmentation techniques that described by Fu and Mui [9] and most of image segmentation algorithms are based on one of those three techniques. The following parts will review these techniques in more detail.

2.1.1 Thresholding

The thresholding method is a fast and the simplest technique which is commonly used in many image processing [10]. In many applications of image processing, the grey levels of pixels belonging to the objects are different from those belonging to the background. Thresholding becomes a simple and effective way to separate objects from the background. It is assumed that the neighbouring pixels whose value based on grey level, colour, texture and within a certain range belong to the same class [9]. There are several thresholding methods and were categorized into six groups based on the information they are using [10]:

(1) Histogram shape-based methods select threshold according to the shape properties of the histogram, for example, the peaks, valleys and curvatures of the smoothed histogram;

(2) Cluster-based methods search a break point to group the grey level samples into two clusters as background and objects;

(3) Entropy-based methods result in algorithm that use the entropy of the background and object regions, the cross-entropy between the original and binarized image;

(4) Object attribute-based methods which select threshold value based on the similarity between the grey level and the binarized images, such as fuzzy shape similarity, edge coincidence;

(5) Spatial methods, which use higher-order probability distribution and/or correlation between pixels;

(6) Local methods calculate threshold at each pixel. The threshold depends on local statistics such as range, variance, or surface-fitting parameters of the pixel neighbourhood.

The thresholding method gives good segmentation result if the image has only two opposite components. This method is more sensitive to noise than other techniques for example edge-based segmentation. It is based on the assumption that different classes of image segments are represented by different clusters according to their similarity of grey level, texture, etc. The grey level values of features are generally image dependent and it is not clear that how these features should be defined in such a way to generate good segmentation results [9].

2.1.2 Edge-based segmentation

Edge-based segmentation technique is based on the pixel values which change quickly at the boundaries between regions. This technique contains two steps. The first step is to find segment boundaries from image by identifying edge pixels. The second step is to generate image regions which are completely surrounded by edge pixels as image segments. However, the problem of this technique is caused by the image noise. For example, the boundaries are presented in the locations where there is no edge in reality.

2.1.3 Region-based segmentation

Region-based segmentation technique is based on the assumption that neighbouring pixels in the same region have similar features such as grey level or colour value. It generates segments by applying homogeneity properties to the candidate pixels. The main advantage of this method is that it works well with noisy images. It can be categorized into region growing and split and merge [11].

Region growing segmentation starts from a randomly selected seed pixel and grows by adding neighbouring pixels as long as the criteria are satisfied. The process is repeated until the whole image is segmented. This method can give better segmentation result than the methods which were mentioned above, because it can give relative thin edges of regions and the ability of handling noise in the image is very good.

Split and merge segmentation has two steps, splitting and merging. The splitting is to divide the whole image into sub-areas in a quadtree fashion, unless the sub-areas satisfy a certain homogeneity criterion. The merging is the second step and the aim is to merge adjacent regions which are not significantly different.

2.2 Image classification

Image classification techniques have been widely used nowadays for various applications in different fields. It is one of the digital image interpretation techniques, which assigns image pixels into different classes according to certain conditions. These conditions are based on the spectral characteristics of different materials on the Earth's surface. The principle of classification is that each pixel is assigned to a class by comparing feature vectors in the feature space [12]. And the classification result may be influenced by some factors such as selected remotely sensed data, complexity on the ground and classification techniques [13]. According to the operators involved into classification process, classification can be separated into unsupervised classification and supervised classification; according to classification element, it can be divided into pixel-based and object-based classification. Generally speaking, classification procedure may include following steps:

(1) Selection of appropriate image data, concerning sensor type, acquisition date, available spectral bands, spatial and spectral resolution;

(2) In supervised classification, training samples are based on spectral characteristics of pixels and operator's knowledge of processing area; and in unsupervised classification, number of clusters which will be generated as classes are defined;

(3) Selection of classification algorithm depends on the purpose of classification and characteristics of image data; and unsupervised classification splits image into pre-defined clusters based on spectral similarity;

(4) Running classification;

(5) Accuracy assessment of classification results from quality and quantity view by comparing it with ground truth and generating error matrix.

Classification of remotely sensed images with different techniques is often used in land cover and land use analysis in urban area. Generally speaking, there are some methods for improving the accuracy. Using high resolution images for classification; considering of pattern recognition and texture analysis; integration of GIS data and remote sensing images [14]. Using high spatial resolution images is important to get more detail of objects on the ground, and classification is mainly based on spatial and spectral characteristics by combining contextual and spectral information; texture analysis is another important factor in classification procedure when it is difficult to classify images through spectral information; integration of GIS data as knowledge to assist classification is a popular way, and this kind of data include digitized land use map, municipality boundaries and cadastral databases.

2.2.1 Conventional classifier

Conventional pixel-based classification approaches mainly make use of the spectral reflectance values of pixels in which three statistical classifiers are generally used; these are box classifier, minimum distance and maximum likelihood (ML) classifier [15]. Box classifier is the fastest and easiest method that the boundary of class will be defined by the minimum and maximum pixel values, or mean and standard deviation in feature space; the disadvantage is that the overlap between classes cannot be handled. Minimum distance algorithm assigns pixels to the cluster depending on the shortest distance to mean value of those clusters; this doesn’t consider the variability of classes. ML classifier is the most common used in these three methods; it assigns pixels to classes by calculating the probability of those pixels based on statistical approach.

Although conventional pixel-based classification is widely used to extract thematic information from images, limitations still exist. Conventional classifiers are hard classifiers; therefore each pixel is assigned to one class only. It means that if one pixel contains two or more different classes, it will be assigned to the class which covers more in that pixel than other classes. For example, if a pixel has 60% information about vegetation and 40% belongs to bare soil then this pixel will be assigned into vegetation class. It will not give us more detailed information within a pixel.

2.2.2 Object-oriented classifier

Pixel-based classification is affected by some factors such as the complexity of landscape or the high variation of spectral reflectance. It may cause the “noise” in classification result; however, object-based classification can solve this kind of problem better. It analyzes image based on image segments and extracts real world objects from those segments, therefore it makes more sense to analyze specific targets or area on the ground. Object-oriented classification contains two main steps which are image segmentation and classification. There are some strategies to generate objects in segmentation steps; one is integration of vector and raster data that vector data as thematic layer can split image into segments and classification is performed based on these segments [13]. If there is no vector data available, another way is that merging pixels into objects depends on the homogeneity of pixel values within an area followed by classification based on objects. Object-based classification for image analysis is extraction of real world objects based on their properties such as shape and size which cannot be fulfilled through pixel-based approaches. To analyze objects and get better result, some features related to objects could be used and grouped as physical features, topological features and context features. In this approach, the difficulty comes from making meaningful image objects. Because there are no standard rules for image segmentation.

2.2.3 Neural network classifier

Artificial Neural Networks (ANNs) as its name, simulating the workings of the human brain, has been developed since 1980s. The rules are hidden behind networks; therefore it is quite difficult to tell how the network works. The early neural network was single layer and can only solve classification problems which boundaries between classes are straight lines [16]. Later networks developed from single layer perception to mulit-layer to solve more complex problems. Basic multi-layer network contains three layers; they are an input layer, a hidden layer and an output layer; and these layers are connected by weighted links. Input layer transmits pixel values to hidden layer where summation and threshold functions will be performed and transmits values to output layer. The advantages of ANN are that it accepts any kind of input data, it has generalisation capability, and it has tolerance to the noise in the training data. The disadvantages of ANN are that there is no standard to design a network but it depends more on experience; it is time consuming for training data [16].

2.2.4 Fuzzy classifier

Various types of uncertainty can influence information extraction from remotely sensed data. First, there are some factors that affect the procedures of data processing and generation; for instance, the differences of earth surface in the same area depend on the season, weather, atmosphere conditions and sensor position. Spatial resolution of image also has an effect on image analysis process and it may lead to mixture classes in one pixel. In this case, a soft classifier will take this uncertainty into account.

Fuzzy sets theory was created by Zadeh; it is a class of objects with levels of memberships [17]. Fuzzy logic is an approach to quantify uncertain situations. The main idea is to express degree of certain states of “false” or “true” through range from 0 to 1 instead of using two exact values “false” and “true”. Some classes in real world don’t have crisp boundaries and cannot be defined by precise membership. To deal with the limitation of conventional hard classifier, fuzzy classifier as a powerful soft classifier based on fuzzy systems which applies membership functions and giving membership value from 0 to 1 to each pixel. Therefore final fuzzy classification result is decided by the maximum membership value.

2.2.5 Decision tree

Decision tree is a more powerful tool to improve classification process than some traditional methods. It supports decision rules and hierarchies that can deal with nonlinear classification. Decision tree is based on independent variables which can split data at maximum dissimilarity [18]. Therefore, the operators need to have knowledge about data to find and decide the threshold value to split the ataset. There are some advantages of classification tree. It is good at handling non homogeneous data; it can reduce the dimensionality of data; it can show the hierarchyl of independent variables and their interactions to have insight into classification.

2.2.6 Ancillary Information

Ancillary information such as texture and contextual information can be extracted from the image and are frequently applied to assist image processing. It could be a powerful way to improve classification accuracy when considered together with spectral information for image classification [15]. Because spectral information is so limited to separate objects, especially in areas that combine natural features with man-made objects, it is better to analyze and use spatial properties instead of spectral [19].

Texture analysis plays an important part in object extraction from many types of imageries. From a broad sense, texture can be defined as the spatial distribution and difference of the grey tones in an image [20]. It can be used to distinguish between objects which have different spectral information; moreover, it also can tell the difference between objects which have similar spectral characteristics. The original applications of texture in remotely sensed images were mentioned by Haralick [21]. If the grey level variation between tonal regions is wide, then it is fine texture; otherwise it is coarse texture [21, 22]. One of texture models is called grey level co-occurrence matrix (GLCM), it describes the spatial interrelationships of the grey level in textural pattern by using specific texture features. Some common texture features in co-occurrence matrix are introduced below:

$$\text{Contrast: } \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i-j)^2 g(i,j) \tag{2-1}$$

$$\text{Correlation: } \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \left[\frac{(i-\mu)(j-\mu)g(i,j)}{\sigma^2} \right] \tag{2-2}$$

$$\text{Entropy: } \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} g(i,j)\log(g(i,j)) \tag{2-3}$$

Homogeneity: $$\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \left[\frac{g(i,j)}{1+|i-j|} \right] \quad (2\text{-}4)$$

Where i, j are the number of row and columns of co-occurrence matrix; n $\times$ n is the dimension of matrix; g(i,j) is the value at row i and column j in matrix; μ and σ are mean and standard deviation, respectively.

Contextual information can be separated as global context and local context. Global context describes sensors, time and location; local context describes the common relationships between image regions and is based on how detailed information (scale) of regions is [5]. Contextual classifier considers the spatial context of each pixel to improve classification results. The spatial context is based on the properties of neighbouring pixels; or it can be seen as the relationship between one pixel and remining pixels in the whole scene [16]. Object-based approach is more meaningful in statistics and texture analysis as it considers topological features and the relationship between objects. When image objects are not only based on their spectral and texture characteristics but also based on sub-level and super-level relation of objects and classes, then objects contain contextual information that also can be seen as semantic information. Therefore, a semantic network is more easily to be described by using object-based method.

2.3 Object extraction

Topographic objects such as buildings, roads, trees and water bodies can be extracted from images. Feature extraction can be seen as object recognition and reconstruction process and is important for updating GIS databases [23]. Conventional manual feature extraction is a time consuming and low efficiency process, therefore automatic or semi-automatic extraction techniques from images become more and more popular but big challenges still exist.

According to the complexity and purposes, feature extraction can be divided into extraction of points, edges and regions with different techniques, respectively [11]. Extraction of points is used for object corners, adjacent points or height points. Edges extraction is mainly for road extraction based on edge detection algorithms. The common method for region extraction of objects such as water bodies and vegetation areas is a region growing algorithm which can detect groups of pixels with homogeneous characteristic in the area. The idea is selecting a single pixel or a small area as seed, determining which properties of the seed would be used and giving threshold to the selected properties; the algorithm will compare the neighbouring pixels and evaluate their homogeneity.

The paper presented by Baltsavias [24] summarised the techniques, current status as well as the trend for object extraction. It includes pure image processing methods such as artificial neural networks and fuzzy logic; 3D methods such as object-based, hierarchical and multi-image which often combine image processing and modelling. Although there are many semi-automatic and automatic object extraction systems, with increasing and various requirements from users, they are limited by some conditions such as lack of prior knowledge, complicated realistic situation and the processing method is not so flexible, in this case manual processing plays an important role. Object-based modelling and knowledge-based modelling for object extraction are being increasingly applied in last few years. The knowledge used in the model may depend on the data used for extraction, objects to be extracted and their context information from the scene such as relationship between two objects and geometric information of objects. The difficulty in these model-based approaches is that it needs to determine all possible building shapes and types which may be very different from one building to the others.

Extraction of building footprints can be seen as extraction of edges or areas depending on which techniques are used. It is a man-made objects recognition problem which is not only used for the detection but also for the reconstruction of buildings in image processing. Region growing and region merging is one way to extract objects, but sometimes this will give artificial results and sometimes it will give uncertain information, for example if the ground is covered by shadows, region growing algorithm will stop and shadow part will not be assigned to ground. Other methods for building extraction such as image classification to classify image into different thematic classes; edge detection algorithm can detect building boundaries according to texture variation; model-based approach focuses on integration of existing knowledge as rules to extract buildings. In recent years, to achieve different purposes, building extraction research was no longer limited to 2 dimensional cases but shifted and focused on 3 dimensional building extraction studies; and related data used for building extraction extended from aerial photos to high resolution images, digital surface model, SAR and Laser scanning data.

2.4 Accuracy assessment

Since classification process has been implemented, it needs to determine the accuracy in final classification result by comparing it with ground truth data. The classification accuracy is not only influenced by the classification approach but also depends on the accuracy assessment method; therefore it is important that data users and researchers have knowledge about the evaluation techniques [25]. Traditional approach commonly used for classification accuracy assessment is the error matrix which can calculate producer's accuracy, user's accuracy, overall accuracy and Kappa coefficient based on randomly selected samples (pixels) [25, 26]. Therefore this method usually indicates the quality of classification related with positions of classes. According to the experience of Zhan [26] in the case of single class, for instance building extraction, randomly generated samples would overestimate the classification result.

Another approach for accuracy assessment, especially for evaluation of building extraction accuracy, is the pixel-based comparison between ground truth and classification of building result and was mentioned by [27-30]. Common method is converting reference polygon layer into raster layer and then overlaying with classification raster layer to get the difference between the two. The idea of this method is based on four factors:

$$\text{Branching factor: } \frac{FP}{TP} \tag{2-5}$$

$$\text{Miss factor: } \frac{FN}{TP} \tag{2-6}$$

$$\text{Detection percentage: } 100 \times \frac{TP}{TP + FN} \tag{2-7}$$

$$\text{Quality percentage: } 100 \times \frac{TP}{TP + FP + FN} \tag{2-8}$$

Where,

TP (True positive), both reference data and extracted result are buildings;
TN (True negative), both reference data and extracted result are background;
FP (False positive), extracted result is building but reference data is not building;
FN (False negative), reference data is building but extracted result is not building.

However, this method is very sensitive to difference of pixels between truth data and extraction result. For example, in most of researches ground truth was obtained from manual digitized polygon and converted into raster file. Therefore, the error from digitizing process and dataset conversion may influence the accuracy. Furthermore, if the building was extracted correctly but may be not matches every pixel from the reference data then there would also be an error.

Extraction process is based on objects, so accuracy assessment should also be based on matching of objects. Object-based accuracy assessment method was presented by Zhan [26]. The idea is overlaying two datasets and assuming that if the common area of overlaid two objects covers at least 50% area of object from reference data then they are the same object. This approach avoids the problems from pixel-based method.

2.5 Related work

Automated building extraction techniques can be studied according to the different data sources. Commercial high resolution multispectral satellite imageries are more and more popular in this field such as Quickbird and IKONOS; aerial photographs are also valuable because of its very high spatial resolution; other data source such as Airborne Laser Scanning (ALS) data has been widely used in extraction of 3D buildings for city models. Building extraction is becoming a more and more challenge task from various data and it helps people for understanding and analyzing ground objects from different data. There are some researches about developing and applying various remote sensing techniques for building extraction.

Many automatic building extraction techniques were presented in recent years. The paper from Mayer [23] reviewed lots of researches especially focused on building extraction. It presented strategies and approaches which were selected for building extraction in previous works, and the criteria for assessment of these approaches. Classification based approaches for building extraction were presented by San [29], Mason [31] and Lee [32]. In San and Mason, multispectral image was classified to separate building class from other classes; digital terrain model (DTM) was used for refining building class through the height differences. Class-base method from Lee which a first supervised classification was used for obtaining approximate locations and shapes of buildings then unsupervised classification (ISODATA) was applied for extraction of low contrast buildings. Hurskainen [33] applied object-based classification on different dates images to detect informal settlements and changing.

Extraction of buildings from images is a complicated process because some building roofs do not have regular shapes and forms, and the materials of roofs are also different. Some literatures presented method that applying texture analysis for building extraction [19, 20, 34]. Integration of texture information extracted from images is used as ancillary data and classification result to improve building extraction result. Using shadows as contextual information for assisting building detection was presented by Wei [35]. The methods were based on mathematical morphological techniques, such as Jin [27] and Shackelford [36]. Differential morphological profile (DMP) was applied by changing various size and shape of structure element to detecting buildings and their shadows. Especially in Jin's work, morphological approach was used for providing building shapes as structural information and shadows as contextual information, then integrating these two ancillary information and spectral characteristics to extract buildings. Lari [37] developed an Artificial Neural Network (ANN) system in their research for automatic building extraction based on structural and spectral information from high resolution images.

Building footprint information is needed for many applications such as updating cadastral databases, management of urban areas, damage assessment after disaster and creation of 3D models. Some building detection approaches were developed for post-disaster cases such as earthquake. Uncollapsed buildings were extracted from post-disaster imagery by using edge detection based on shadow information; then extracted buildings were compared with pre-disaster building polygons to obtain collapsed buildings result [38]. Bitelli [39] compared pixel-based and object-based classification for extraction earthquake damage information then applying change detection method to obtain damaged areas and buildings.

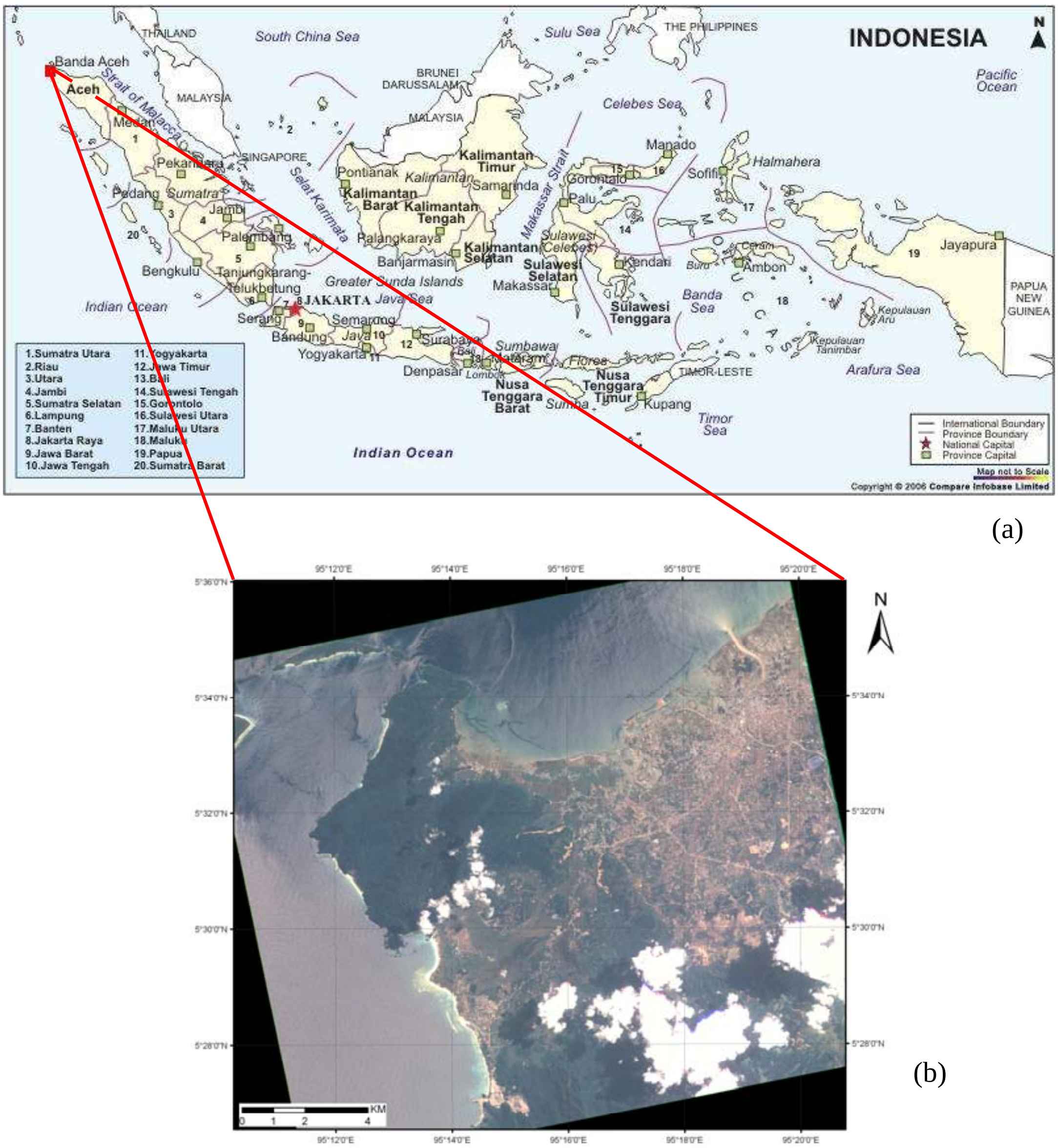
2.6 Summary

This chapter reviewed the knowledge about different image segmentation techniques which are thresholding method, edge-based segmentation and region-based segmentation; various classifiers which are traditional pixel-based classifier, object-based classifier, fuzzy classifier, decision tree and some ancillary information for assisting the classification. Then it reviewed the knowledge of object extraction and accuracy assessment approaches; some previous studies were reviewed and focused on building extraction.

3 Study area and data description

3.1 Study area

The study area of this research is Banda Aceh (Figure 3.1(a)). It is the capital city of Nanggroe Aceh Darussalam (NAD) Province and located in northern part of Sumatra. The approximate geographic location of Banda Aceh is 5°33'N and 95°19'E and the elevation of Banda Aceh is 21 meters.



(a)

(b)

Figure 3.1 (a) Map of Indonesia shows the location of Banda Aceh; (b) KOMPSAT-2 multispectral bands show the scene of Banda Aceh
(Source: http://www.mapsofworld.com/indonesia/indonesia-political-map.html)

On December 26, 2004, a giant tsunami was triggered by an Indian Ocean earthquake and Banda Aceh was the closest city to the epicentre and the worst hit area. After the 2004 tsunami, there were several other earthquakes in this area, the population of Banda Aceh reduced from original 264628 persons to nearly 203553 persons [40]; most of buildings, road networks and infrastructures were destroyed during the tsunami. A quick assessment of the settlement area damages was made by LAPAN [41]. Approximately, 74% of settlement area had been destroyed. Thousands of residential buildings and other kinds of buildings had various degrees of damages. Soon after the disaster, some new houses were built up in this area by international donors.

3.2 Data description

KOMPSAT-2 Satellite Imagery

KOMPSAT-2 (Korean Multi-Purpose Satellite) was launched by Korean Aerospace Research Institute (KARI) in July 2006. KOMPSAT-2 imageries were provided by KARI under agreement from INTOSAI (International Organization of Supreme Audit Institutions) Tsunami Task Force, distributed by SIM-Centre, BRR NAD-Nias. The imagery used in this study was “200704_NP043_Banda Aceh”, the original image contains four multispectral bands and one panchromatic band.

The characteristics of KOMPSAT-2 imagery are shown in Table 3.1. There were several imageries from different angles of view which covered Banda Aceh, and one of them was named as “200704_NP043_Banda Aceh” that had no clouds of influence. The scene information of this imagery was shown in Table 3.2.

Table 3.1 Characteristics of KOMPSAT-2 imagery
(Source: SPOT IMAGE)

Spatial resolution	Panchromatic: 1m; Multispectral: 4m;
Dynamic range	16 bits;
Band wavelength	Multispectral bands: Band1 (Blue): 0.45 ~ 0.52 µm; Band2 (Green): 0.52 ~ 0.6 µm; Band3 (Red): 0.63 ~ 0.69 µm; Band4 (NIR): 0.76 ~ 0.9 µm;
	Panchromatic band: 0.50 ~ 0.9 µm;
Orbit height	685 KM;

Orbit type	Sun Synchronous;
Footprint	15 KM x 15 KM;

Table 3.2 Scene information of “200704_NP043_Banda Aceh”
(Source: BRR-SIM Centre)

Image Set	Acquisition			Multi-spectral	Pan-chromatic
	Date yyyy-mm-dd	Time UT	Time WIB		
200704_NP043_Banda Aceh	2007-05-25	04:34:42	11:34:42	☑	☑

	Geographic		Projected		Image			
	Latitude	Longitude	Longitude	Latitude	Multi-Spectral		Pan	
Vertex	Northing (dd[1])	Easting (dd)	Easting (m)	Northing (m)	**X** [2]	**Y**[3]	**X**	**Y**
Top Left[4]	5.57497868	95.16139546	739435.232620	616661.034313	1	1	1	1
Top Right	5.60806394	95.32796834	757883.033178	620391.585761	3750	1	15000	1
Lower Left	5.42893723	95.19158988	742840.990896	600518.254018	1	3750	1	15000
Lower Right	5.46204132	95.35811909	761289.032622	604250.037063	3750	3750	15000	15000

Orthophotos

Orthophotos (total number is 44) covering the whole area of Banda Aceh are available from Bakosurtanal (Indonesian Cartographic Institute) via SIM-Centre. These orthophotos were generated from aerial photographs of 2005 and have been georeferenced in UTM projection coordinate system zone 46(N), WGS 84 datum, with spatial resolution 30cm. These orthophotos were used for georeferencing KARI images.

[1] Decimal Degrees, World Geodetic System 1984 (WGS 84)

[2] Image X is Column or Sample

[3] Image Y is Row or Line

[4] In ENVI top left is image location 1,1, (x,y)

Vector datasets

Vector data layers which were already georeferenced to UTM zone 46(N), WGS 84 datum and digitized from orthophotos mentioned above. These vector layers contain information about the polygons of survived buildings from tsunami, the polygons of rebuilt buildings in pre-reconstruction, the road polygons of Banda Aceh.

4 Methodology

The general workflow for this research is shown in Figure 4.1. Firstly, imagery was registered to the same projection system by using image to image registration. Secondly, building footprint extraction based on object-based image classification was performed. There were two main sub-steps which are segmentation and object-oriented classification with decision rules in this part. And then different approaches of accuracy assessment were used in the research for evaluating the accuracy of building extraction. Last but not the least, new buildings were separated from old buildings.

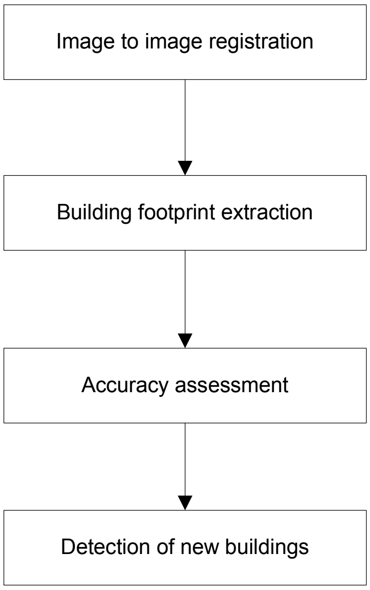


Figure 4.1 General frame of the research

4.1 Image to image registration

There are several approaches for image registration, such as map based registration and image to image registration [42]. In this research, the latter method was used. The principle of image to image registration is setting an already georeferenced image with higher resolution as master data to register the raw image which is so the called slave. To compare with map based registration, the advantage of this image to image registration method is that it is a one step process by selecting the reference points on the master image and finding the corresponding points in the slave image.

In this research, firstly orthophotos of 2005 of Banda Aceh were used as master image to register the panchromatic band (as slave) of KOMPSAT-2 imagery. Then this registered panchromatic band was used as master to register multispectral bands of KOMPSAT-2 imagery. The main workflow of image to image registration is shown in Figure 4.2.

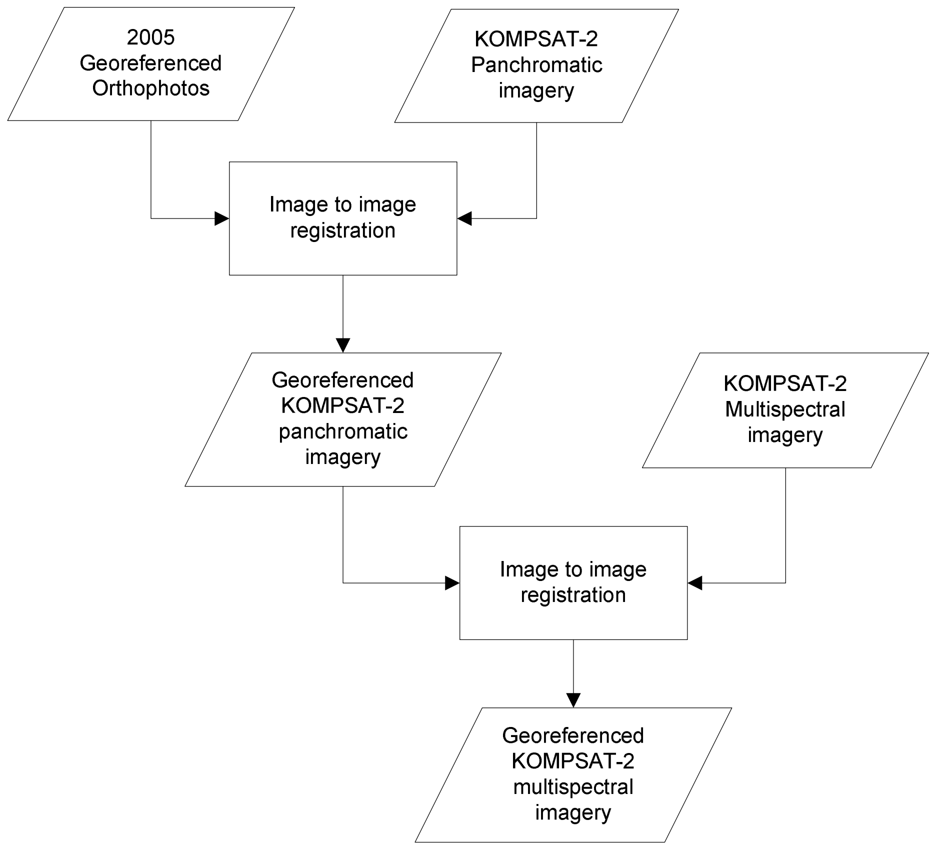


Figure 4.2 Image to image registration process

4.2 Building footprint extraction

Man-made objects extraction from high resolution satellite imageries based on different techniques were studied frequently in recent years. Proposed method in this research was building footprint extraction based on object-oriented classification with decision rules, then integrating with GIS tools to verify existing buildings. Object-based building footprint extraction is mainly divided into segmentation and object-based classification steps. The applied workflow for the building extraction process is shown in Figure 4.3.

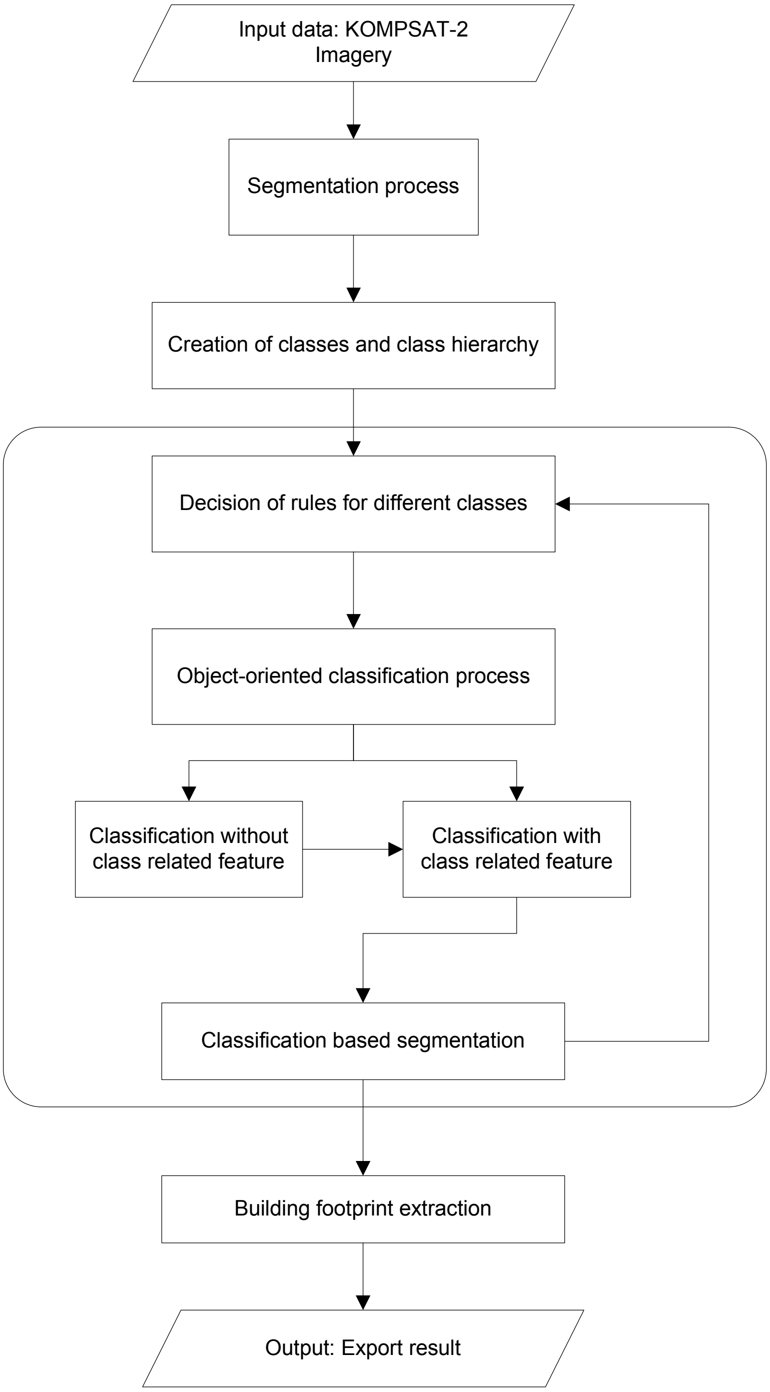


Figure 4.3 Applied workflow of building footprint extraction

4.2.1 Segmentation

In object-oriented image analysis, segmentation is the first and an important step that is carried out before performing object-oriented classification. The principle of segmentation is to split the whole image into different objects according to their spectral characteristics and size, and each object has its own properties [43]. There are mainly two types of segmentation: one method is bottom-up and the other is top-down segmentation. The bottom-up method can be seen as data compression, in this method the segments are generated based on statistical methods and parameters for processing the whole image. The top-down method is a kind of “knowledge driven method” [44] and implemented by generating a model of target objects. Bottom-up segmentation method was applied in this research, which is the most common method and can be handled relatively easy.

The segmentation was starting at the level of pixels. Pixels having similar spectral values were grouped into the same object. On each level, the objects are based on sub-objects from the previous level and merged into super-objects on the next level. Finally, a hierarchical network of image objects should be created after the segmentation (Figure 4.4). This process was repeated several times to get suitable segmentation parameters. The result of segmentation depends on the segmentation algorithms and parameters, and also depends on the homogeneity of spectral reflectance from ground objects [44]. A thematic layer can be used as an additional source. It is quite useful ancillary data for segmentation.

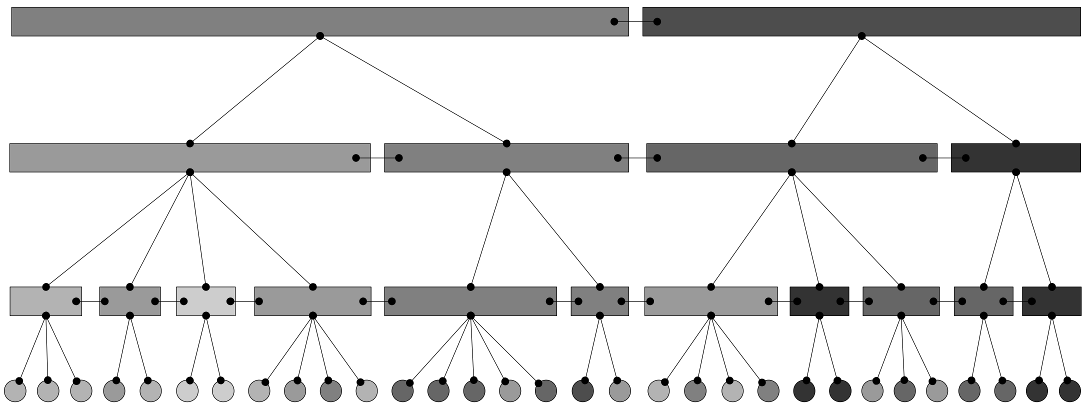


Figure 4.4 Hierarchical networks of image objects [44]

4.2.2 Object-oriented classification

Object-oriented classification was performed after segmentation. It is not like conventional pixel-based classification, but it is based on objects which have been obtained from segmentation. Objects such as building roofs, roads and bare land have similar spectral values, which makes it difficult to separate them using only spectral information. In addition, even the same objects, for example building roofs, can be made of different materials and may be separated into different classes. Therefore, spectral, spatial, textural, contextual and semantic information can be used in the classification. Conventional multispectral classification is mainly based on spectral values but most of the time this is not enough to get a satisfactory result. Combination of spectral characteristics with other information is a way to improve classification accuracy. Texture is an important factor used in image classification and analysis; it can be measured from one pixel and its neighbourhood pixels values [45]. In this case, brightness value and variance of grey value can be used. Contextual and semantic information, for instance, spatial relationship between two objects, can be applied during the classification.

Object-based classification was done in eCognition software. The strategy of object-based classification was a kind of decision tree approach. For each branch in this decision tree, there were two classes defined according to the complementary conditions. It means that the objects would be assigned to class I if they were under condition “A”, otherwise they would belong to class II as is shown in Figure 4.5. In the study, classes were mainly based on the colour of building roofs which can be observed from multispectral bands with true colours.

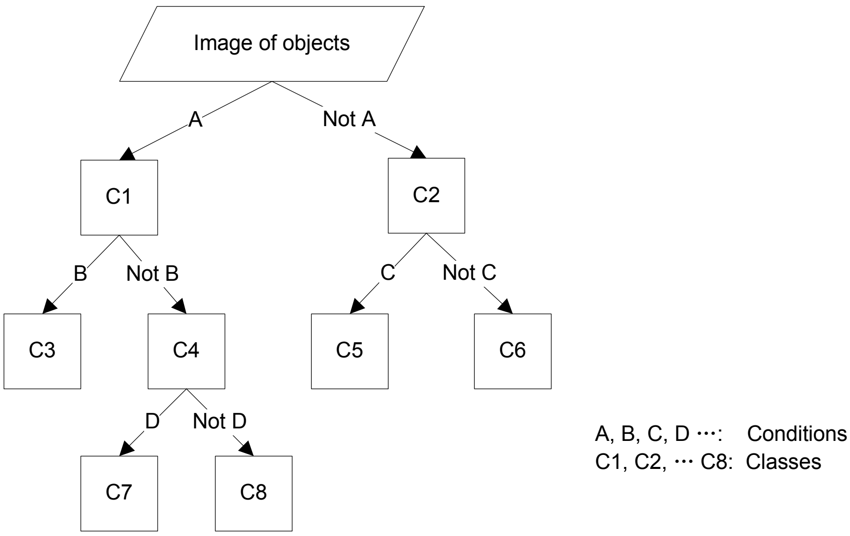


Figure 4.5 Decision tree based image classification

4.3 Classification accuracy assessment

4.3.1 Traditional approaches

Accuracy assessment is performed after classification and consists of comparing classification result with ground truth (reference data) to assess how accurate the classification is. Traditional accuracy assessment method for classification is based on generating random points and error matrix, then calculating user’s accuracy, producer’s accuracy, overall accuracy and K-statistics [25, 26]. Most of studies have been using this method for evaluation of classification accuracy.

4.3.2 Object-based approaches

In this research, object-based approaches for accuracy assessment were applied. The first one was overlaying the extracted building footprint with the reference data of building footprint based on their geometric centres, and it was called “GC” method in this study; to take the shape and size conditions into account, the second method was based on using a “bounding box” to both reference data and result followed by calculation of the shape and size ratio, and it was called “BB” method; then the third one is the combination of “GC” and “BB”.

The first approach was the geometric centres from extracted objects (polygons) overlaid with reference objects (polygons), and applying certain conditions to judge whether they may be the same or not. Those conditions were, (a) overlaying extracted building footprint with reference data, if the geometric centre of extracted building footprint is located into the reference data then this extracted footprint should be a house in real; (b) intersection of reference data with those already judged as real houses (from previous step) then the intersected reference data should be recorded as detected buildings.

The second method was using a “BB” on both the reference data and the extracted result to consider the shape and the size of buildings in accuracy assessment. The process of applying “BB” method is shown in Figure 4.6. “BB” is one way to simplify the polygons. It transforms the original polygon to a bounding rectangle according to its horizontal and vertical maximum range. After a “BB” was applied, the polygons became more regular shape than before; in addition, the ratio of length and width of each “BB” was calculated (S1 and S2), the area of each “BB” was also calculated (A1 and A2). The shape condition was based on the ratio of S2 and S1, and the size condition was based on the ratio of A2 and A1. After the shape and size conditions were obtained, some extracted objects were limited by either size or shape and not considered as buildings. One way to complement one condition to another is the combination of shape and size. To combine these two conditions, the average value of the shape and size conditions was calculated.

The third method was the combination of “GC” method and “BB” method. The detailed procedure is shown in Figure 4.7. After applying “BB” to the reference data, it was found that for some houses the area or the shape were changed considerably but some were not. This was caused by the different orientation of the houses. Therefore, the combination of those two methods was considered. The reference data was separated into two parts. One part was applied “BB” on the reference data of which the shape or the area were not changed considerably from original reference data; the other part was the rest of the reference data of which the area or the shape were changed. The latter part of the reference data was kept in the original format and used for applying “GC” method.

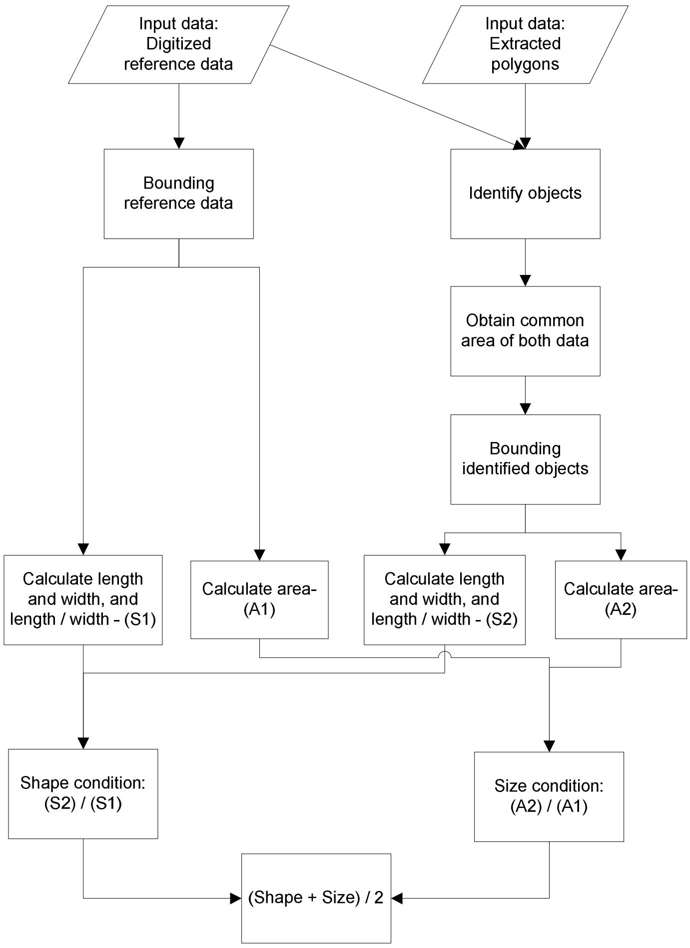


Figure 4.6 The procedure of applying “BB” for evaluation

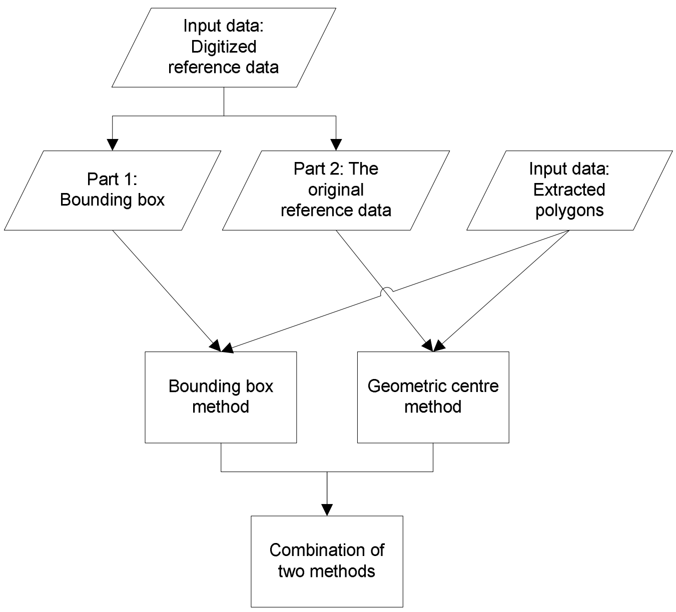


Figure 4.7 The procedure of the combination of “GC” and “BB”

4.4 Detection of new buildings

One of the objectives of this research is to detect the new buildings, and it is a kind of change detection problem. Change detection techniques have been used commonly in damage assessment after disaster [39], for example in earthquake case, the comparison of building data between pre- and post earthquake is to assess the damage of buildings. This technique has also been used for updating of databases [46, 47], for example in urban management, comparing multi-datasets of the same area from different time to analysis the urbanization.

In the past, most of change detection studies were based on comparing two or more raster datasets of different time (years). In this study the building footprints were extracted and evaluated based on objects, therefore the detection of new buildings was also based on the comparison of objects. From the 2005 orthophotos, it can be observed that there existed some buildings before reconstruction projects started. The building footprints which were extracted from satellite imagery showed all the buildings on the ground no matter whether they came from projects or not. Therefore, extracted result needs to be compared with the old building polygons so that the new buildings and old ones can be separated from each other. The main procedure of separating old buildings from the new buildings is shown in Figure 4.8.

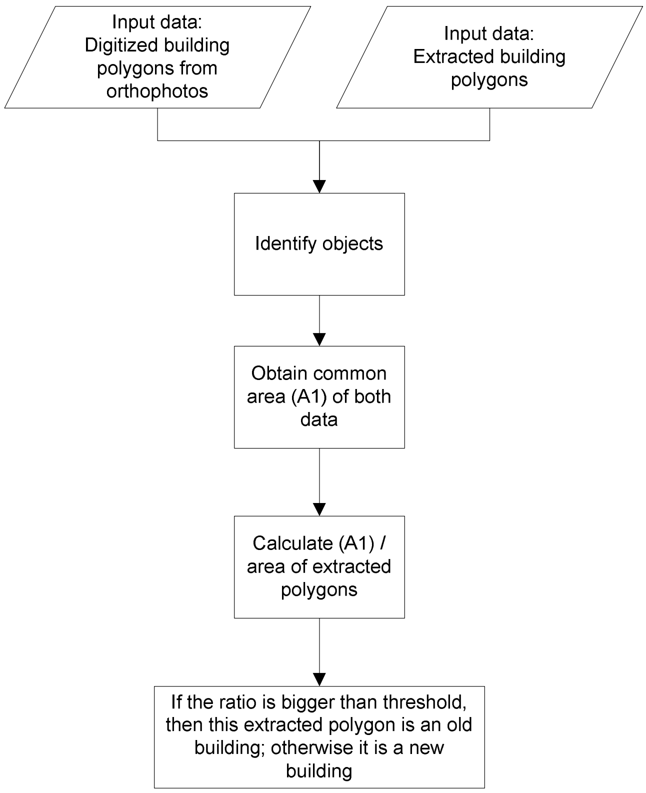


Figure 4.8 The process of detection of new buildings

4.5 Application of eCognition

More and more studies interpret and analyze remote sensing imageries from an object point of view instead of pixels; not only considering about spectral values but also spatial relationships between objects. eCognition is an advanced and powerful software for object-oriented classification. It is developed by Definiens Imaging. The three main procedures are performed through segmentation, object-oriented classification and exporting the result.

4.5.1 Image segmentation

There are several segmentation functions for various image characteristics and purposes. The most common segmentation functions in eCognition are multi-resolution segmentation and spectral-difference segmentation. Multi-resolution segmentation is a bottom up technique and starts with one pixel. In the subsequence processes, small objects are merged into bigger ones [5]. The image objects generation is determined by several factors [44]:

Image layer weights: this parameter can assess image bands (layers) differently depending on their importance or suitability for the segmentation result. The higher the weight which is assigned to one band (layer), the more information of this band will be used during the segmentation.

Scale parameter: this parameter determines the maximum allowed heterogeneity of image objects and influences the size of image objects.

Shape and colour: these are two complementary parameters which influence the way of grouping pixels. The more the shape criteria are set the less the colour similarity influences objects generation.

Smoothness and compactness: as long as the shape parameter is larger than 0, the user can decide whether the objects should be more smooth or more compact.

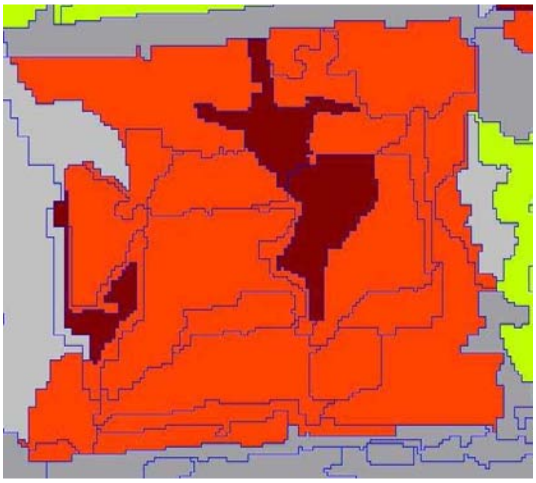
However, it is a big challenge to select the suitable parameters to obtain the optimal result. There is no specific standard for setting segmentation parameters. It depends on the objects of interest. Sometimes the shape factor is given more weight for extracting urban man-made features, and colour is better for vegetation area and water bodies.

4.5.2 Object-oriented classification

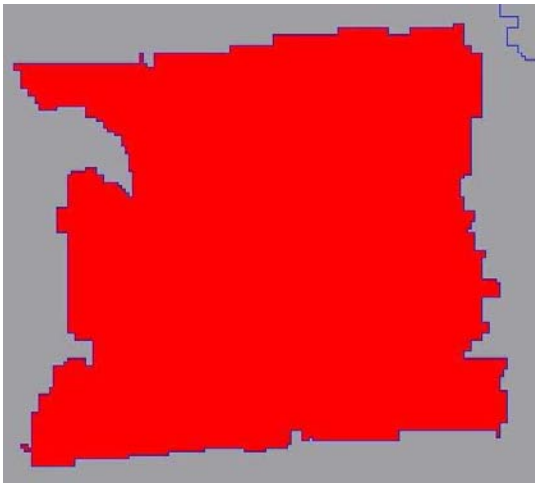
Object-oriented classification in eCognition is mainly based on fuzzy logic. Two classification methods are available: nearest neighbour classifier and classification hierarchy. The former is based on selecting objects as training samples and minimum distance measurement; and the latter is based on designing a classification strategy according to the knowledge about characteristics of image objects. Classification based segmentation is a useful tool for refining classification result through merging objects within the same class or between different classes. An example of this merging function is shown below (Figure 4.9), and the whole example which includes data and processing strategy is from eCogntion example and given by Definiens [48].

Explanation of the example:

From Figure 4.9, it can be seen that (a) shows the objects of sub-level are in different classes. The reddish objects belong to one building roof but classified into two classes; and even in one class there are several objects. (b) shows the objects of super- level. The red part is still the same building in sub-level but as an entire object. Not only the objects that belong to the same class are merged to become one, but also two classes are merged to an entire building object.



(a)



(b)

Figure 4.9 Merging example from eCognition software [48]

Features used for classification are calculated based on image objects but not on pixels. Besides common features contain such as spectral values, others like texture properties, class-related features and object-related features, features based on shape and size are also available. These commonly used object features are shown below:

(1) Layer mean value $\overline{C_L}$, which is calculated from the average of layer value C_x of total number of pixels N contained in an image object;

$$\overline{C_L} = \frac{1}{N}\sum_{x=1}^{N} C_x \qquad (4\text{-}1)$$

(2) Brightness B of layers (L), which is the sum of mean value of all layers divided by the total weight of layers (W);

$$B = \frac{1}{W}\sum_{l=1}^{L} W_l \overline{C_l} \quad (4\text{-}2)$$

(3) Ratio of layer L (R_L), which is the ratio between layer mean value and brightness;

$$R_L = \frac{\overline{C_L}}{B} \quad (4\text{-}3)$$

The target objects are houses therefore shape features were used based on different aspects in this study. These aspects were (i) area of one object which is the sum of pixels number in object; and (ii) ratio between length and width of one object.

The texture attribute which commonly used in eCognition is called “texture after Haralick”. It is based on grey level co-occurrence matrix (GLCM). Each matrix is normalized according to the formula:

$$M(i,j) = \frac{P(i,j)}{\sum_{i,j=0}^{n-1} P(i,j)} \quad (4\text{-}4)$$

where i, j are the row number and column number of matrix respectively, n is the dimension of matrix, P(i,j) is the value in the cell (i,j) of the matrix and M(i,j) is normalized value in the cell (i,j). Based on the experience had been tried in this study and some previous studies, texture features such as homogeneity, contrast and dissimilarity maybe considered more than others [19, 22, 49].

(1) Homogeneity, which measures the distance of elements in the GLCM to the diagonal of the matrix.

$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} \frac{M_{i,j}}{1+(i-j)^2} \quad (4\text{-}5)$$

(2) Contrast, which is the opposite of homogeneity. And it is a measure of the amount of local variation in the image.

$$\text{Contrast} = \sum_{i,j=0}^{N-1} M_{i,j}(i-j)^2 \quad (4\text{-}6)$$

(3) Dissimilarity, which is similar to contrast but increases linearly and it would be high if the local region in the image has high contrast.

$$\text{Dissimilarity} = \sum_{i,j=0}^{N-1} M_{i,j}\left|i-j\right| \quad (4\text{-}7)$$

A class-related feature is a kind of contextual feature attribute. It contains relations to the neighbour objects, to the sub-objects and to the super-objects. The “relations to the neighbour objects” refers to the relationship between assignment of image objects and existing class on the same level of image objects hierarchy; “relations to the sub-objects” refers to the relationship between assignment of image objects and lower level image objects in objects hierarchy.

4.6 Software

Three kinds of software were mainly used in this study. They are eCognition, Erdas and ArcGIS. ECognition was used for segmentation and object-oriented classification; Erdas was used for image registration and generating random points in accuracy assessment; then ArcGIS was used for object-based accuracy assessment and detection of new buildings.

4.7 Summary

This chapter described the methods that were applied in this study. They consist of image to image registration which was followed by building footprints extraction and accuracy assessment, then at the end the new buildings were separated from old ones. The building footprints extraction was done by segmentation and object-based classification which based on spectral properties, texture and contextual information. Accuracy assessment was based on traditional method and object-based method; new buildings were detected by comparing extracted result with orthophotos of 2005. The whole extraction process was done in eCognition software because it is suitable for this approach.

5 Results

This chapter shows the results which the order is following Figure 4.1. Then the results are discussed in chapter 6.

5.1 Image to image registration

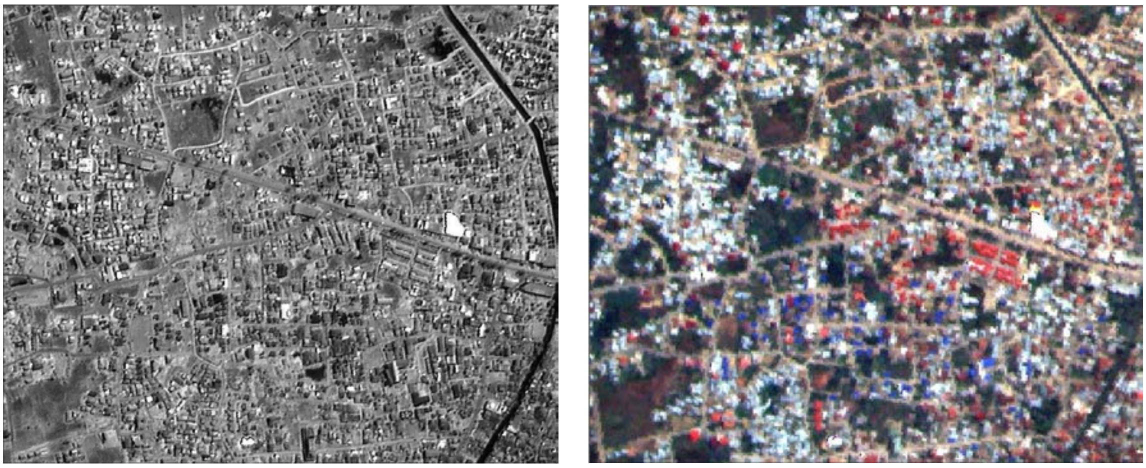
Kompsat-2 imagery "200704_NP043_Banda Aceh" was registered by using image to image registration method. Georeferenced orthophotos of 2005 covering the whole of Banda Aceh were available from Bakosurtanal via SIM-Centre and BRR, and registered in UTM zone 46(N) / WGS84. Therefore, I first used these orthophotos as master image to register the panchromatic band (as slave image). The Root Mean Square Error (RMSE) of registration panchromatic band was 0.0378 m in X, 0.0397 m in Y, and 0.0548 m for the total error. After registration of panchromatic band, then I used this georeferenced panchromatic image as the master image to register multispectral bands (as slave image) resulting in an RMSE equal to 0.0437 m in X, 0.0529 m in Y, and 0.0687 m for the total error. Finally, both the panchromatic image and multispectral images were registered to the UTM Zone 46(N) / WGS 84. This resulted into a spatial resolution 1m and 4m for the panchromatic image and multispectral image, respectively.

5.2 Building footprint extraction

5.2.1 Test area description

The georeferenced KOMPSAT-2 image of Banda Aceh is shown as a true colour image in Figure 5.1. A test area of a size of 1048m × 862m was created as subset from the whole image (the red rectangle box) and is shown as both panchromatic and multispectral image in Figure 5.2. This test area contains vegetation area, water bodies, roads, bare ground and buildings. Buildings which have different colours such as red, blue and bright can all be observed from multi-spectral layers (Figure 5.2 (b)).

Figure 5.1 The Banda Aceh scene from KOMPSAT-2 imagery.
The red square is the test area for this study



(a) (b)

Figure 5.2 (a) The panchromatic band of the test area (spatial resolution: 1m, RMSE: 0.0548m); (b) the multispectral bands (true colour) of the test area (spatial resolution: 4m, RMSE: 0.0687m).

5.2.2 Segmentation

Segmentation is the first step before an object-based classification is performed. Applying segmentation we split the image into segments according to homogeneity of pixels. In this study, segmentation was based on bottom-up method which starts with one pixel and merges small objects into larger ones. KOMPSAT-2 imagery was segmented in two levels by setting different scale parameters and layer weights. To refine the segmentation and improve following object-based classification, road vector layer has been used during segmentation process as thematic layer.

Segmentation strategies are shown in Table 5.1. And the process was repeated several times by using different parameters leading to the optimal result in the end. In the first level, green, near infrared and panchromatic bands were used for multi-resolution segmentation; the reason is that study area is largely covered by vegetation and small houses, therefore it was given more weights on shape factor and the scale parameter cannot be set too large; road vector layer was used to give objects with road attribute. Then in the second level, segmentation based on spectral differences was performed by giving weight 1 to multi-spectral bands and weight 2 to panchromatic band; vegetation area and water bodies were grouped into larger objects, respectively. Building roofs have different colours which can be observed from multi-spectral bands such as red and blue; therefore, blue and red bands were used for segmentation in this level; road vector layer was still used.

Table 5.1 Segmentation parameters were used for each level

Level	Segmentation Mode	Bands Weight	Scale Parameter	Colour / Shape	Compactness / Smoothness
1	Multiresolution segmentation	Green = 1; Blue = 0; Red = 0; NIR = 1; Pan = 2;	7	0.6 / 0.4	0.5 / 0.5
		Thematic: road			
2	Spectral difference segmentation	Green = 1; Blue = 1; Red = 1; NIR = 1; Pan = 2;	12	0.7 / 0.3	0.5 / 0.5
		Thematic: road			

5.2.3 Object-based Classification

Object-based classification on different levels was performed after segmentation; level one was used for providing information of sub-objects and level two was used for mainly for classification procedures.

In level one, objects with the thematic attribute “road” were assigned to road class; shadows were classified by using their low spectral reflectance in panchromatic layer and high texture dissimilarity. Some very bright roofs were also separated from other objects in this level depending on their high reflectance. The result of classification of level one is shown in Figure 5.4.

In level two, a rule-based classification decision tree was set to classify different features. The detailed strategy for assigning classes is shown in Figure 5.3. In this level, normalized difference vegetation index (NDVI) was calculated by using NDVI algorithm to separate vegetation area and water bodies easily:

$$\mathrm{NDVI} = \frac{NIR - \mathrm{Re}\,d}{NIR + \mathrm{Re}\,d}$$

During classification, at first only the spectral values were considered. This was not sufficient to get a good result. In the test area many natural features are mixed with man-made features and the spatial distribution of those man-made features is irregular; therefore other conditions such as texture variance of panchromatic imagery, contextual information and semantic relationship between level one and two and class related features were used to assist classification.

Texture dissimilarity was used to separate objects (e.g. houses) from their neighbour objects. Class-related feature, for instance, “relations to sub-objects” was used to include or exclude the objects which were in the lower level. Thematic layer was also used for classification and its attribute was used in the same way as image layer attributes. Houses were classified based on the colour of their roofs, which can be seen from multispectral bands, and assigned to classes such as blue, red, bright and dark. To refine the classification result in level two, object merging function was using for merging objects which belonged to the same class. The classification result of level two is shown in Figure 5.5.

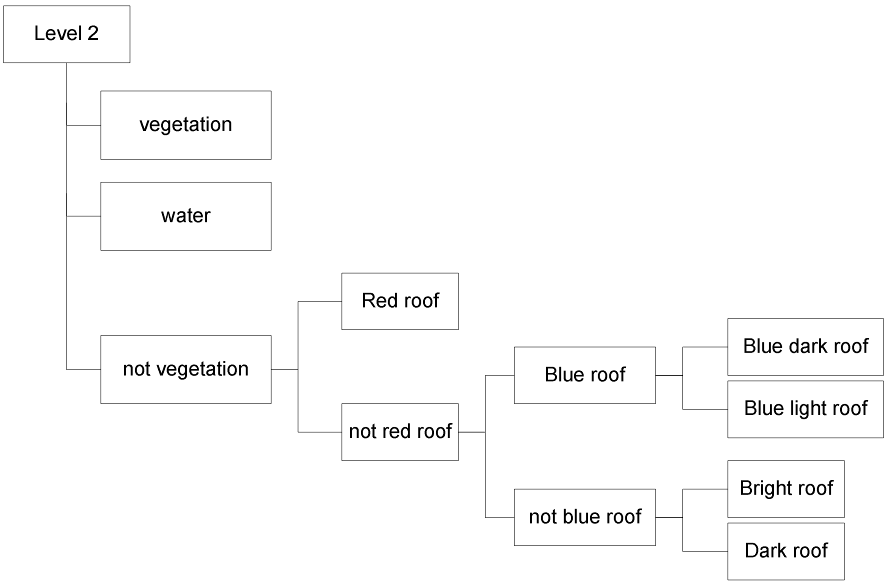


Figure 5.3 Rule-based classification decision tree in level two

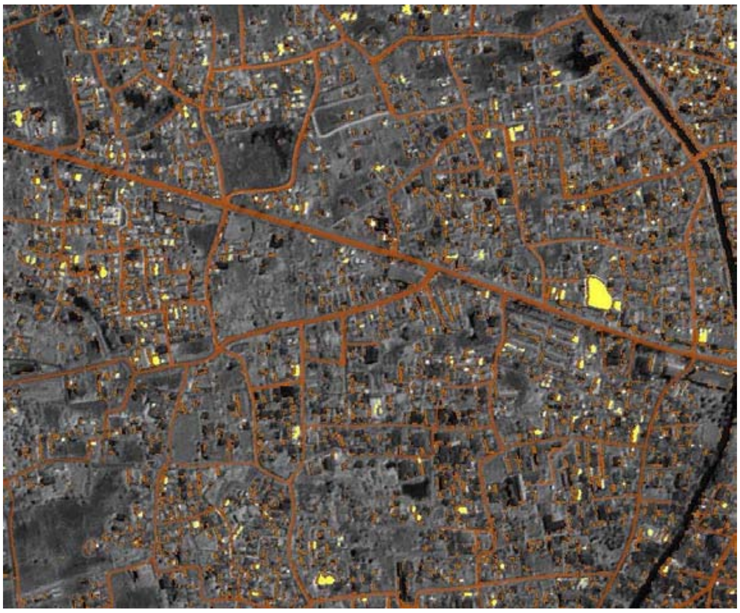


Figure 5.4 Object-oriented classification in level one

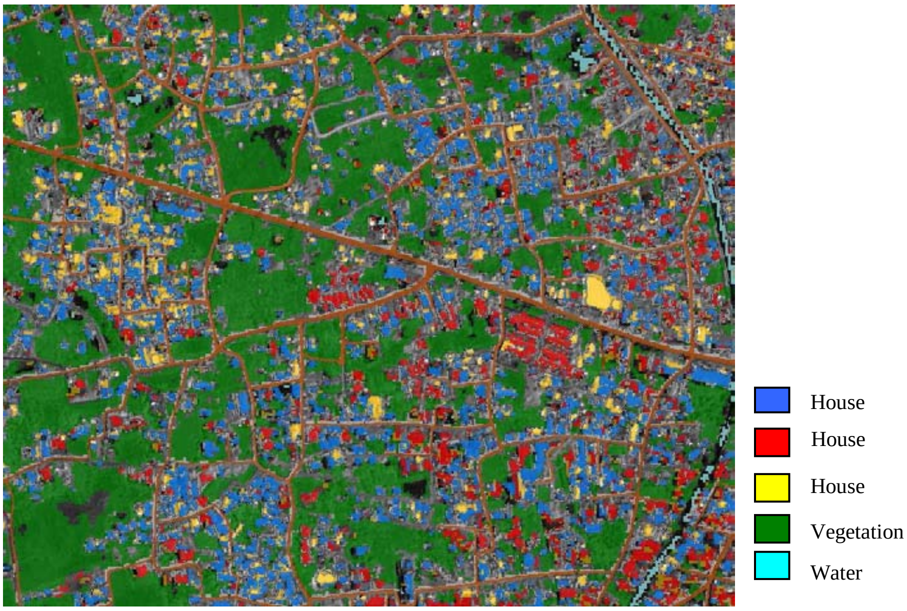


Figure 5.5 Object-oriented classification in level two. Blue, red and yellow are the type of houses; green is vegetation area and bright blue is water body

5.2.4 Building footprint extraction

To extract building footprint from the classification result, level three was created from classification based segmentation. Two classes were generated, namely buildings and background, by applying class-related feature to building classes of level two. Then building classes were extracted as polygon layer and shown in Figure 5.6.

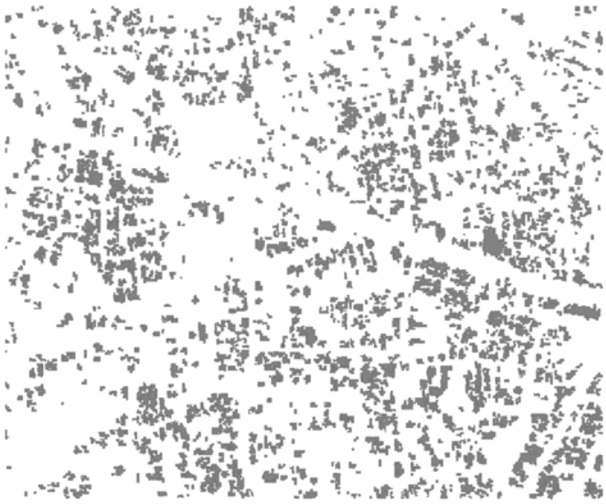


Figure 5.6 Extracted building polygons from object-based classification

5.3 Accuracy assessment

5.3.1 Accuracy based on error matrix

To evaluate classification accuracy, the traditional method is to create an error matrix by comparing the classification result with reference data. In this case, classification of building roofs in level three was exported as raster data. Then 500 random points were generated and given reference attributes either houses or background by visual interpretation of the imagery. The error matrix was created and is shown in Table 5.2 with 63.39% producer's accuracy and 98.61% user's accuracy in building class and 91.60% overall accuracy, and the overall K-statistics equals to 0.7232. The discussion of this error matrix is shown in the section 6.2.

Table 5.2 Error matrix

	Reference data			
Classification	Roof	Background	Total	**User's accuracy**
Roof	71	1	72	98.61%
Background	41	387	428	90.42%
Total	112	388	500	
Producer's accuracy	63.39%	99.74%		**Overall accuracy:** 91.60%
K-statistics	0.7232			

5.3.2 Accuracy based on object-based methods

Building extraction based on object-oriented classification was the main method in this research; therefore accuracy assessment in this case should also be based on objects and not on pixels. Object-based accuracy assessment was presented by Zhan [26].

Because there is no more accurate or detailed (higher resolution) data available for evaluation, the reference data (ground truth) was collected by manual digitizing of KOMPSAT-2 imagery which was already pansharpened by fusing panchromatic band and multispectral bands. A part of the test area (Figure 5.7(b)) was selected for manual digitizing and accuracy assessment. Object delineation by hand requires interpreter with good visual interpretation skills. The interpreter combines colour, size, texture, location, shape and pattern visible in these images with knowledge from the area to detect and delineate the buildings [50]. Building footprints were manually digitized to polygons as reference data; and it was assumed that digitized data had enough accuracy to be used as ground truth. Then building footprints from image classification were exported as vector layer.

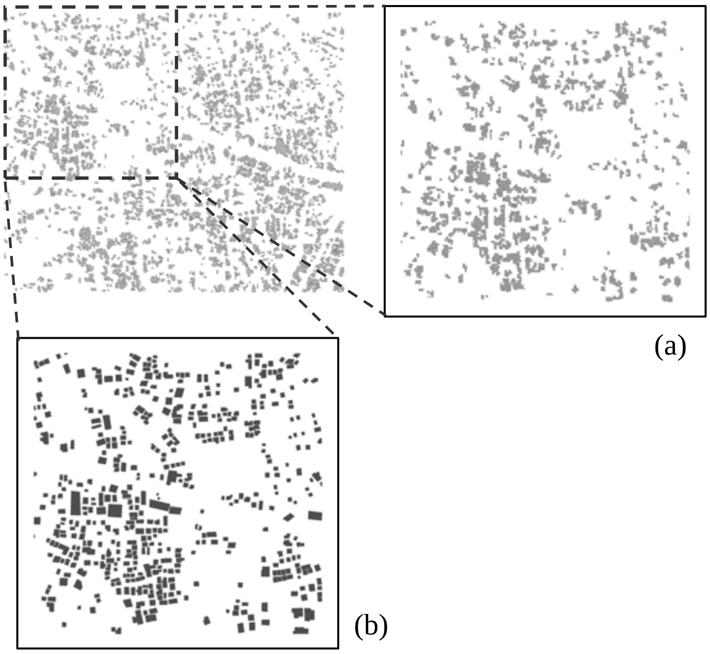


Figure 5.7 Area in red box for object-based accuracy assessment;
(a) is extracted polygons; (b) is manually digitized polygons.

1. Accuracy based on geometric centre ("GC")

To assess the accuracy of the digital object extraction, first all polygons of extracted building footprints (Figure 5.7(a)) were overlaid with reference polygons (Figure 5.7(b)). To decide whether an extracted building footprint was correct, two conditions were used. The first condition was that if the "GC" of (a) was located in (b) then this extracted polygon was considered a real building on the ground. The second condition was that intersection of (b) was applied with those "real buildings" from the previous step, and those intersected reference data were recorded as detected buildings. The result of this method is shown in Table 5.3. "Over" means building footprints which were extracted using object-based classification but not in digitized data; "Missed" means building footprints which were digitized as reference data but not detected by using classification method.

Table 5.3 Accuracy of using "GC" method and without threshold

	(a)	(b)
Recorded number of objects	508	447
Over	181	0
Missed	0	49
Total number of objects	689	496
	Correctness: 73.73%	Completeness: 90.12%
Overall accuracy	80.59%	

Notes: (a) represents extracted building footprints by using object-based classification;
(b) represents manually digitized building footprints.

Selection of threshold

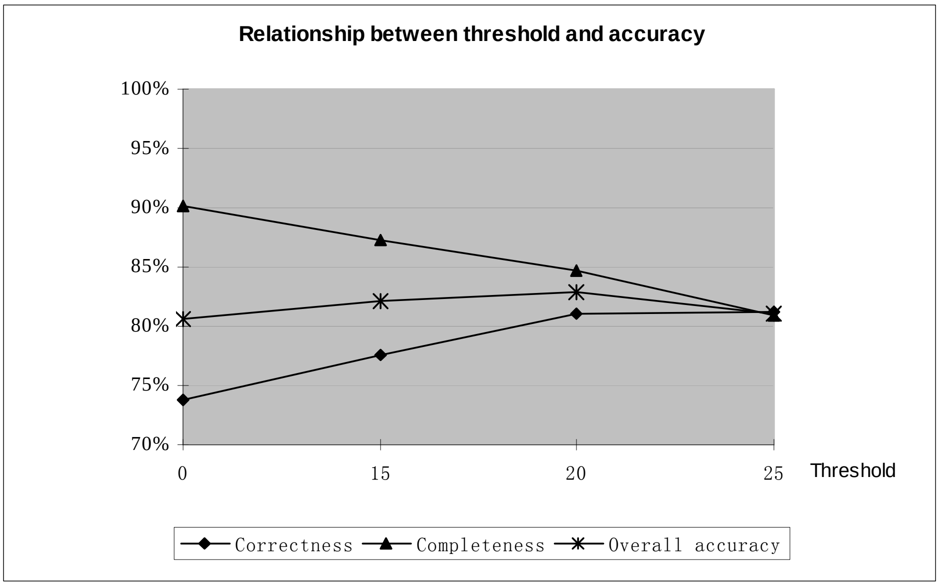
To remove small objects which may influence accuracy, three different thresholds of the area were used (Table 5.4); correctness which also refers to user's accuracy and completeness refers to producer's accuracy; overall accuracy which was calculated by using the number of matched objects divided by total number of objects. The relationship between threshold and accuracy is shown in Figure 5.8. A high threshold corresponds with a high correctness but with a low completeness. The overall accuracy increased with increasing area threshold and reached the highest value at threshold 20 m^2; but it decreased when the threshold was larger than or equal to 25 m^2. Therefore, according to the curves of overall accuracy, correctness and completeness (Figure 5.8), threshold larger than or equal to 20 m^2 was the optimal value in this case.

Table 5.4 Accuracy based on selection of area threshold

	(a) >= 15 m^2	(b)	(a) >= 20 m^2	(b)	(a) >= 25 m^2	(b)
Recorded number of objects	451	433	409	420	373	401
Over	133	0	96	0	86	0
Missed	0	63	0	76	0	95
Total number of objects	580	496	505	496	459	496
	Correctness:	Completeness:	Correctness:	Completeness:	Correctness:	Completeness:

	77.59%	87.30%	80.99%	84.68%	81.26%	80.85%
Overall accuracy	82.16%		82.82%		81.05%	

Notes: (a) represents extracted building footprints by using object-based classification;
(b) represents manually digitized building footprints.



Notes: The unit of threshold is m^2

Figure 5.8 The relationship between threshold and accuracy

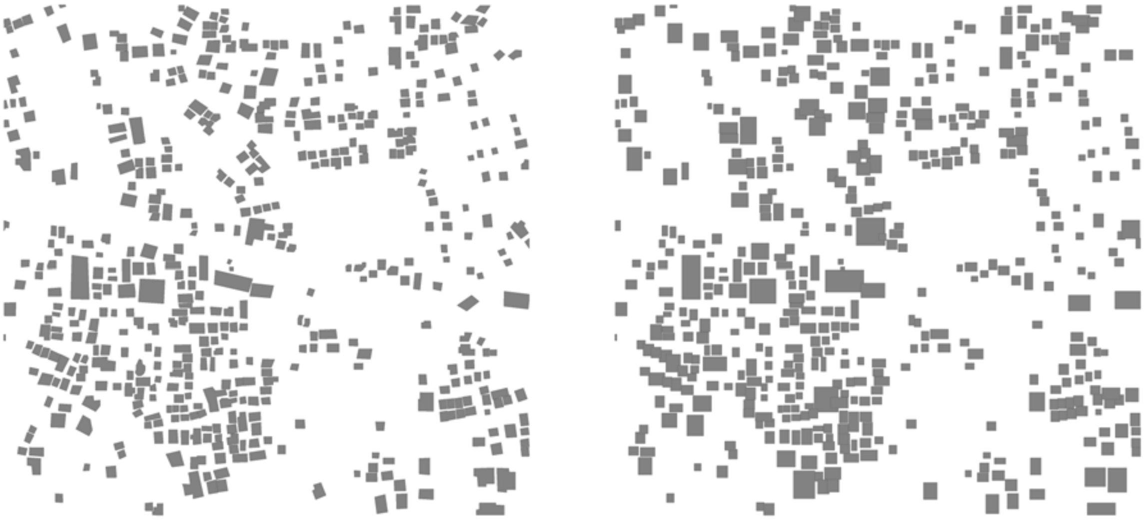
2. Accuracy based on “bounding box” (“BB”)

The whole process of applying “BB” was shown in Figure 4.6. From the flowchart, it can be seen that before applied “BB” to extracted polygons, it was identified by overlaying digitized polygons to obtain the “index ID” from digitized polygons so that the two datasets were linked together. Then the ratio of length and width and the area of each bounding box in two datasets were calculated separately; the ratio of length and width ratio of two datasets (S1 and S2) and the ratio of areas of two datasets (A1 and A2) were calculated.

To combine the shape and size conditions, the average value of shape ratio and size ratio was taken. The results of applying “BB” to two datasets are shown in Figure 5.9(b) and Figure 5.10(b). The comparison of differences between bounding reference data and bounding extracted polygons was applied to the average value from integration of shape and size conditions. The threshold of combination of shape and size was based on the histogram of that average value (Figure 5.11). In this case, threshold of combination value between 0.5 and 1.5 was chosen and the accuracy was calculated and shown in Table 5.5.

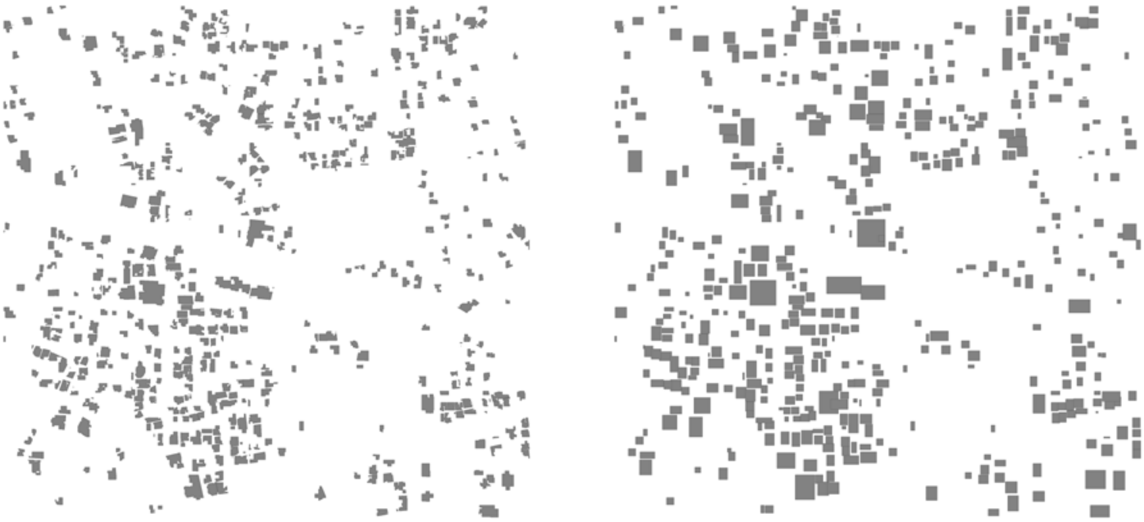
Table 5.5 Accuracy based on “BB” method

	Extracted result	Reference data
Recorded number of objects	567	409
Over	122	0
Missed	0	87
Total number of objects	689	496
	Correctness: 82.29%	Completeness: 82.46%



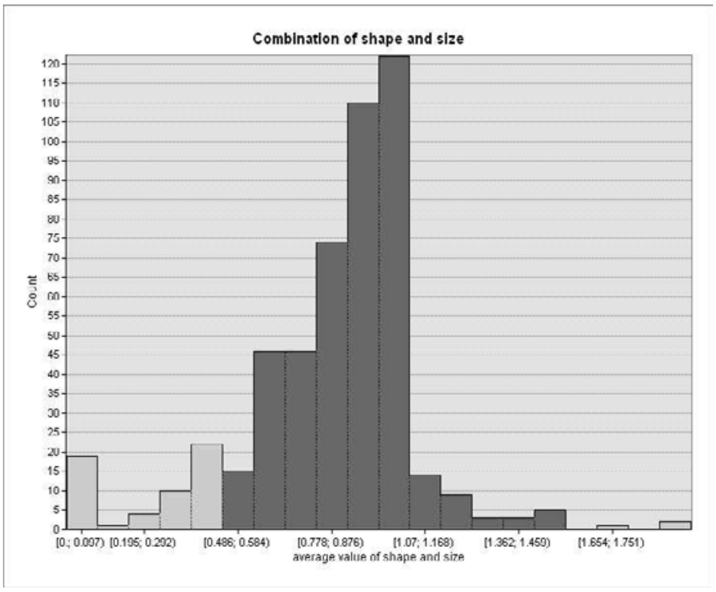
(a) (b)

Figure 5.9 (a) Digitized reference polygons; (b) Bounding box on reference polygons



(a) (b)

Figure 5.10 (a) Identified extracted polygons by using reference polygons; (b) Bounding box on identified polygons



Notes: The dark grey represents average values from 0.5 to 1.5; the light grey represents other values.

Figure 5.11 The histogram of combination of shape and size conditions

3. Accuracy based on the combination of "GC" and "BB"

After applying "BB" to the reference data, it was found that for some houses the area or the shape was changed considerably but for others this was not the case (Figure 5.13). This was caused by the different orientation of the houses. Therefore, the combination of "GC" method and "BB" method was applied.

In the "BB" part, reference data were selected on the condition that the ratio of the area of original reference data to the area of "BB" data is larger than a threshold between 0.5 and 0.9. The way to determine the threshold value is shown below. Then the extracted data were evaluated by following the same process as "BB" method. Meanwhile, the rest of the reference data were processed by using "GC" method. Then the final result was the combination of the two results from those two parts.

Determination of threshold:

There is a rectangle with a certain angle (α) of rotation to its "BB", the sides of the rectangle are "a" and "b" (Figure 5.12). The area of rectangle is: "$A_r = a \times b$"; according to the geometric relationship, the area of bounding box is:

"$A_B = (a\times\cos\alpha + b\times\sin\alpha) \times (a\times\sin\alpha + b\times\cos\alpha) = \frac{1}{2}\left(a^2 + b^2\right)\times \sin 2\alpha + a\times b$;"

Therefore, it can be seen that when α=45°, the area of bounding box reaches the largest; when α=45° and a=b, the ratio of A_r to A_B is 0.5; if α=45° and a≠b, the ratio of A_r to A_B is smaller than 0.5. In this case, the threshold starts from 0.5 up to 0.9; when the threshold is equal to 1, the rectangle doesn't have rotation. If the ratio of A_r to A_B of objects is above the threshold, then those objects used "BB" method; otherwise, the objects used "GC" method.

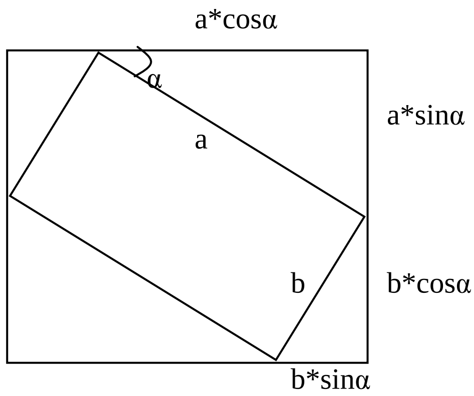


Figure 5.12 A rectangle with a rotation angle "α", the sides are "a" and "b"; and it is surrounded by its "BB", the sides of BB are calculated by using geometric principle

Through applying different thresholds from 0.5 to 0.9 with step 0.1, we obtained different correctness and completeness values. The tables of accuracy based on the different thresholds are shown in (Appendix I). Calculated overall accuracy values based on those two, relationship between those three values is shown in Figure 5.14. And according to the curves, it can be seen that the threshold equal to 0.7 is the optimal value.

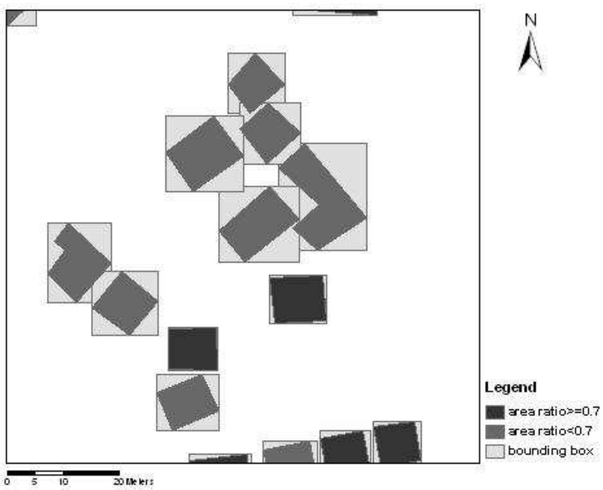


Figure 5.13 The area and the shape changed after applying "BB" to reference data

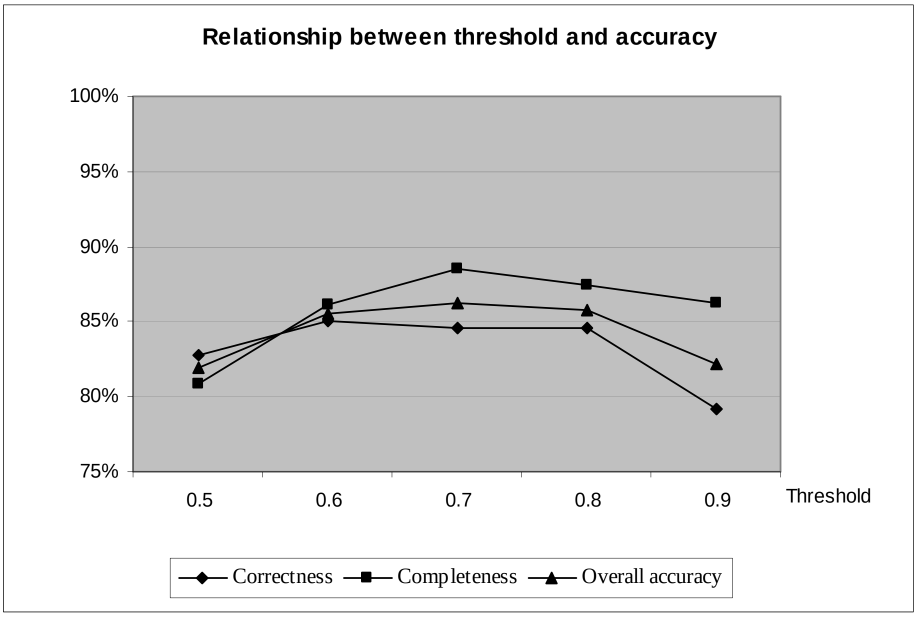
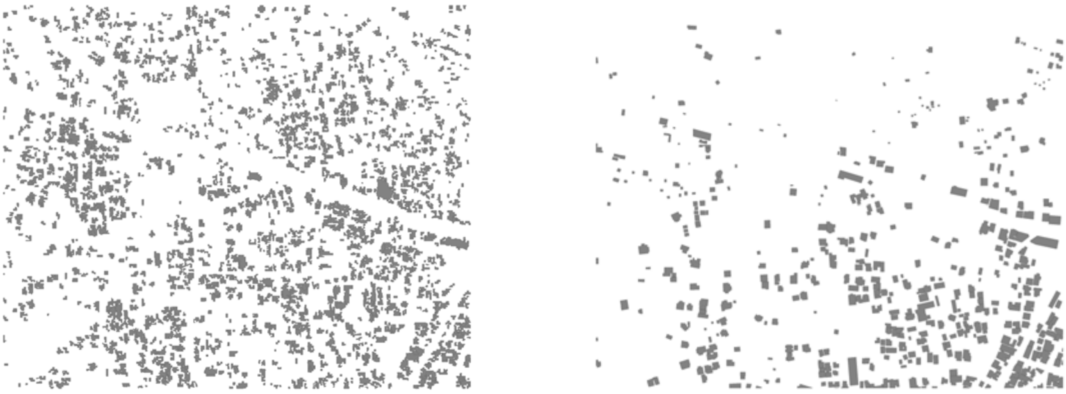


Figure 5.14 The relationship between the thresholds and the accuracy

5.4 Detection of new buildings

Building footprints were extracted using object-based classification as described in Section 5.2. These are all buildings that can be detected from KOMPSAT-2 imagery. We next compared extracted buildings with old orthophotos to separate those new buildings from old ones. The old buildings were digitized from orthophotos of 2005, it contains all buildings that still existed after the tsunami but before reconstruction projects started.

The polygons representing the old buildings of the whole test area, as mapped from orthophotos, are shown in Figure 5.15(b). These polygons were overlaid with the extracted objects (building footprints) from the KOMPSAT-2 data (Figure 5.15(a)), the area that overlapping polygons had in common was calculated and compared with the area of extracted objects. It was assumed that if the common area was larger than or equal to 50% then these extracted buildings were recorded as old buildings and separated, otherwise these would be new buildings. The final result contains new buildings and old buildings in the test area and is shown in Figure 5.16.



(a) (b)

Figure 5.15 (a) Extracted buildings of the test area from KOMPSAT-2 imagery; (b) Old building polygons of the test area from orthophotos of 2005.

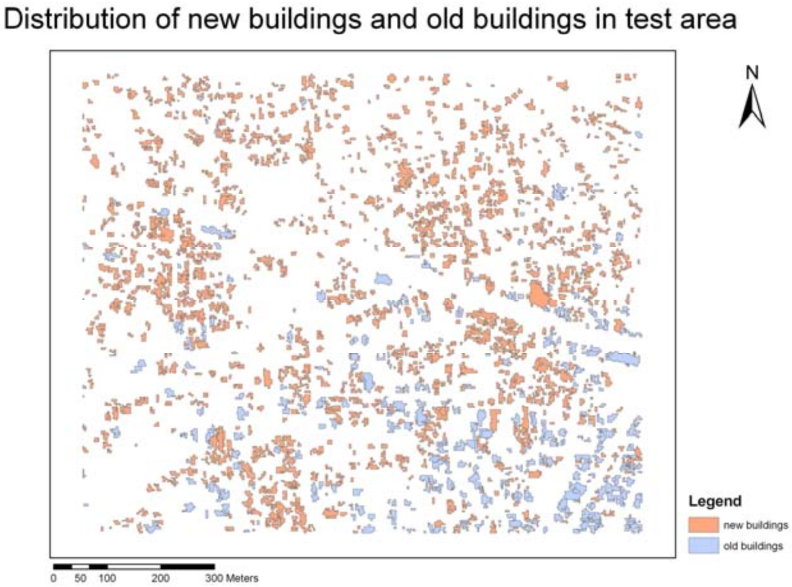


Figure 5.16 The distribution of new buildings and old buildings in the test area

6 Discussion

This study showed how the building footprints were extracted by using object-oriented classification approach from high resolution satellite imagery, how the different approaches were applied for evaluation of the accuracy and how the new buildings were detected by comparing with old data. This chapter will discuss the results and follow the same order as in chapter 5.

6.1 Building footprint extraction

In this research, an object-based classification approach was combined with decision rules and ancillary information. This combined method was applied for extraction of building footprint from high resolution satellite imagery. It started with a bottom-up segmentation which merged the pixels into objects. During segmentation different parameters were set for different purposes. For example, the NIR and the green layers in level one were used to segment vegetation areas then the blue and the red layers were used to segment the different colours of roofs; the scale parameter was not too large because the houses were small and many of them were individual buildings; the colour scale was more than shape scale because of the different colours of roofs and large vegetation areas; for the using of thematic layer, the road vector layer was not changed so much by tsunami and used as thematic data to assist segmentation; however, most of the cadastral databases such as residential areas were destroyed by tsunami and not available. The object-oriented classification strategy was based on decision tree. NDVI was used for separating vegetation area and residential area first, and then classification of buildings based on the colour of roofs; some of these building roofs had different spectral characteristics because of diverse materials, multi-facets and different colours. The texture information and the contextual information were used for separating those connected objects which were separated in reality. The merging function was applied for merging the objects which belong to the same class.

Some characteristics of the study area posed challenges during the classification and made it more difficult to extract building footprint. One situation is that building roofs are multi-facets (Figure 6. 2(d)), which means the different parts of the roof of the same building have different angles and reflectance; therefore it appeared that one roof was separated into multiple objects. During the object-based classification, classification based segmentation method and object merging rules were used to merge parts of roofs to entire roofs as much as possible. Another situation is that two or more independent buildings in reality are very near to each other, the gaps between them are quite narrow, even some of them are connected together (Figure 6. 2(a) and (b)). Therefore, it was hard to separate these kinds of roofs from classification even if the texture properties were applied.

6.2 Accuracy assessment

The research used different data and approaches to evaluate the accuracy of building footprint extraction result. Thus through these methods, different results of accuracy have been obtained. The reference data used in evaluation was based on visual interpretation of the image. It was assumed that the accuracy of the reference data was 100%.

Accuracy based on error matrix

As it was shown in Table 5.2 that 500 random points were generated for accuracy assessment. The producer's accuracy of "roof" class was 63.39% and "background" (non-roofs) was 99.74%. The user's accuracy were 98.61% for "roof" class and 90.42% for "background" class. This low producer's accuracy and high user's accuracy situation of "roof" class also happened in Zhan's study [26] and it was called "single-class" case. In Zhan's study, they selected 1000 random samples, and their producer's accuracy in "building" class was 78.8% which is higher than this study; but their user's accuracy was 74.3% and is lower than this case.

There are some problems by using traditional error matrix to evaluate accuracy in this case. The first problem is that the test area was covered by vegetation, water bodies and bare ground more than house area, therefore, most of random points were distributed on background area and a few of points were on building roofs (as in Table 5.2). The second problem is caused by the fact that there were only two classes, namely "roof" and "background" in the error matrix, and it is not so reliable to assess whole classification results only based on building and non-building classes. Because of those problems, the assessment of "background" class is mostly overestimated and "roof" class maybe underestimated. However, this accuracy assessment method is still making sense. Producer's accuracy in "roof" class is not very high, and high accuracy is in user's accuracy. It means that some houses were missed by using classification but most of those extracted houses were real houses.

Accuracy based on "GC"

The evaluation of extracted building footprint was based on the "GC" method which got better result and was making more sense. First result of accuracy was not so high (Table 5.3), so some small objects were removed to see whether they influenced accuracy or not. Indeed, according to applying the different thresholds of area of objects, the accuracy had increased after removing the small objects. From Figure 5.8, it can be seen clearly that the maximum value of overall accuracy was 82.82% when the threshold was $20\,m^2$. At this threshold value, the correctness was 80.99% and the completeness was 84.68%, which means that 84.68% of all reference data (visual interpretation) were detected and extracted, and 80.99% of those extracted buildings were real buildings.

Comparing this result with Zhan's study [26], they tested their method in two areas and both had higher accuracy. The difference between the method used in this study and the method used in their study is the way of determining the threshold. Our method was overlaying extracted building footprints with reference data by using their "GC" then applied different threshold on the size of extracted buildings; their method was overlaying two datasets and took the common area which were at least 50% of both two datasets and at least 10 m^2 .

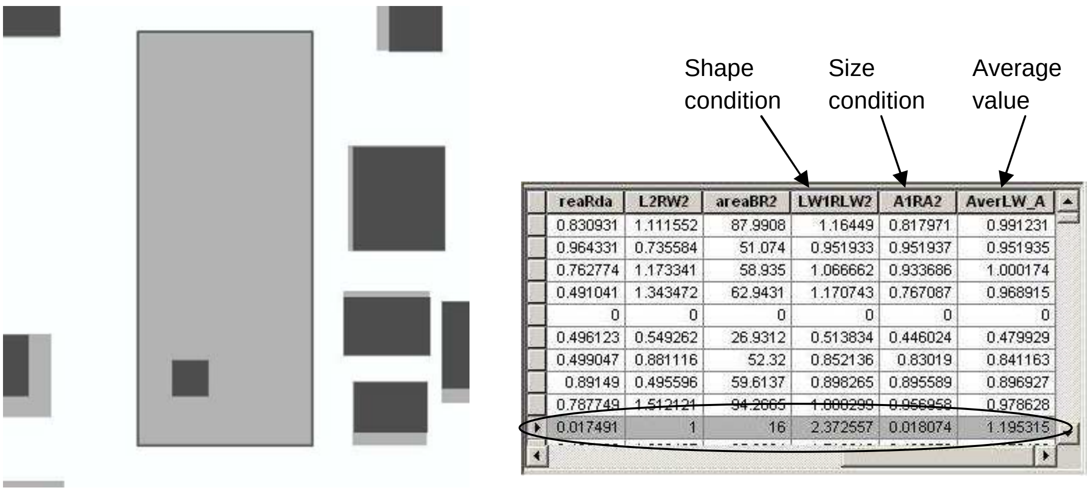
Comparing the "GC" method with traditional method (error matrix), it can be seen that the accuracy of building extraction had increased and the problem from the traditional method had been overcome by considering buildings as objects. However, it also can be observed from the processing that there are still some other problems remaining. First, this method did not consider shape differences between extracted objects and reference data. Second, the relationship between extracted objects and reference data was not one to one but "one to many" and "many to one". It means that one extracted object may represent two or more buildings in reality, and two or more extracted objects may belong to one building. This was caused by the complicated structure of building roofs and also some buildings were connected or very near to the neighbouring buildings.

Accuracy based on "BB"

To consider the shape and size conditions together in analysis, the extracted building footprint was simplified by applying "BB" method. The shape condition was based on the ratio of length and width of each bounding rectangle. Some corresponding objects from extracted result and reference data, which had the same length and width ratio but differ in size, therefore the size (area) condition also needed to be included. After integrating these two conditions for evaluation by taking the average value of them, the threshold of average value was selected between 0.5 and 1.5, in this case the correctness was 82.29% and completeness was 82.46%.

Comparing to the study of Shackelford [36], in their study, they used IKONOS panchromatic imagery, the opening and closing differential morphological profiles were applied for extraction of buildings and shadows. They defined the minimum length edge of bounding rectangle was longer than 5m as shape condition and the ratio of object's area to the approximating polygon was greater than 0.6 as size condition to identify buildings; they identified shadows based on the length of edges of objects. By combining these two results, they had 89.1% in correctness and 64.7% in completeness. The extracted buildings in our study were more completed than theirs, but we had low correctness.

Comparing with the "GC" method, this "BB" method included the shape and size as a whole condition to evaluate building extraction result. It had the same problem as in the "GC" method, concerning the "one to many" and "many to one" relationship between reference data and extracted result. Another problem in this "BB" method was caused by using average value to combine the shape and size conditions. There exist extreme situation as shown in Figure 6.1, the average value of shape and size is between 0.5 and 1.5, but the shape value is 2.37 and size value is 0.01. In this case, the object shouldn't be recorded as a building.



reaRda	L2RW2	areaBR2	LW1RLW2	A1RA2	AverLW_A
0.830931	1.111552	87.9908	1.16449	0.817971	0.991231
0.964331	0.735584	51.074	0.951933	0.951937	0.951935
0.762774	1.173341	58.935	1.066662	0.933686	1.000174
0.491041	1.343472	62.9431	1.170743	0.767087	0.968915
0	0	0	0	0	0
0.496123	0.549262	26.9312	0.513834	0.446024	0.479929
0.499047	0.881116	52.32	0.852136	0.83019	0.841163
0.89149	0.495596	59.6137	0.898265	0.895589	0.896927
0.787749	1.512121	94.2665	1.000299	0.956958	0.978628
0.017491	1	16	2.372557	0.018074	1.195315

Figure 6.1 Errors in the “BB” method. “Black” is the “BB” of identified extracted objects; “grey” is the “BB” of reference data. A small square in a big rectangle in the left picture, the error is from its large shape condition value and small size condition value (right side).

Accuracy based on the combination of “GC” and “BB”

After applying “BB” to reference data, it was found that the areas or the shape of some reference polygons were changed. Therefore, the combination of previous two methods was applied. By applying various thresholds on the ratio of area of original reference data to the bounding reference data, and the optimal threshold value was decided by the curve of overall accuracy. In this case the optimal threshold was 0.7, the correctness was 84.61% and the completeness was 88.51%. This combined result was higher than the results of both previous methods.

Comparing this combined method with previous two methods separately, it can be seen that the overall accuracy at the optimal threshold was higher than the previous methods (Table 5.4, Table 5.5 and Figure 5.14). In the “GC” method, the threshold was based on eliminating small areas of the extracted objects, and the accuracy was 82.82% at the optimal threshold. In the “BB” method, the threshold was selected on the shape of histogram and the accuracy was very similar to the “GC” method. In the combination method, the threshold was applied to the reference data and based on ratio of the original area to the “bounding” area. Because of the complicated situation of the test area, from this study, the optimal method for evaluating the building extraction result is the combination of “GC” and “BB” method.

6.3 Detection of new buildings

During the period between post-tsunami and pre-reconstruction projects, there were still houses which survived from the tsunami and some of which were (re-)built by people without organizing. These houses can be observed from orthophotos of 2005. The extracted buildings from KOMPSAT-2 imagery contained both new buildings and old buildings. Therefore it is needed to know which are the new buildings.

The method for detection of new buildings was the comparison of the common areas of extracted buildings with digitized buildings from orthophotos of 2005. The common areas were obtained by overlapping these two datasets. There are a few difficult situations in comparing the two. One is that some old buildings were removed and new buildings were built in the same location, and it is hard to tell whether these two are the same building or not; another is that some buildings were in "halfway" when the orthophotos were taken and completed when the satellite imagery was recorded. In this case, we assumed that if the common area was less than 50% of extracted building then it was considered as a new building.

The difference between this method and others is that the new buildings were detected based on objects but not pixels. Other studies detected changes based on post-classification comparison change detection technique which is the comparison of two raster data pixel by pixel and obtain the difference between them. The reason why we chose object-based method for detection of new buildings is that the whole study was based on objects but not pixels.

6.4 Methods analysis

The methods applied in this study were analyzed from strong and weak points and are shown in Table 6.1. The classification rules were limited by the test area. If it is changed to another area, the spectral information will change and these specific rules would fail. The classification result partly depends on the image. The spatial resolution of multispectral bands is coarse, therefore it is hard to separate individual buildings.

To consider the building extraction result from various aspects such as the location of geometric centres, the shape and size of buildings, different methods for evaluating the building extraction result were performed. These methods can be applied to other studies depending on the different situations and purposes. For example, if the users focus on the shape and size of buildings, then the "BB" method would be used; if some buildings have rotation angles to their "BB", then the combination of "GC" and "BB" would be a good choice.

There are some problems from field which posed a challenge in this study. Firstly, some buildings were extended by the owners, some buildings are very near to each other and even connected together (Figure 6. 2(a) and (b)), therefore it is hard to detect the gap between two buildings in this case from satellite imagery; secondly, building roofs are multi-facets (Figure 6. 2(d)), the spectral reflectance of the facets is quite different even from the same roof, so it is hard to merge one roof to one object. When it comes to combine with GIS techniques for further work, there are also some limitations. The main problem is the lack of available data. The tsunami destroyed everything on the ground and the cadastral databases of the disaster area were also destroyed. The new cadastral databases are being established and not available yet, therefore the ground truth is not enough for this study. For example, the reference data for object-based accuracy assessment was manual digitized from fused panchromatic and multi-spectral images.

Table 6.1 Analyzing methods from strong and weak points

Methods	Strong points	Weak points
Object-oriented classification	1. The rules were based on decision tree and had a clearly classes hierarchy; 2. Applying texture and contextual information to assist the classification;	1. The rules were limited in the specific area; 2. Some of the individual buildings were classified as object, one building was split into multiple objects;
Error matrix	1. Traditional method for accuracy assessment;	1. Can’t handle “single-class” situation;
“GC” method	1. Consider buildings as objects; 2. Solve “single-class” problem; 3. From the location point of view to assess buildings;	1. Can’t handle “1 to 1” relationship between extracted objects and reference data;
“BB” method	1. Consider the shape and size of buildings together; 2. Using average value between shape and size to complement one condition to the other;	1. The problem exists when there is a certain angle between buildings and their “BB”; 2. Can’t handle “1 to 1” relationship between extracted objects and reference data;
“GC” & “BB” method	1. Combine “GC” and “BB” methods, avoid the first problem in “BB” method;	1. Can’t handle “1 to 1” relationship between extracted objects and reference data;

(a) (b)

(c) (d)

Figure 6. 2 (a), (b) and (c) are pictures from field; (d) is from orthophotos of 2005.
(a) Houses are very near to each other; (b) Houses are connected together;
(c) Some new houses are not occupied; (d) Roofs are multi-facets.

7 Conclusions and recommendations

During the past few years, object-oriented image analysis has proven that it has great potential in feature extraction because it combines the spatial information and spectral information of the objects. Moreover, different techniques and different data sources have been applied in the previous studies to extract building footprint information which is needed for many applications. In this study, the applied method was to develop rules and extract buildings based on object-oriented classification; in addition, three different methods were applied to evaluate the building extraction result. Corresponding to these objectives, several research questions have been proposed and answered in this study.

7.1 Conclusions

(1) What kind of rules could be the optimal choice for classification in this study?

The structure of classification rules was based on the decision tree. In each level of the decision tree there were two complementary conditions to classify the previous classification result into two sub-classes. The rules combined spectral values with texture and contextual information. The applied spectral values were emphasized on the colour of roofs, for example red and blue. The texture information such as dissimilarity and contrast in the panchromatic band were used. Contextual attributes such as “relations to the sub-objects” and “relations to the neighbour objects” were used to assign the objects based on their relationships between each other.

(2) How successful is the extraction of buildings?

There were four methods to assess the accuracy of extracted buildings. The accuracy based on error matrix was 98.61% in user’s accuracy and 63.39% in producer’s accuracy. Because there are only two classes in the error matrix, it caused a “single-class” situation.

The evaluation approaches based on objects were “GC” method, “BB” method and the combination of those two. By applying different thresholds, the optimal accuracy in each method was achieved. The completeness of “GC” method (84.68%) was higher than in “BB” method (82.46%), and the correctness of “GC” method (80.99%) was lower than in “BB” method (82.29%). After the combination of “GC” and “BB”, the completeness was 88.51% and the correctness was 84.61%, both were higher than either of those two methods.

(3) What kind of method is the optimal choice to evaluate the building extraction result by using reference data?

In this case, it can be seen from the accuracy that the combination of “GC” and “BB” method is the optimal choice to evaluate the building extraction result. It combined the location of geometric centres and the shape and size of the objects together. This method was not based on only one condition but more conditions.

(4) How to detect new buildings from the extraction result?

The new buildings were detected and separated from the old ones by using object overlaying method. If the common area was less than 50% of extracted result, these objects were recorded as new buildings.

(5) Which level of detail of reconstruction project can be verified?

This study showed that it is possible and successful to extract building footprint from high resolution satellite imagery (KOMPSAT-2) by using an object-based classification method. The level of details that can be extracted is limited by the different kinds of real situations. First is that some individual buildings are closed to each other, and it is hard to separate them. The relationship between extracted objects and reference objects is not “one to one” but “one to many” and “many to one”. Second is that the lack of ancillary data. If other dataset such as residential parcels or elevation data was available, the result of building extraction would be more accurate.

7.2 Recommendations

Because of the lack of data and time constraints, there are some limitations in the research. Therefore, some recommendations are given for the further study.

From data aspect

(1) The updated cadastre of residential parcels should be helpful in segmentation process, and can give better and more accurate and meaningful image segments;

(2) The elevation dataset such as digital terrain model and digital surface model with high resolution should be useful during the classification process as it can help to separate buildings sharply from the neighbour objects which are without height;

(3) The higher spatial resolution image should be better to identify individual buildings;

(4) More and more accurate ground truth can help with accuracy assessment;

From method aspect

(1) The mathematic morphology technique can be applied during segmentation and gives more regular shape of the objects;

(2) This study was limited to a specific area; a further study can be done by generalization of the classification rules and make it suitable for a general area;

(3) This study was limited to separate individual buildings; a further method can be developed for “one to one” relationship between extracted buildings and reference data;

(4) A further research related with GIS techniques can be explored to link the building extraction result with ground data and data from other databases, e.g. services, and see whether those buildings are occupied or not.

From application aspect

This work has a direct application to post-disaster audit. Buildings in the disaster area were successfully detected by the method developed in this study, and could be identified as new (due to reconstruction projects) and old (not destroyed) by using a map of existing buildings. This method only identified presence of buildings, not their height, construction materials and occupancy, which are also important for auditors. Nevertheless, this presence information can be used for field planning and can form part of an integrated audit, along with other data source such as water and electric services.

References

1. Nasution, A., *Auditing of Tsunami Funds with Geophysical and Environmental Perspectives*. 2005.

2. *ReliefWeb*. [cited; Available from: http://www.reliefweb.int/rw/RWB.NSF/db900SID/EVIU-6AJJ3C?OpenDocument&rc=3&cc=tha.

3. Steinberg, F., *Housing reconstruction and rehabilitation in Aceh and Nias, Indonesia--Rebuilding lives.* Habitat International, 2007. **31**(1): p. 150-166.

4. BPK, *Terms of Reference for the project "Use of Geographical Information Systems for audit of disaster-related aid" for Badan Pemeriksa Keunangan RI (BPK), Strenthening Accountability for the Audit of Disaster Related Aid (SAADRA) Program.* 24 September 2007, Schedule A of contract number 04/BPK-RI/TF-05742/09/2007.

5. Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., and Heynen, M., *Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information.* ISPRS Journal of Photogrammetry and Remote Sensing, 2004. **58**(3-4): p. 239-258.

6. Marangoz, A.M., Oruc, M., and Buyuksalih, G., *Object-Oriented Image Analysis and Semantic Network For Extracting The Roads and Buildings From IKONOS Pan-sharpened Images*. 2003.

7. Hofmann, P., Strobl, J., Blaschke, T., and Kux, H., *Detecting informal settlements from Quickbird data in Rio De Janeiro using an object based approach*. 2006.

8. Mohamed Rafeek, M.T., *Extraction of building facade information from terrestrial image using segmentation based object - oriented image analysis*. 2007, ITC: Enschede. p. 104.

9. Fu, K.S. and Mui, J.K., *A Survey on Image Segmentation.* Pattern Recognition, 1981. **13**: p. 3-16.

10. Sezgin, M. and Sankur, B., *Survey over image thresholding techniques and quantitative performance evaluation.* In: Journal of electronic imaging, 13(2004)1, pp. 146-165, 2004.

11. Vosselman, M.G., *Semi-automatic Feature Extraction.* Paper presented at the course "Digital Photogrammetry" at the University of Bonn, 1998.

12. Kerle, N., Janssen, L.L.F., Huurneman, G.C., Bakker, W.H., Grabmaier, K.A., et al., *Principles of remote sensing (ITC Textbook).* Third edition ed. ITC Educational Textbook Series;2. 2004, Enschede: ITC. 250.

13. Lu, D. and Weng, Q., *A survey of image classification methods and techniques for improving classification performance.* International Journal of Remote Sensing, 2007. **28**(5): p. 823-870.

14. Michalak, W.Z., *GIS in land use change analysis: integration of remotely sensed data into GIS.* Applied Geography, 1993. **13**(1): p. 28-44.

15. Tso, B. and Mather, P.M., *Classification methods for remotely sensed data*. 2001, London and New York: Taylor & Francis. 332.

16. Mather, P.M., *Computer processing of remotely - sensed images : an introduction*. Third edition ed. 1999, Chichester etc.: Wiley & Sons. 324.

17. Zadeh, L.A., *Fuzzy sets.* Information and Control, 1965. **8**: p. 338-353.

18. Hansen, M., Dubayah, R., and Defries, R., *Classification trees: an alternative to traditional land cover classifiers.* International Journal of Remote Sensing, 1996. **17**(5): p. 1075-1081.

19. Zhang, Y., *Optimisation of building detection in satellite images by combining multispectral classification and texture filtering.* ISPRS Journal of Photogrammetry and Remote Sensing, 1999. **54**(1): p. 50-60.

20. Kiema, J.B.K., *Texture analysis and data fusion in the extraction of topographic objects from satellite imagery.* International Journal of Remote Sensing, 2002. **23**(4): p. 767-776.

21. Haralick, R.M., Shanmugam, K., and Dinstein, I.H., *Textural Features for Image Classification.* IEEE Transactions on Systems, Man, and Cybernetics, 1973. **SMC-3**(6): p. 610-621.

22. Haralick, R.M., *Statistical and Structural Approaches to Texture.* Proceedings of the IEEE, 1979. **67**(5): p. 786-804.

23. Mayer, H., *Automatic Object Extraction from Aerial Imagery-A Survey Focusing on Buildings.* Computer Vision and Image Understanding, 1999. **74**(2): p. pp. 138-149.

24. Baltsavias, E.P., *Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems.* ISPRS Journal of Photogrammetry and Remote Sensing, 2004. **58**(3-4): p. 129-151.

25. Congalton, R.G., *A review of assessing the accuracy of classifications of remotely sensed data.* Remote Sensing of Environment, 1991. **37**(1): p. 35-46.

26. Zhan, Q., Molenaar, M., Tempfli, K., and Shi, W., *Quality assessment for geo-spatial objects derived from remotely sensed data.* International Journal of Remote Sensing, 2005. **26**(14): p. 2953-2974.

27. Jin, X. and Davis, C.H., *Automated Building Extraction from High-Resolution Satellite Imagery in Urban Areas Using Structural, Contextual, and Spectral Information.* EURASIP Journal on Applied Signal Processing, 2005(14): p. 2196-2206.

28. Shamaoma, H., Kerle, N., and Alkema, D., *Extraction of flood - modelling related base - data from multi - source remote sensing imagery.* In: ISPRS 2006 : ISPRS mid-term symposium 2006 remote sensing : from pixels to processes, 8-11 May 2006, Enschede, the Netherlands. Enschede : ITC, 2006. 7 p., 2006.

29. San, D.K. and Turker, M., *Automatic Building Extraction from High Resolution Stereo Satellite Images.* ISPRS Commission VII, 2007.

30. Shufelt, J.A., *Performance Evaluation and Analysis of Monocular Building Extraction From Aerial Imagery*. IEEE Pattern Analysis And Machine Intelligence, 1999. **21**(4): p. p311-326.

31. Mason, S. and Baltsavias, E., *Image-Based Reconstruction of Informal Settlements*, in *Automatic Extraction of Man-made Objects from Aerial and Space Images (II)*. 1997: Basel,Switzerland. p. 97-108.

32. Lee, D.S., Shan, J., and Bethel, J.S., *Class-Guided Building Extraction from Ikonos Imagery*. Photogrammetric Engineering & Remote Sensing, 2003. **69**(2): p. 143-150.

33. Hurskainen, P. and Pellikka, P., *Change Detection of Informal Settlements Using Multi-temporal Aerial Photographs - The Case of VOI, SE-KENYA*, in *Department of Geography*. 2004, University of Helsinki.

34. Zhang, Y., *Detection of urban housing development by fusing multisensor satellite data and performing spatial feature post-classification.* International Journal of Remote Sensing, 2001. **22**(17): p. 3339-3355.

35. Wei, Y., Zhao, Z., and Song, J. *Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection*. in *IEEE International Geoscience and Remote Sensing Symposium, 20 - 24 september 2004, Anchorage, Alaska. on CD - ROM*. 2004: Institute of Electrical and Electronics Engineers (IEEE).

36. Shackelford, A.K., Davis, C.H., and Wang, X. *Automated 2-D Building Footprint Extraction from High-Resolution Satellite Multispectral Imagery*. in *IEEE International Geoscience and Remote Sensing Symposium, 20 - 24 september 2004, Anchorage, Alaska. on CD - ROM*. 2004: Institute of Electrical and Electronics Engineers (IEEE).

37. Lari, Z. and Ebadi, H. *Automatic Extraction of Building Features From High Resolution Satellite Images Using Artificial Neural Networks, ISPRS Commission VII, WG WG VII/2 & VII/7*. 2007.

38. Turker, M. and San, B.T., *Detection of collapsed buildings caused by the 1999 Izmit, Turkey earthquake through digital analysis of post-event aerial photographs.* International Journal of Remote Sensing, 2004. **25**(21): p. 4701-4714.

39. Bitelli, G., Camassi, R., Gusella, L., and Mognol, A., *Image Change Detection on Urban Area: The Earthquake Case*. 2001.

40. Haala, N. and Brenner, C., *Extraction of buildings and trees in urban environments.* ISPRS Journal of Photogrammetry and Remote Sensing, 1999. **54**(2-3): p. 130-137.

41. Rizatus, S., Ratih, D., Dimyati, Agus, K., and Wahyunto, *Tsunami effect in Nanggroe Aceh Darussalam and North Sumatra Provinces, Indonesia.* Asian Journal of Geoinformatics, 2005. **Vol. 5 (2)**.

42. Richards, J.A. and Xiuping, J., *Remote sensing digital image analysis : an introduction*. Fourth edition ed. 2006, Berlin etc.: Springer-Verlag. 439.

43. Soille, P., *Morphological image analysis : principles and applications*. 1999, Berlin etc.: Springer. 316.

44. Definiens, *eCognition Reference book 5*. 2006.

45. Aronoff, S., *Remote sensing for GIS managers*. 2005, Redlands: ESRI. 487.

46. Zhang, Q., Wang, J., Peng, X., Gong, P., and Shi, P., *Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data.* International Journal of Remote Sensing, 2002. **23**(15): p. 3057-3078.

47. Maktav, D. and Erbek, F.S., *Analysis of urban growth using multi-temporal satellite data in Istanbul, Turkey.* International Journal of Remote Sensing, 2005. **26**(4): p. 797-810.

48. Definiens, *eCognition User Guide 4*. 2004.

49. Shamaoma, H., *Extraction of flood risk - related base - data from multi source remote sensing imagery*. 2005, ITC: Enschede. p. 92.

50. Bijker, W. and Sanjaya, H., *Use of Geographical Information System for audit of disaster-related aid (Mid-term report), December, 2007*.

Appendix I

These tables show the different accuracy by applying various thresholds in the combination of the “GC” method and the “BB” method.

Ratio >= 0.5	**Extracted result**		**Reference data**	
	Bounding box	Geometric centre	Bounding box	Geometric centre
Recorded number of objects	558	12	393	8
			486	10
Total number of objects	689		496	
	Correctness: 82.73%		Completeness: 80.85%	
Overall accuracy	81.94%			

Ratio >= 0.6	**Extracted result**		**Reference data**	
	Bounding box	Geometric centre	Bounding box	Geometric centre
Recorded number of objects	521	65	383	44
			444	52
Total number of objects	689		496	
	Correctness: 85.05%		Completeness: 86.1%	
Overall accuracy	85.48%			

Ratio >= 0.7	**Extracted result**		**Reference data**	
	Bounding box	Geometric centre	Bounding box	Geometric centre
Recorded number of objects	468	115	343	96
			396	110
Total number of objects	689		496	
	Correctness: 84.61%		Completeness: 88.51%	
Overall accuracy	86.24%			

Ratio >= 0.8	**Extracted result**		**Reference data**	
	Bounding box	Geometric centre	Bounding box	Geometric centre
Recorded number of objects	379	204	268	166
			308	177
Total number of objects	689		496	
	Correctness: 84.61%		Completeness: 87.5%	
Overall accuracy	85.82%			

Ratio >= 0.9	**Extracted result**		**Reference data**	
	Bounding box	Geometric centre	Bounding box	Geometric centre
Recorded number of objects	220	326	157	271
			178	318
Total number of objects	689		496	
	Correctness: 79.24%		Completeness: 86.3%	
Overall accuracy	82.2%			