ANALYSIS OF PHOTOGRAMMETRICALLY-DERIVED DIGITAL SURFACE AND TERRAIN MODELS FOR BUILDING RECOGNITION

MTSHATSHA BANDILE

Submitted to the Department of Geomatics, University of Cape Town in fulfilment of the requirements of the degree of Master of Science in Engineering

Cape Town

November 1997

Acknowledgements

I would like to express my appreciation and indebtedness to my supervisor, Dr. Scott Mason for his valuable suggestions towards the successful accomplishment of this research work and for providing the research topic, and guidance.

A particular note of thanks is given to my family and friends for their love and financial support. I would also like to acknowledge the suggestions, comments and technical assistance received from Mr. D. Wilson and Mrs. S. Binedell, Julian Smith, Sidney Smith and Michael Haywood.

Li Jun, Ouma Yashon and Kwabena-Forkou Eric not only encouraged me but helped me by reading the draft of the manuscript, and offering very helpful critical evaluations. To them I am particularly grateful. Many thanks are also extended to all members of staff in the Department of Geomatics, University of Cape Town, for their valuable suggestions and comments on the draft of the manuscript.

Last but not least, I want to thank the following colleagues and friends, for their assistance and co-operation during my stay in Cape Town: Mufaro Chivasa, Danso Antwi Adjei, Osei Samuel and Yirenkyi Samuel. Since the list is too long to mention, it is hoped that those who have given me help in this manner will accept this anonymous recognition.

ABSTRACT

Buildings are one of the most frequently occurring man-made objects and in urban scenes their detection and reconstruction, e.g., in the form of three-dimensional CAD (computer aided design) models, is very important to many users such as architects, town planners and telecommunications and environmental engineers. This thesis examines the role of digital terrain and surface models in supporting this reconstruction process.

The thesis is structured into four main parts, namely, image matching to derive the data sets, building detection to delineate buildings from other man-made objects in DSM (digital surface model), DSM quality analysis to determine the reliability of the data, hydrological analysis to determine flood zones as an additional example of DTM application and finally conclusions and possible future outlook. Image matching was performed using an in-house image matching software in the Geomatics department. Off-the-shelf GIS functionality was used to tackle building detection, DSM quality analysis and hydrological analysis. A key feature of GIS functionality is the ability to exploit standard functions for the input/output, management, spatial analysis, editing and visualisation. It also aims at enhancing the accessibility of developed tools to end users.

Image matching is the most time consuming and complex of digital photogrammetric tasks. This is the process of locating corresponding points of interest in conjugate images for the purpose of three dimensional (3D) point position location in object space. The image matching algorithm used follows two distinct, yet linked algorithms, in order to minimise the need for extensive a priori knowledge of points to be matched. Firstly, the use of epipolar geometry of the conjugate images and cross correlation of the reference and search patches, in the reference and target images respectively is employed to locate the initial approximations of the corresponding search position for the point of interest, as well as providing initial estimates for the affine transformation parameters between conjugate image patches. These

approximations are then used in a least squares multi-image matching algorithm with imposed geometric constraints, in the form of collinearity equations.

Two data sets namely of an informal and a formal settlement were available for the study. Building detection in the formal settlement follows a vector based strategy. The DSM is first segmented into contour lines. Buildings as well as other man-made structures such as cars are expected to be closed contours with certain shape and area characteristics. Thus, closed polygons are isolated from the contour lines. As stated above, these closed polygons may result from other man-made objects other than buildings. This calls for the use of context rules such as area, elongation and perimeter to separate the buildings from other man-made objects and also the use of height statistics. The detected buildings are further verified as buildings using image data.

In the informal settlement case both raster and vector approaches are employed. The raster approach attempts to segment surface blobs in the DSM. The first step is to reduce the DSM to above surface objects. This is achieved by subtracting the DTM from the DSM. Because the difference data set is now approximately on a plane it can be thresholded to segment objects of a given height above the terrain. Buildings in the DSM are connected due to the limitation of the image matching algorithm. Subtracting the shadows from the thresholded data set results in segmentation and detection of the buildings. A filter is then used to achieve a fine segmentation of the results.

The vector approach in informal settlement case aims at detecting building without using a DTM. This approach uses the DSM to directly detected buildings. The DSM is transformed into contour lines. The next step is to extract closed polygons because buildings are expected to have closed boundaries. The buildings in the polygon coverage are merged together because of the close proximity of buildings and limitations of the image matching algorithm. Also the boundaries of individual buildings in a group are not closed. This means that it is impossible, in the informal settlement case, to extract the closed polygons pertaining to individual buildings.

Quality analysis of the DSMs is important to verify their usefulness. Two approaches to the analysis were employed namely statistical and visual inspection methods. Visual inspection makes use of histograms and scattergrams and overlaying the DSMs onto the ground truth data. One DSM was produced in a Wallis filtered image while the other in a Median filtered image. Wallis filtering results in a slightly superior DSM than Median filtering.

A DTM was used in hydrological analysis to identify flood zones in the informal settlement as this is one of the pressing concerns in the community. Areas likely to flood are those that lie lower than their surroundings. Drainage networks were derived to determine how water is flowing in the terrain as well as directions of flow.

TABLE OF CONTENTS

Glossary

Affine transformation: a spatial transformation that preserves the affine properties of embedded objects such as parallelism.

Bilinear interpolation: a resampling technique used for the geometric correction of images.

Bit: binary digit (0 or 1).

Bundle adjustment: the iterative computation of the relations between image co-ordinates and object co-ordinates without introducing model co-ordinates as an intermediate step.

Collinearity condition: the mathematical relationship between the image co-ordinates one image and the image co-ordinates of the corresponding image.

Digital data: data displayed, recorded or stored in binary notation.

Digital image filtering: spatial filtering to smooth, edge enhance or texturally enhance digital image data.

Digital surface model(DSM): is a geometric decription of the terrain surface and objects located on and above this surface.

Digital terrain model(DTM): is a geometric description of the terrain.

Feature extraction: finding and labelling of parts of a 2D image of scene that corresponds to objects in a scene.

Filtering: the removal of certain spectral or spatial frequencies to enhance features in the remaining image.

Geographic information system(GIS): a computer-based information system that enables capture, modelling, manipulation, retrieval, analysis and presentation of geographically referenced data.

Gradient operator: is a sharpening technique useful primarily as an enhancement tool for highlighting edges in an image

Grid: is a structure that allows the vertical and horizontal subdivision lines to be arbitrary points taking into account the distribution of the data points.

Histogram: the graphical display of a set of data which shows the frequency of occurrence of individual values.

Model: is an artificial construction in which parts of one domain (source domain) are represented by means of a structure preserving function in another domain (target domain).

Pixel: a picture element having both spatial and spectral aspects.

Polygon: the area enclosed by a simple closed polyline.

Polyline: a finite set of line segments (called edges) such that each edge end-point is shared by exactly two edges.

Raster data structure: the representation of a field as a regular grid of pixels.

Smoothing: the averaging of values or tones in adjacent image areas to produce more gradual transitions.

Vector data structure: the representation of an object-based model as a collection of nodes, arcs and polygons.

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

1. INTRODUCTION

Automatic methods for three dimensional (3D) reconstruction of man-made objects are an important issue to many users and providers of 3D city data, including urban planners, architects, and telecommunications and environmental engineers. Therefore the necessity to interpret, classify and quantitatively process aerial images in an automatic mode and to integrate the results into a GIS (geographic information system) is more urgent than ever. A large number and diversity of applications require 3D city models, for example: architectural and town planning, telecommunications planning, property management, utilities planning and management and reconstruction of past cityscapes. 3D city models store sufficient information about each object (buildings, utilities, etc.) such that it can be displayed in three dimensions, i.e. they are 3D CAD (computer aided design) models of the urban areas. A great advantage of 3D models is that they provide a representation of the physical form of the city. 3D city models as components of GIS provide for the integration of the vast and heterogeneous data required in city administration (e.g., social statistics, utilities, etc.) and for the co-ordination of the diverse bodies engaged in the administration and development of an urban area through the sharing and use of common databases.

Buildings are fundamental objects in urban landscapes. They form, therefore, a critical component of a 3D city model. Informal settlements are characterised by rapid, unstructured and unplanned expansion, poorly constructed buildings constructed often from scrap material, destruction of the local ecosystem through erosion, poor water quality and sanitation and severe social problems due to adverse living conditions and high unemployment rate. The improvement and development of informal settlements have been identified as basic needs to be met. Such development is best tackled by planning. Effective planning can only be made with accurate, relevant and up-to-date information.

To date, informal and formal settlement modelling has been carried out with conventional mapping techniques, primarily photogrammetrically using large format, analogue aerial photography. This is uneconomical over the relatively small, dense areas covered by squatter settlements and expensive to fly on a regular basis. Data acquisition is currently performed by manual, labour-intensive and hence slow methods. These limitations are compounded by the rapid growth of informal settlements. More rapid mapping can be achieved by developing techniques for (semi) automatic extraction of information from imagery, the use of such techniques will also reduce the level of expertise required (Mason *et al*, 1997).

A number of appearance characteristics of informal settlements, which bear relevance to automating shack detection, can be drawn. First, typical of nadir aerial imagery, primarily shack roofs, and not walls, can be detected and reconstructed. Second, shack detection and extraction is a difficult task for several reasons:

- Shack roofs are composed of numerous different materials often for a single structure. Uniform texture and radiometric response can therefore not be assumed across an individual roof or between neighbouring shacks.
- Small structures on the roof like small stones, bricks, and other objects like cars adjacent to the shacks cause further fragmentation.
- Shacks cast shadows and often there are objects such as cars on the ground next to the structures, which lead to changes in the contrast along the roof sides. Shadows and other surface markings on the roofs cause similar problems.
- Shacks generally cover small areas, typical size 4 by 4 m, hence the number of pixels they occupy is limited. They are generally single story (2 ~ 2.5m high) and often very densely packed.
- Shack roofs in RGB image show extremely varied colours, ranging from dark grey, to brown, light blue and the occasional white and red roof. In many instance, a single shack roof is composed of materials with different colours.

It is anticipated that shacks cannot be reliably detected and extracted using the commonly applied techniques to reconstruct buildings in formal urban areas due to

their properties(e.g., see Henricsson *et al.*, 1996). Most shacks in the Marconi Beam informal settlement in Cape Town, one of the sites studied in this thesis, have simple geometric descriptions; almost all are flat roofed and approximately 84% can be described by four-sided polygons. It is this characteristic which bears most promise for automating shack extraction (Mason *et al.*, 1997).

DSMs (digital surface models) can play a crucial role in the detection of buildings in both formal and informal settlements. Buildings appear as blobs in a DSM and can be detected using various strategies. In this work the detection of blobs in informal settlement is achieved by subtracting an available DTM from a DSM and thresholding the difference. The blobs in the difference data set are then segmented by subtracting shadows. In the formal settlement the DSM is used directly to extract the building blobs because the buildings are detached.

1.1 Objectives of the research

The objective of the research presented in this thesis is the development of methodologies using standard GIS functionality for automating the geo-spatial modelling of urban settlements and their effectiveness in building detection. Geo-spatial models comprise environmental data and relations that are referenced to a co-ordinate system. Two classes of settlements will be considered:

(a) Informal settlements

Provision of adequate housing, infrastructure and services to the underpreveleged population of South Africa, predominantly located in the so-called townships or squatter settlements, is of great importance. Primary attention in this research is placed on quantitatively modelling informal urban settlements, particularly buildings. Shack data is important for monitoring grouth at regional level, shack counting for electoral boundary determination, generation of GIS databases of infrastructure for service upgrading and for environmental quality assessment (Mason *et al*, 1997). This information should be cheap and readily available, hence the need to automate.

(b) Formal settlements

3D urban models are a prerequisite to a number of important applications. These include planning for mobile phone communications, town planning, architectural design, dispatching systems for police and fire departments, environmental simulations and studies of pollution and lighting.

1.2 Motivation for the study

Interest in the use of DSMs (digital surface models) to support building reconstruction has recently arisen for a number of reasons. First, algorithms for automated building reconstruction based, for example, on grouping of line features extracted from imagery have not yet proven reliable. DSMs may help to reduce the search space for building hypotheses and provide a coarse estimate of the general shape of the building which can guide line-based reconstruction. Second, some applications of 3D urban models, e.g. for placement of cellular phone network nodes, do not require precise building models. 3D objects extracted from DSMs may provide sufficient building modelling for these applications.

1.3 Background to the study

The first step in the use of DSMs is obviously generation of the DSM itself. A number of possibilities exist. The most common approach is using image matching (e.g. stereo correlation) on stereo-pairs of large scale aerial imagery. Airborne Laser Scanners are also expected to provide viable alternatives to support DSM generation by image based techniques. In both cases it is important that a sufficiently dense surface description be obtained, e.g., providing a raster with an interval in the range between 0.5 and 5m, depending on the size of the buildings to be detected.

One limitation of DSMs is that not only the topographic surface and the buildings are measured, but also other objects such as trees, bushes and cars etc. which, from the point of view of building reconstruction are regarded as outliers. All these objects

appear as "blobs" on the topography. Identification of such blobs is the next step towards building recognition, accompanied by blob extraction to delineate the related object's spatial extent. Clearly, a further step is needed wherein the object type is established, i.e. building or otherwise. To this end, context rules in the form of shape and dimensional constraints may be applied (e.g. on size, area, elongation). Further, analysis of image information associated with a blob can be employed. Baltsavias et al (1995) distinguish buildings from trees on the basis of the strength and orientation of edges in image areas of blobs. The histograms belonging to buildings will contain significant peaks approximately 90° apart for regularly shaped buildings. Histograms of more complex buildings contain a few additional peaks. Histograms for trees, on the other hand, are typically random with no observable peaks. Other additional classification cues include the number and length of long straight lines, texture, and color.

In our research we build on the ideas for DSM exploitation in 3D building reconstruction proposed by Baltsavias et al (1995). They suggest four methods for building detection. The first method is to extract edges from the DSM, i.e. at discontinuities in the surface. Edge detectors extract most of a building's outlines but they often do not deliver closed contours. Other structures with a much smaller height than buildings, such as road borders, will also be detected. A second approach is to smooth the peaks in the DSM using morphological operators. By differencing the smoothed and original DSMs, blobs are detected at the peak regions in the surface. Morphological operators are sensitive to the choice of structuring element size, particularly in dense urban areas. They also present problems when other DSM blobs are situated close to the buildings or when the terrain is steep and irregular. A third method is to subtract an existing DTM (digital terrain model) from the DSM leaving only surface blobs. In practice however, DTMs even if available, do not usually have sufficient density and accuracy.

The fourth alternative method of blob extraction from DSMs that has been proposed by Baltsavias et al (1996) is based on use of a contour description of the DSM. In the so-called multiple height bin (MBH) method, contours in the DSM heights are grouped into height ranges (bins) of a certain size. This results in segmentation of the DSM into

relatively few regions that are always closed and easy to extract. The method is applied hierarchically whereby large bins are used to detect possible buildings, and the small ones to verify and refine the coarse detection and separate buildings that are close to each other. The maximum and minimum bin sizes are determined from the known height accuracy of the DSM and the estimated minimum building height in the image.

Other authors have approached this same problem. The strategy used by Weidner and Förstner (1995) towards automatic building extraction is object-related i.e. they use a geometric model of the building. This is implemented in the form of specific structuring elements used for morphological filtering. In a first pass, minimum filtering - a special form of erosion – is carried out to remove smaller surface peaks in the DSM. A second step of maximum filtering - a special form of dilation – expands surface peaks in the original DSM. Subtraction of these two results ideally results in a set of detected blobs. As noted above, this method requires the correct choice of element which may, however, generally not be uniform for all buildings in a scene.

The strategy used by Wang and Schenk (1992) for blob detection is to generate a DEM (digital elevation model) and then transform it into a grey-value image, in which blobs appear as locally bright clusters. This image is then segmented to form contour lines with closed boundaries. Non-blob boundaries are eliminated using closure of boundaries and constraints on the heights of boundaries. Shape detectors are used to detect blunders. All contours which belong to one blob are grouped together. The resulting blobs are classified into contours on the topographic surface or above it. These methods are further examined in detail in chapter 3.

1.4 Research methodology

The general methodology used in this research has been as follows:

(1) The feasibility of separating buildings, modelled as 'blobs' in the DSM, from the natural terrain has been investigated.

(2) Where building-terrain separation from DSMs has been shown to be potentially feasible, a method for automating this separation has been investigated. Because a

'blob' in a DSM can have causes other than a building, e.g., trees, techniques for verification have been studied.

(3) The DSM produced by the in-house multi-photo image matching software in the Surveying department has been subjected to quality analysis.

(4) Hydrological analysis of the terrain has been carried out to extract topological data for GIS analysis, with the objective of determining possible flood areas.

1.5 Limitations of the research

The following are the limitations of this research:

(1) Only two data sets have been used, one for the informal settlement and one for the formal settlement.

(2) Building detection in formal settlements is fairly well documented while the detection of shacks in an informal settlement is a new concept. Thus the availability of related work in this field of study is challenging because of the lack of documented strategies for shack detection.

(3) Currently the standard GIS software is limited to 2.5D data processing.

1.6 Structure of thesis

In chapter 2 the process of deriving the data sets used in the research is described. This chapter starts with an overview of the different image matching techniques and then gives a brief account of the matching techniques adopted to derive the data set. Results of the matching are also given at the end of the chapter. Chapter 3 deals with the most pressing issue, that of building detection. Here two data sets are used to test the two different strategies for building detection outlined in the chapter. One data set represents a formal settlement while the other is an informal settlement. The results of the employed strategies are given in the chapter as well as the analysis of the results. Chapter 4 deals with the quality analysis of the data sets derived in chapter 2. The analysis is based on statistical procedures as well as visual inspection of the data and results are depicted. Chapter 5 considers a further application of DTM analysis in urban environments for the detection of flood zones and the results of such an analysis

are presented. The thesis ends by giving the conclusions from this study and possible future developments in the field of building detection.

CHAPTER 2

DSM GENERATION

2.1 Introduction

Photogrammetry has been used to generate maps and to compute highly accurate point positions in three dimensions for about a hundred years. Digital sensors have made it possible to use photogrammetry on a computer and apply its algorithms to digital imagery. Some techniques of digital photogrammetry include image matching and DSM generation, as well as digital orthophotography. Depending upon the features used to detect similarities, image matching can be classified into different techniques. The most popular technique is area-based matching where a grey-value matrix of one image is compared to a grey-value window of another image on a pixel by pixel basis. Other matching techniques are based on features which could be lines or characteristic points in the image. They must first be extracted before matching can begin. This technique is less accurate, but more robust since correct matches can be found over the whole image area without any approximation (Novak *et al*, 1992).

The matching of points can take place in object space or in image space. Collinearity equations, which form the basis of all photogrammetry point positioning methods, give the relationship between object and image space. Any matched points on the image can immediately be projected to the ground. Once on the ground (object space) the three dimensional points can be used to generate a DSM stored as a matrix of grey-values. Good DSM interpolation techniques allow us to approximate the surface smooth functions by using reference points (matched points). It is important to smooth reference points derived automatically in order to eliminate wrong matches and reduce noise. The surface point represented in the DSM pixel can be projected into the original image using the perspective relationship between the ground and the image, its location in the original image can be identified and its corresponding grey-value can be extracted by resampling techniques to derive a digital orthophoto of the whole image. These data can then be integrated into a GIS together with the raster data.

2.2 DSM generation strategy

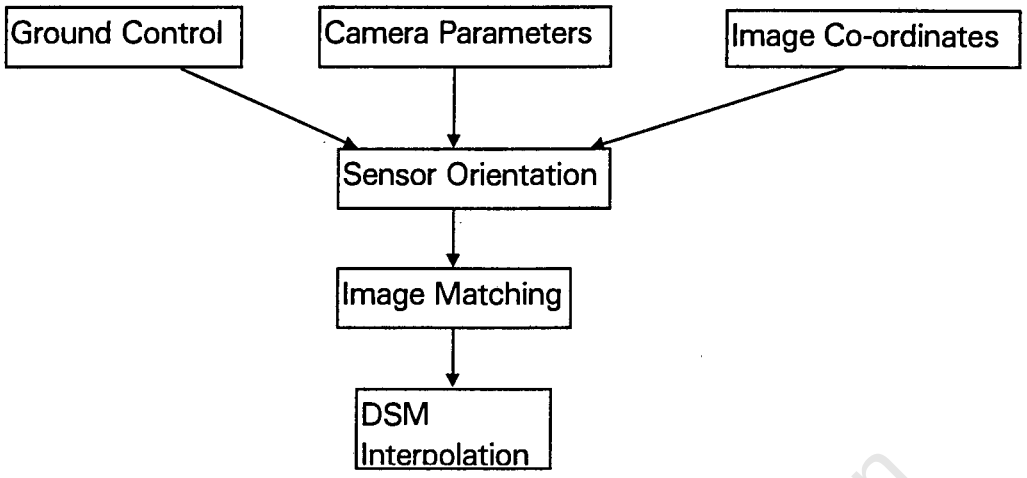


Fig. 2.1 Strategy for DSM generation

The first step in DSM generation is the preparation of data files needed for the sensor orientation and as approximations for image matching as outlined in Fig. 2.1. These data files include the ground control points, image co-ordinates and camera calibration information. The operator identifies and measures image co-ordinates to within a pixel or by using a resampling technique, even to a quarter of a pixel (Novak *et al*, 1992). These measurements are done manually on the screen, but some points such as fiducial marks are automatically measured. Ground control points are identified and measured on an existing map, otherwise geodetic or terrestrial survey methods are used, as well GPS(global positioning systems) . The sensor orientation parameters are computed in a bundle adjustment. The bundle adjustment is based on collinearity equations and solves for the exterior orientation parameters of the photograph, densifies ground points and estimate the parameters of a model of the sensor distortions. Sensor orientation parameters are important for three dimensional positioning points to create a random DSM (Novak *et al,* 1992). The image matching technique employed follows the area-based approach. Once the sets of corresponding points are available and have been edited, an intersection is computed to project the image co-ordinates into object space using the known orientation parameters. In object space they form a random DSM of reference points to be used to interpolate a grid DSM. The quality of the DSM is highly dependent on the success of the image matching and thus it will be discussed further below.

2.3 Overview of Image Matching techniques

2.3.1 Introduction

A variety of matching techniques have been developed over the last few decades. Unfortunately there is no unifying theoretical background behind all these techniques. In photogrammetry and remote sensing, matching is defined as the establishment of the correspondence between various data sets. Digital image matching automatically establishes the correspondence between primitives extracted from two or more digital images depicting at least partially the same scene.

Using image matching we try to reconstruct the three-dimensional object surface from two-dimensional projections. Information is lost during this projection. This is most evident in the case of occlusions. Image matching is therefore an ill-posed problem. A problem is ill-posed if there is no guarantee that a solution exists, is unique, and is stable with respect to small variations in the input data. Image matching is ill-posed because, for a given point in one image, a corresponding point may not exist due to occlusions, there may be more than one possible match due to repetitive patterns or a semi-transparent object surface, and a solution may be unstable with respect to noise due to poor texture (Heipke 1996).

An ill-posed problem can be converted into a well-posed problem by introducing additional information about the problem. Several assumptions usually hold true when dealing with photogrammetric imagery (Heipke 1996):

- The grey values of the various images have been acquired using one and the same or at least similar spectral bands.
- The illumination and possible atmospheric effects are constant throughout the time interval for image acquisition.
- The scene depicted in the images is rigid, i.e. it is not deformable.
- The object surface is piecewise smooth.
- The object surface is opaque.
- Initial values such as the approximate overlap between image or an average object height are known.

Most matching algorithms contain a combination of assumptions about the depicted scene. Rather than describing these algorithms as a whole it is more appropriate to decompose them into smaller modules. The questions to be answered are :

- Which primitives are selected for matching?
- Which models are used for defining the geometric and radiometric mapping between primitives of various images?
- How is the similarity between primitives from different images measured, how is the optimal match computed?
- Which strategy is employed in order to control the matching algorithm?

2.3.2 Matching primitives

The primitives fall into two broad categories : either windows composed of grey values or features extracted in each image a priori are used in the matching. The resulting algorithms are usually called area-based matching and feature-based matching respectively. In both cases there is a choice between local and global support for the primitives.

In area-based matching grey values in a small window serve as matching primitives. The grey values are regarded as quantised samples of the continuous brightness function in image space. The windows can be extracted very fast. Area-based matching has a high accuracy potential in well-textured image regions. Area-based matching is sensitive to the grey value changes in radiometry, e.g. due to illumination changes, the large search space for matching, and the large data volume to be handled (Heipke 1996).

In feature-based matching features are extracted in each image individually prior to matching them. Local features are points, edges and small regions. Global features comprise polygons and more complex descriptions of the image content. Features are more abstract descriptions of the image content. They are more invariant with respect to geometric and radiometric influences than the grey value windows.

2.3.3 Models for mapping primitives

Mapping between primitives of various images is defined by two models : a sensor model and a model for the object surface. Two-dimensional transformations from one image the next such as two-dimensional translation or an affine transformation implicitly contain a combination of these two models. They are rough approximations of the situation during image acquisition and should only be used if the selected matching primitives have local support. If the selected primitives have global support the mapping between images should be modelled more rigorously. There is an advantage if the mapping between the primitives is formulated in terms of object space parameters which are common for more than two images: multiple images can be matched simultaneously (Heipke 1996).

2.3.4 Similarity and optimisation procedures

For area-based matching the similarity between grey value windows is defined as a function of the differences between the corresponding grey values. This function can be the covariance or the cross correlation coefficient between the windows, the sum of the absolute differences between corresponding pixels, or the sum of the squares of the differences. Similarity measures for feature-based matching are more complicated as they must be based on attributes of the features. The optimisation procedure depends on the choice of the matching primitives. In a local area-based matching an exhaustive search can be carried out as is the case in cross correlation. Gradient-based iterative schemes such as ordinary or robust least squares adjustment could be used. In global area-based matching the optimisation procedure, the generation of three-dimensional information and the estimation of parameters describing the object surface are integrated into one model. In local feature-based matching for each given feature in one image a small search area is defined in the other image(s) using the selected mapping transformation. An exhaustive search is carried out in this search area. Blunders are detected through global consistency checks similar to local area-based matching. Alternative schemes for global consistency include dynamic programming and relaxation labelling etc (Heipke 1996).

2.3.5 The matching strategy

In the matching strategy the individual steps carried out within the algorithm are determined. This includes the input of prior information from a human operator, and the presentation of results for final visual verification. Hierarchical methods are used in many matching algorithms in order to reduce the ambiguity problem. They are employed from coarse to fine, and results achieved on one resolution are considered as approximations for the next finer level. Also the blunder rate for individually matched points can be rather high. Effective blunder detection is only possible if there is a large redundancy in the system. Another issue related to redundancy is that of multi-image matching (more than two images). In a conventional photogrammetric block with 60% end overlap and 20% side overlap only 24% of on image in the interior of the block is covered by two images, and the same area is covered by six images. Thus, multi-image matching is advantageous for DSM /DTM generation without having to acquire more images.

2.4 Multiphoto Geometrically Constrained Matching

In this work we use the Multiphoto Geometrically Constrained Matching to generate the 3D points used to derive a DSM. Multiphoto Geometrically Constrained Matching is based on the least squares matching which is an area-based matching method. It establishes a fit between small image patches by an affine geometric and a two-parameter linear radiometric transformation. The least squares matching is considerable improved by two new elements :

(1) exploitation of a priori known geometric information to constrain the solution and,

(2) simultaneous use of any number of images.

These elements reduce the search space and improve the accuracy, success rate and reliability of matching (Baltsavias 1991).

Observation equations consist of equations formulating the grey level matching and the geometric constraints. These two parts are related to each other through common known parameters. The geometric constraints are the collinearity conditions extended by additional parameters modelling systematic errors and assuming that interior and exterior orientation of the sensors are known. The radiometric parameters are can be included in the mathematical model but are applied during the

iterations. Also a Wallis and/or Median filter is used for radiometric equalisation of the images, as an alternative, before matching.

Multiphoto geometrically constrained matching is based on linear least squares estimation. The model is non-linear and must be linearised. The approximations are derived by a pyramid-based approach (coarse to fine). The image pyramid approach is an efficient way to derive approximate values, and increases the convergence radius, convergence rate and computational speed. The strategy consists of choosing good match points in the reference image and matching them in all pyramid levels. Thus the multiphoto geometrically constrained matching is a combination of area-based and feature-based, especially edge-based, matching (Baltsavias 1991).

2.5 Experiments and results

A pair of overlapping digital images, acquired with Kodak DCS 460 still video CCD camera, of a portion of the Marconi Beam informal settlement were used in this experiment. The bundle adjustment was applied to solve for the sensor orientations of the images. After the sensor orientations the digital images were processed to remove noise and to smooth them. Two filters were used to achieve this goal, namely the Wallis and Median filters. Filters that decrease noise, thus removing high frequencies, lead to an increase of the convergence radius and rate, an increase of the correlation coefficient and a decrease of the estimated σ_o. The use of smoothing is particularly beneficial when the approximations are poor, by cutting the number of iterations and processing time by about a third (Baltsavias 1991).

The Wallis filter forces the mean and especially the contrast of an image to fit some given values. This transformation is done locally and thus assures that different parts of the processed image have similar contrast. These characteristics are particularly useful because there are very often global grey level differences between images and, locally within each image, big differences in contrast (Baltsavias 1991). The median filter is a smoothing method that reduces blurring of edges. It reduces the current point in the image by the median of the brightnesses in its neighbourhood. This median of the brightnesses in the neighbourhood is not affected by individual noise spikes and so the median filter eliminates impulse noise well and does not blur edges much. The disadvantage of the median filter in a rectangular neighbourhood is that it damages thin lines and sharp corners in the image (Sonka *et al* 1993).

The Wallis enhanced image has an area of 12070 square meters while the Median enhanced image has an area of 11680 square meters. Both images have a density of 0.25 meters and the method of interpolation is the linear interpolation. The number of points used to generate the DSMs is 11608 for the Wallis and 14234 for the Median. Digital surface models of the two smoothed images were generated by the Geometric Constrained Matching described above. The output 3D cloud from the image matching was modelled using Arc/Info GIS. The Arc/Info **generate** command was used to generate a point coverage. The command is as follows :

Createtin ***name of tin***

generate in_file point

The command takes a file of co-ordinate points in x,y,z format and generates a point coverage. The point coverage is converted into a grid using the Arc/Info **tinlattice** command. This command allows one to specify the type of interpolation, which linear in this case. The command is as follows:

tinlattice in_tin out_lattice interpolation_method

Continuous surfaces of the digital surface models were generated using GRID as shown in Fig.2.2 and Fig.2.3. The bright parts in the images indicate high areas while the dark parts indicate low areas. Figure 2.4 represents the absolute differences between the Wallis and Median enhanced images while figure 2.5 shows the distribution of the differences.

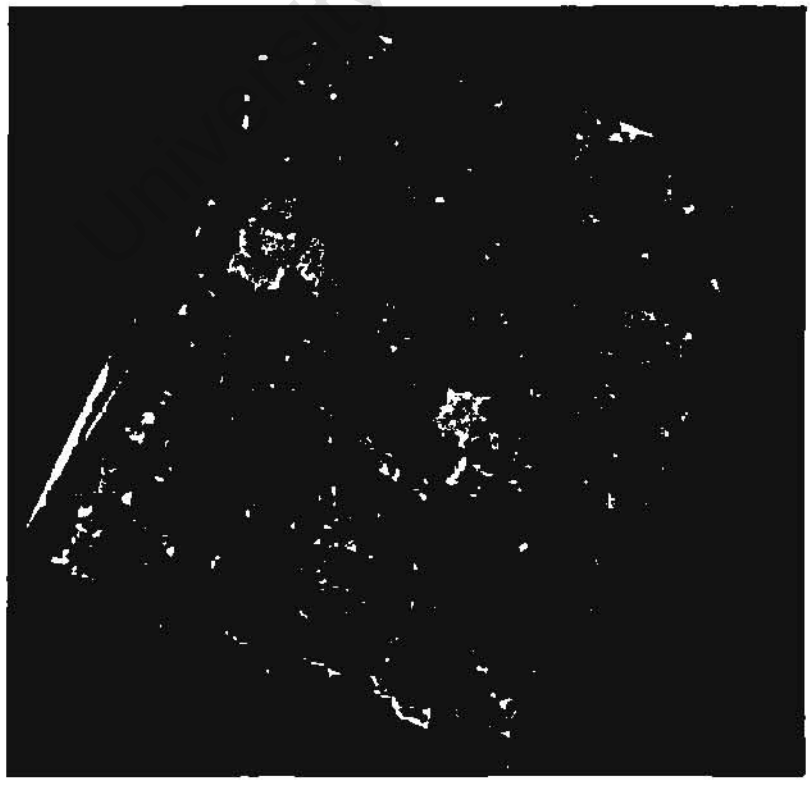


Fig. 2.2 Wallis filtered DSM

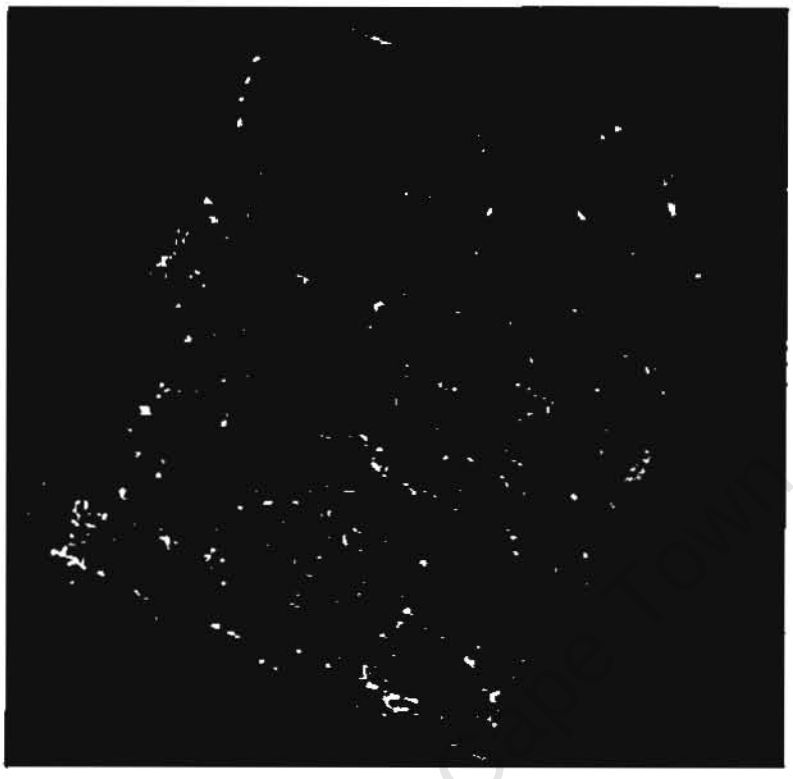


Fig. 2.3 Median filtered DSM

Fig. 2.4 Absolute difference of Wallis - Median DSM

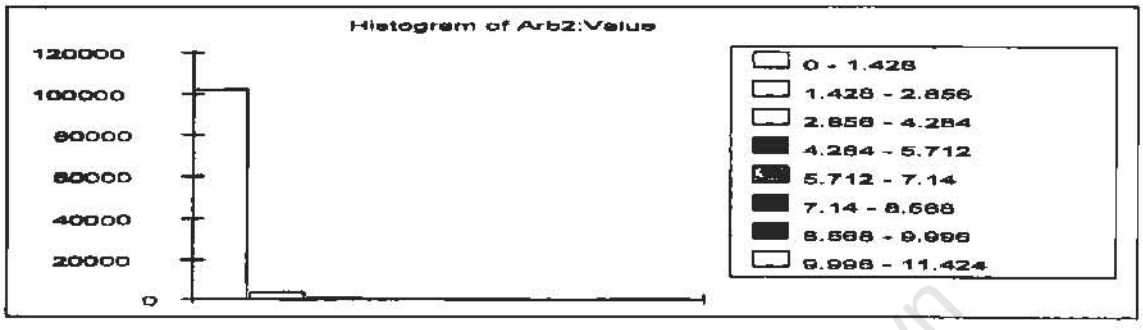


Fig. 2.5 Histogram of Absolute differences

Figure 2.5 is a histogram showing the absolute differences between the Wallis filtered DSM and the Median filtered DSM. In figure 2.4 the bright parts have high differences while the dark parts have low differences. As can be observed in figure 2.5 most absolute differences are in the region of 0 to 1.5 meters and correspond to the dark parts of figure 2.4. The bright parts in figure 2.4 are areas where the two DSM have significant differences but there are few of them.

2.6 Chapter Summary

This chapter described the strategy employed in DSM generation. The major focus is then given to the image matching process because of its complexity. The overview of the different assumptions employed in image matching is presented. Geometrically Constrained Multi-photo Matching techniques are described in brief. The Geometrically Constrained Multi-photo Matching algorithms have been applied to the Wallis and Median filtered images. The results are analysed for quality in chapter 4.

The strategy for DSM generation has four stages. The first stage is preparation data files such as ground control, camera parameters and image co-ordinates. These data files are used in the sensor orientation computed by the bundle adjustment. The bundle adjustment is based on collinearity equations. The image matching employed in this work was the Multiphoto Geometrically Constrained Image Matching which is a area-based matching technique. Then the cloud of three-dimensional co-ordinates are used to generate a DSM surface.

CHAPTER 3

Building detection

3.1 Introduction

The reconstruction of buildings and other man-made objects in 3D is currently a very active research area and an issue of high importance to many users of Geographic Information Systems(GIS), including urban planners, architects, and telecommunication and environmental engineers. Manual 3D processing of aerial images is time consuming and requires the expertise of highly qualified personnel and expensive instruments. The purpose in DSM generation is the derivation of a complete 3D model of the visible surface with the highest possible accuracy and taking into account terrain discontinuities (Baltsavias *et al*, 1995). The DSM in this work is generated by image matching and modelled in ARC/Info GIS. The available city plans contain 2D information about the ground plan of buildings. 3D information extraction requires a great deal of automation. The classical photogrammetric techniques are expensive and time consuming due to the higher amount of information required. State-of-the-art automatic algorithms based on stereo imagery are not able to distinguish between the terrain surface and objects on and above this surface e.g. man-made objects such as buildings or natural features such as trees. Automatic interpretation of digital data requires explicit modelling of the scene to be extracted. The amount of modelling depends on the complexity of the scene and the type of sensor data. Automation requires the derivation of 3D co-ordinates using matching techniques. Computational stereo is difficult to achieve in urban areas because of the presence of occlusions and the need to represent vertical surface structures. On the other hand laser scanners are starting to provide alternatives to support DSM generation by image based techniques. Unfortunately they leave unsolved the same problem of vertical walls being inaccessible; their detection is essential for deriving city models.

3.2 Uses of digital terrain and surface models

Aerial imagery can be used to derive both DTMs and DSMs. DTMs must be manually measured in urban areas. This is because state-of-the-art automatic algorithms based

on stereo imagery and image matching for DTM generation are not able to distinguish between the terrain surface and objects on and above this surface. Their output is a 2.5D surface model as opposed to the desired terrain model. However, the DSM offers a number of possible uses :

- Detection of buildings by filtering out 'blobs' in the DSM. The detection can be expected to perform well for detached buildings although the detection of individual buildings decreases with an increase in building density. Analysis of the image content associated with the blob region can be used to classify buildings over other non-terrain objects, e.g. trees.
- DSMs can be used to support 3D feature matching by providing approximations and reducing the number of candidate matches.
- DSMs provide information which allows the inference of 3D object hypothesis in model-based reconstruction of buildings. When the DSM is fine-grained major building features such as roof ridge lines can be discerned.
- The removal of non-terrain DSM blobs from the terrain DSM blobs and interpolation of the terrain can be used to generate a DTM.
- DSMs can be used to generate ortho-images and ortho-rectified stereo-pairs. The ortho-rectified stereo-pairs can in turn be used to detect DSM errors.

DTMs have many uses. Among the most important are the following:

- Storage of elevation data.
- Solving cut-and fill problems in road design and other civil and military engineering projects.
- 3D display of landforms for landscape design and planning.
- Planning of routes for roads.
- Planning for the location of dams.
- Computing slope maps, aspect maps and assist geomorphological studies.

3.3 Strategies used in building detection

Buildings are the most predominantly and frequently occurring 3D man-made objects in high resolution stereo images of urban areas. Their reconstruction requires many components such as image processing, matching, geometric processing and reasoning

as well as object modelling. The main features of the strategy employed by Baltsavias *et al* (1995) are the following :

- Automatic generation of a DSM exploiting all overlapping images.
- Automatic generation of coarse buildings for feature matching and model-based building reconstruction exploiting the DSM and the extracted 2D features.
- Multi-image matching of 2D image features to derive 3D information.
- Structuring of the 3D linear features into roofs by generic object models.

Weidner and Förstner (1995) present a strategy consisting of three steps:

- Automatic generation of a high resolution DSM.
- Detection of buildings in DSM and,
- Reconstruction of each detected building.

In the detection of buildings attention is focused on areas where buildings can be expected. An approximation of the topographic surface is first computed using mathematical morphology. The difference between the measured DSM and the approximate topographic surface gives information about the buildings. Thresholding the difference data set results in the detection of buildings.

Schenk and Wang (1992) use the following strategy for building detection :

- Automatic generation of a DSM
- Transforming the DSM into a grey-value image
- Detection of buildings
- Classification of the detected buildings.

3.4 Literature review

Baltsavias *et al* (1995) developed three schemes for building detection from DSMs. Scheme 1 is based on applying morphological operation 'opening' to smooth the DSM. The difference between the smoothed DSM and the measured DSM results in the extraction of smoothed peaks. The peaks from the difference data set are segmented by thresholding. The threshold can be chosen according to the expected building height range of the scene. Shape descriptors and context rules are applied to judge whether or not an individual component is hypothesised as a building. This scheme is sensitive to the choice of structuring element size, particularly in dense

urban areas, and has problems when other DSM blobs are situated close to the buildings or when the terrain is irregular.

Weidner and Förstner (1995) proposed a similar approach. As stated before the difference between the measured DSM and the approximation of the topographic surface gives information about the buildings. The approximate topographic surface represented in the DSM is computed using matbematical morphology. The first step is minimum filtering with the structural element a square window. The minimum filtering is a special erosion, in terms of mathematical morphology, where the structural element has constant z values. This step is followed by maximum filtering with the structural element which is a special dilation. Both steps perform an opening on the set of heights to deliver an approximate topographic surface. If the topographic surface converges towards the DSM in non building parts, the difference set, measured DSM minus the approximate topographic surface, consists of buildings approximately put on a plane. Thresholding the difference data set is necessary because the buildings are put on a reference surface which is approximately a plane. The threshold can be chosen to be the expected height of vertical walls in the segmentation stage.

Scheme 2 is based on segmenting the DSM with a 3D edge detector. Edge detectors extract most of the buildings outlines but do not deliver closed contours. Shape detectors and context rules still need to be used to hypothesise individual components as buildings but the choice of structuring element is avoided. However, other structures with a much smaller height than buildings, such as road borders, are also detected.

Schenk and Wang (1992) proposed a similar strategy to scheme 2. The generated DSM is first transformed into a grey-value grid. After the transformation blobs show up as bright clusters on the grey-value grid. This grid is an 8-bit representation of the DSM and therefore there is loss of data. In order to find all blobs the grey-value grid is segmented to form contour lines. The interval used in the segmentation should be smaller than the lowest height of the blobs in a given scene. In the contour image the blobs are characterised as closed boundaries. All non-blob boundaries have to be

eliminated. This is achieved by using generic properties such as closure and length properties. A boundary for a blob is always closed and it should not be too long or too short. Also redundant boundaries are eliminated by choosing the outermost boundary.

Scheme 3 is based on the subtraction of an existing DTM from the DSM. The DTM must have sufficient density and accuracy to deliver good results which is not the case in most instances such as informal settlements because they have not been mapped accuratetly. Shape descriptors and context rules still need to be applied to hypothesize individual components as buildings. An alternative method consists of grouping the DSM heights into consecutive bins (height ranges) of a certain size. This corresponds to cutting equidistant slices through the DSM. The resulting DSM is segmented into relatively few regions, that are always closed and easy to extract. This method is applied hierarchically whereby the coarse bins are used to detect possible buildings whilst the finer bins verify the coarse detection. The fine bins also provide information for an approximate building model and separate buildings close to each other.

3.5 Test data sets

3.5.1 Avenches data

A data set from Avenches (Switzerland) was acquired for use in the Avenches project (Mason *et al.* 1993). The data set consists of a residential and an industrial scene with the following characteristics : 1:5000 image scale, near vertical aerial photography, four-way image overlap, colour imagery, geometrically accurate film scanning with 15 μm per pixel size, precise sensor orientation, and accurate ground truth including a DTM and buildings and a manually measured CAD model of the houses (for more details see Mason *et al*, 1994). A DSM of the formal residential area if shown in Fig. 3.1.

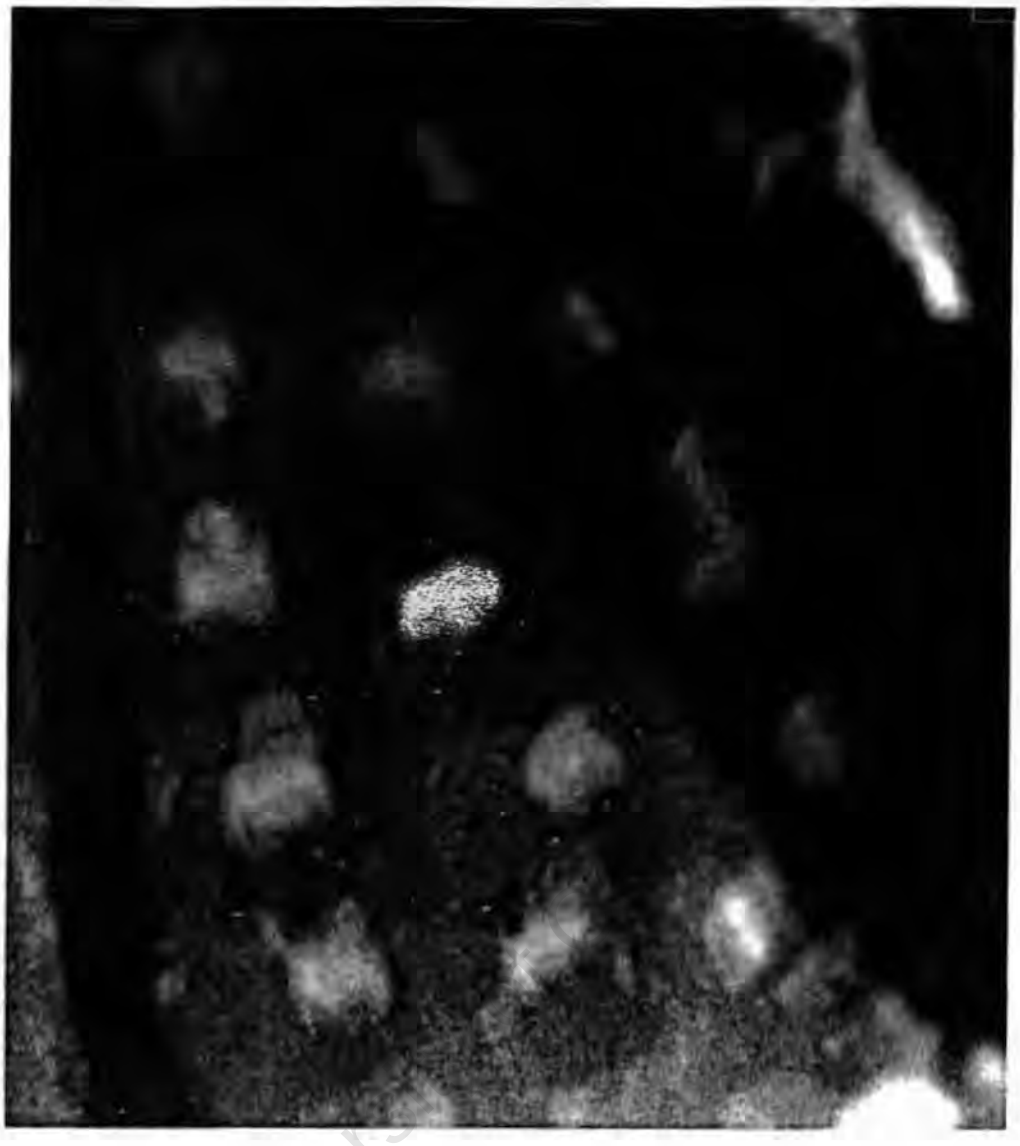


Fig. 3.1 DSM of Avenches settlement

The bright parts in figure 3.1 represent high areas and the dark parts correspond to low areas.

3.5.2 Marconi Beam data

Informal settlements can take on many different forms subject to, in varying degrees, topography, climate, available building materials, political and social influences, etc. For these investigations a data set was chosen to reflect: (1) broadly typical informal settlement conditions for the Cape Town region; and (2) low-cost image sources potentially relevant to a rapid mapping environment. To this end, high resolution (3060 by 2036 pixels) digital colour aerial imagery acquired with a Kodak DCS460 still video camera at a scale of about 1:20,000 in Marconi Beam, a densely populated shack settlement located in Milnerton, Cape Town, was selected. This imagery is well-suited in terms of accuracy for detail mapping exercises; however, its cost

advantage over conventional aerial photography rapidly deteriorates as the size of the area covered increases (Mason *et al.*,1997). The ground pixel size of this imagery is approximately 18cm while the effective resolution is estimated at 30 ~ 40 cm. This is sufficient to recognise even small objects such as block toilets. Also a DTM was generated with an Adam Topocart stereo plotter. Shadow data derived by classification of the orthoimage using ERDAS Imagine was obtain from Li *et al*, 1997. Figures 3.2-3.4 below depicts an orthoimage, a DTM and shadow data of Marconi Beam respectively.

Fig.3.2 Enhanced RGB (Red,Green and Blue) stretched orthoimage

Fig. 3.3 DTM of a section of Marconi Beam

Fig. 3.4 Shadows (after Li *et al*, 1997)

3.6 Experiments and results

In this section we describe the strategies and outline the results of the experiments carried out in this research. Two data sets were used in this research namely a section of a formal urban area (Avenches project) in Avenches Switzerland and a section of an informal settlement in Cape Town (Marconi Beam project). The two experiments were implemented in the Arc/Info Geographical Information System.

The use of a GIS environment implies performing building/shack detection in object space. Orthoimages therefore form the image source employed. The advantage of object space reconstruction is that all geocoded sources of information, e.g. from earlier mapping epochs, can be exploited directly (Mason *et al*, 1997). The detection of buildings/shacks is achieved by employing the AML (Arc Macro Language) in Arc/Info GIS.

3.6.1 Marconi Beam project

In this project two approaches were investigated, one being raster based while the other is vector and hybrid method. The two approaches are described below.

3.6.1.1 Raster based approach

This approach is based on the raster representation of the DSM. A raster representation is an array of pixels in 2D. Each pixel is referred to by a row and column number.

The strategy used in this experiment is based on the following principles :

- Automatic extraction of the simplest shacks, e.g. 4-sided shacks.
- Multiple cues such as edge contours, shadows, DSM, DTM etc. are fused in a two-step procedure of shack detection followed by shack extraction.
- Manual support for automated procedure where cues are inadequate is accommodated in shack extraction.

Below is a diagram outlining the strategy.

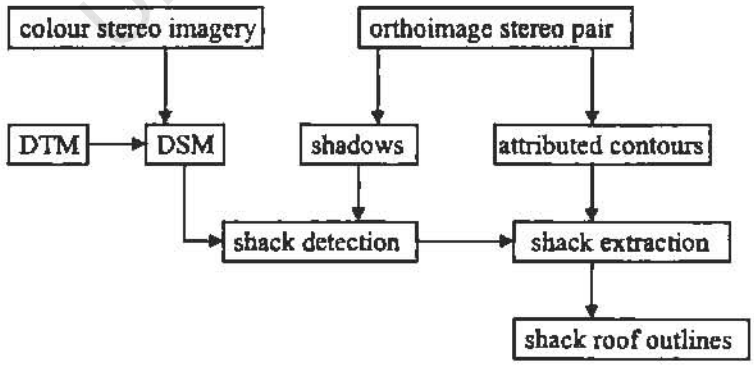


Fig.3.5 Shack reconstruction strategy (adapted from Mason et al 1997)

The approach assumes an a priori DTM (fig.3.3) of the site. Hypotheses of shacks can be derived from analysis of the DSM (fig.3.6) from overlapping images, but due to

the connectivity of the shacks it is clear that a DSM alone is not sufficient to detect individual shacks. Blob features represent clusters of shacks on the terrain. Identifying these blobs can be used to infer recognition of possible shack locations. The brighter parts on the DSM are higher than their surrounding areas which are dark. These brighter parts are possible blob locations.

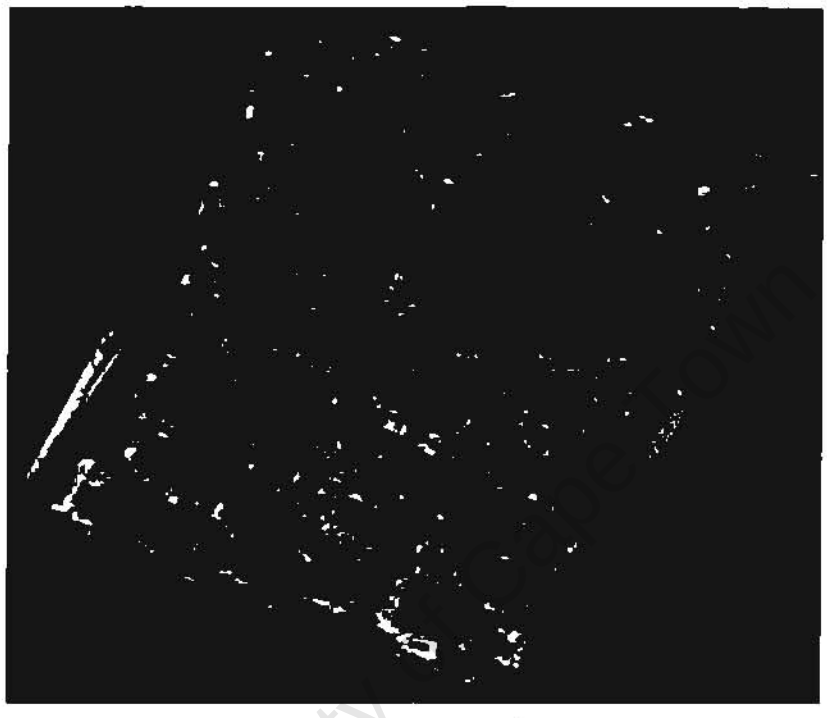


Fig. 3.6 DSM of Marconi Beam

If an accurate DTM (fig.3.3) is present, e.g. from a previous survey, subtraction of the DTM from the DSM can be used to coarsely delineate blobs. The DTM used in this work was measured at 2-5 metre spacing and interpolated to 0.25 metres.

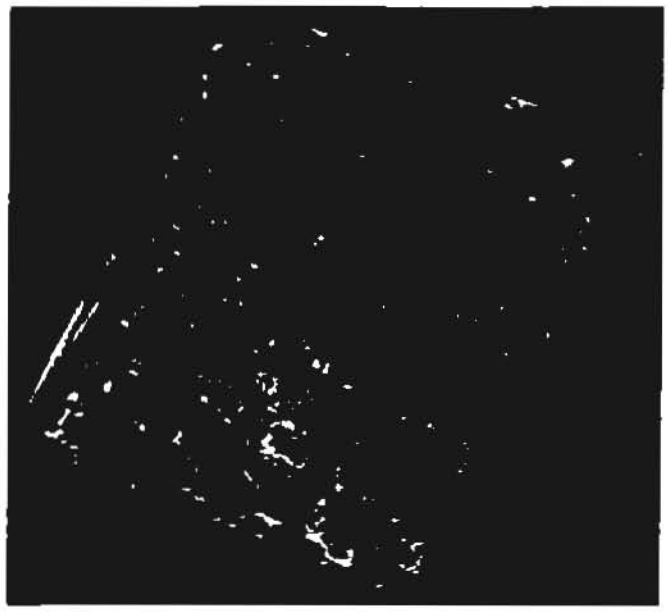


Fig. 3.7 DSM - DTM

The difference data in figure 3.7 is put on a reference surface which is approximately a plane. This permits thresholding to be used to ensure that only information above the surface is retained.

The result of the difference data is then thresholded for the anticipated lowest shack height. A threshold of 1.5m was chosen here, which is a minimum building height, so as not to exclude the lowest shacks accounting for some smoothing in the DSM peaks, but to exclude possible surface features. Figure3.8 depicts the results of thresholding the difference data set. This figure represents only information above the topographic surface.

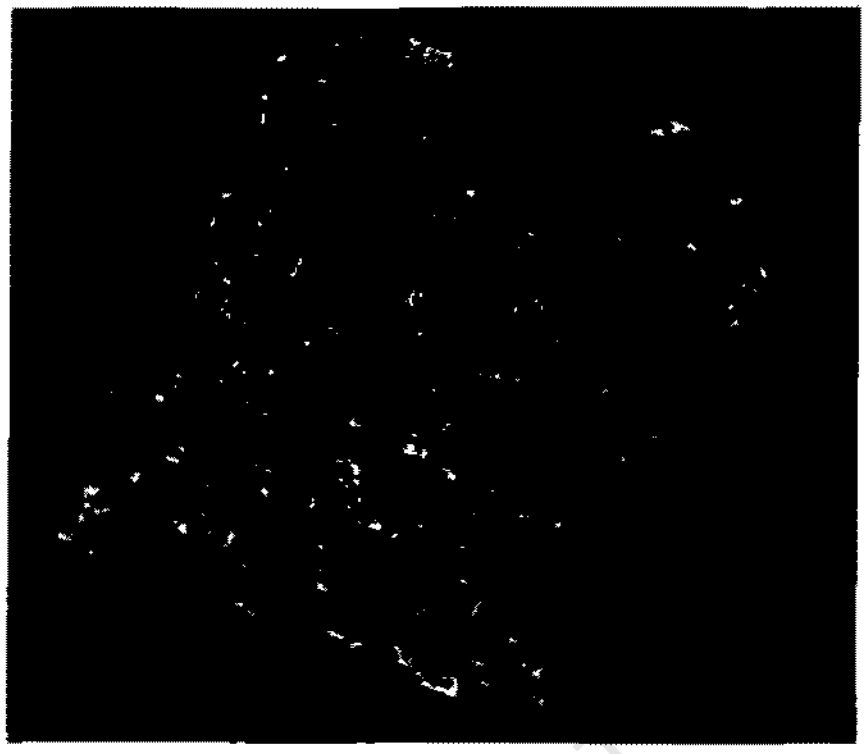


Fig. 3.8 DSM - DTM threshold

Under favourable conditions, such as no cloud cover, elevated structures such as shacks will cast shadows. Thus the orthoimage is segmented into shadows and non-shadow classes. Because the extracted shadows and the DSM are geo-referenced the shadows can be used to mask out the blobs to produce shack hypotheses. Many shacks are connected, thus leading to difficulties in individual shack detection. Masking eliminates the blobs that overlay shadow regions. The following Arc/Info commands achieve the differencing, thresholding and segmentation of blobs.

Diff = DSM - DTM (1)

thresh = **select(diff, 'value > 1.5')** (2)

out = thresh - shadows (3)

DSM and DTM are the input grids. In equation 2 we select from the difference grid all values greater than 1.5. Out is the result of segmenting the DSM to extract shack blobs. Figure 3.9 depicts the extracted shack blobs.

Fig. 3.9 DSM - DTM threshold - shadows

Verification of blobs as shacks, e.g. using image texture measures, would need to be employed in areas where vegetation exists. In this case there is hardly any vegetation. Histograms of gradient orientations can be used to distinguish buildings from trees. A histogram of the gradient orientations in the range 0 to 360 degrees at all pixels belonging to strong gradients in an image region associated with a DSM blob outline is computed to distinguish buildings from trees. Histograms belonging to buildings contain significant peaks 90 degrees apart for regularly shaped buildings with additional peaks being detected for more angular buildings. On the contrary, histograms belonging to trees are random with few observable peaks (Baltsavias *et al*, 1995). Figure 3.10 represents a shack from Marconi Beam. The gradient of the shack is shown in figure 3.11 and figure 3.12 represents the gradient orientations of the shack. As stated above the histogram contains four pronounced peaks 90 degrees apart. Figure 3.13 shows blobs overlaid on CAD data.

Fig. 3.10 Shack

Fig.3.11 Gradient of shack

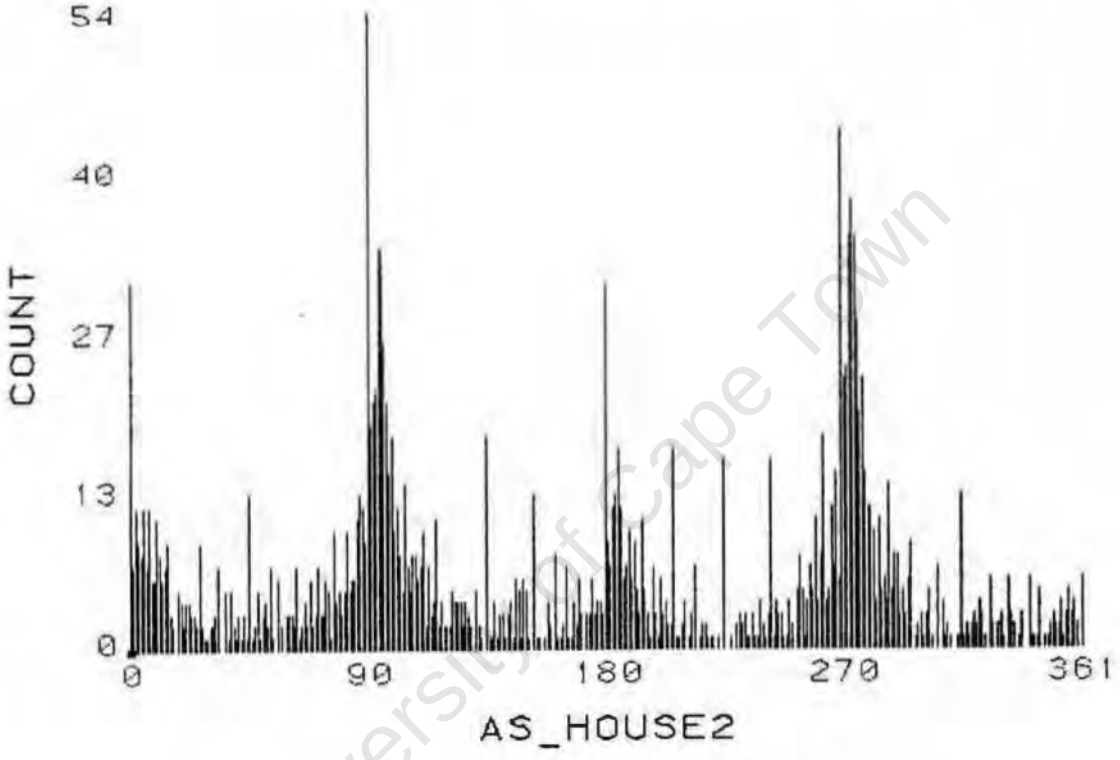


Fig. 3.12 Histogram of shack gradient orientations

Fig.3.13 CAD overlaid on blobs.

The raster based approach yields favourable results in shack detection, as depicted in figure 3.9. As stated above, the difference between an approximate topographic surface (DTM) and a DSM yields shack blobs. Due to the nature of the informal settlement, close proximity of shacks, the shack blobs are connected. Thus the use of shadows to separate the DSM shack blobs is very useful in this case. The result is shown in figure 3.9. The area consists of 80 shacks and a car at the bottom of the image (next to the 4th blobs from the left). Of the 80 shacks 73 were detected and 7 omitted. The percentage area of shacks correctly identified is 67%. The car has the same geometry as some of the detected shacks and thus cannot be eliminated automatically in the shack detection process.

The result of the shack detection is sensitive to the choice of the threshold used A threshold of 1.5m was chosen here to ensure that we do not exclude the lowest shacks accounting for some smoothing in the DSM peaks, but to exclude possible features such as cars. 1.5m is the expected minimum shack height. Figure 3.131 below shows the result of choosing a threshold greater than 2.5m.

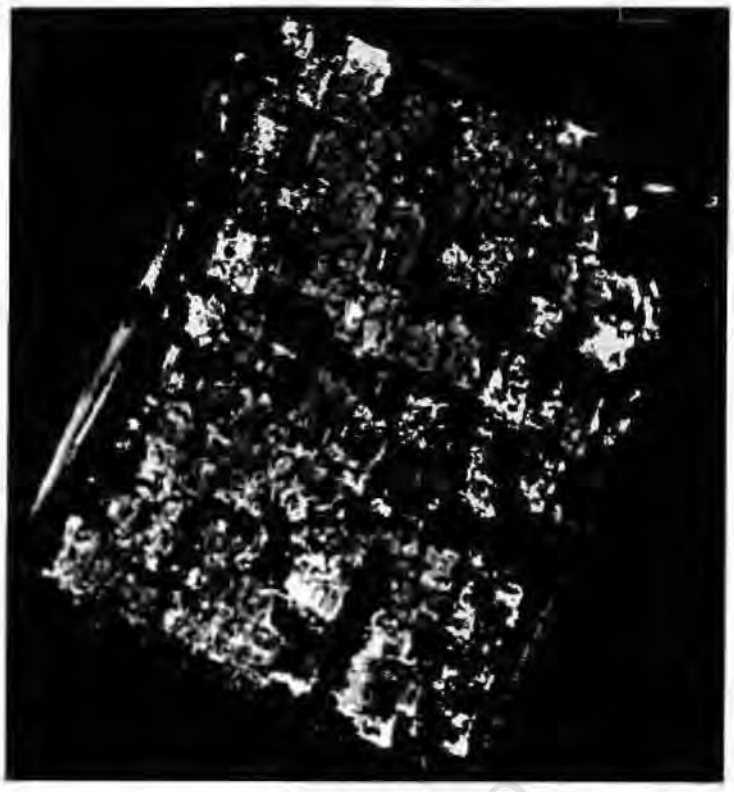


Fig. 3.131 DSM-DTM threshold(>2.5m)

Comparison of figure 3.131 with figure 3.8 shows that at a threshold of greater that 2.5m we start losing shacks. This is further illustrated by a using a threshold greater than 3m as shown in figure 3.132 below.

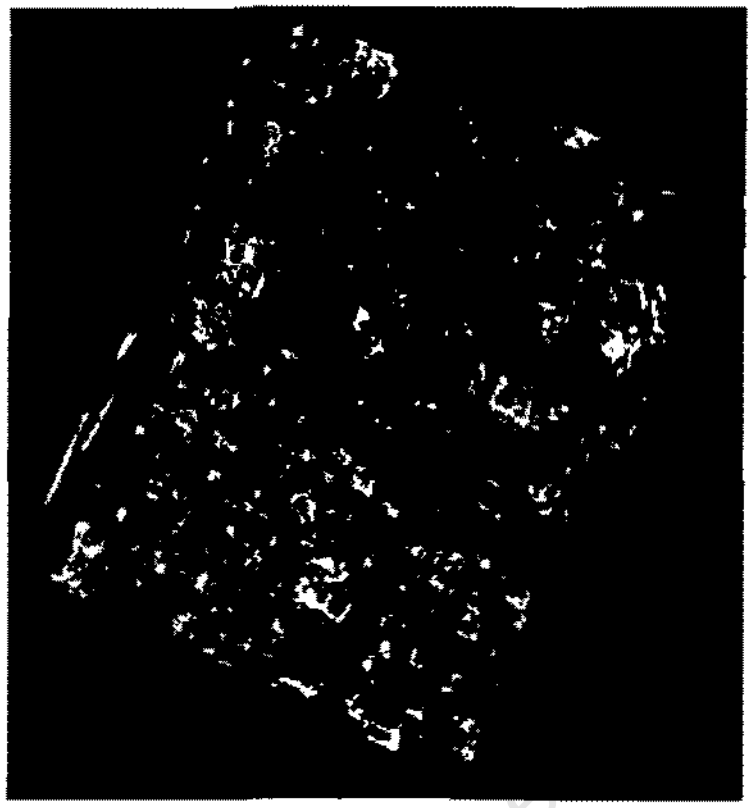


Fig. 3.132 DSM - DTM threshold(>3m)

Figure 3.132 clearly demonstrates that as the threshold is increased above 1.5 shacks are lost. Hence the use of 1.5m threshold, as stated above, is justifiable in this environment.

Also a DTM was derived from the DSM by eliminating the areas where blobs occur. The DSM was first smoothed to reduce noise prior to eliminating the blob areas. The result of the blob elimination was used to interpolate a DTM using a linear interpolation. The derived DTM was compared to the original DTM to identify erroneous areas. Figure 3.133 represents the derived DTM while figure 3.134 represents the difference between the derived DTM and the original DTM.

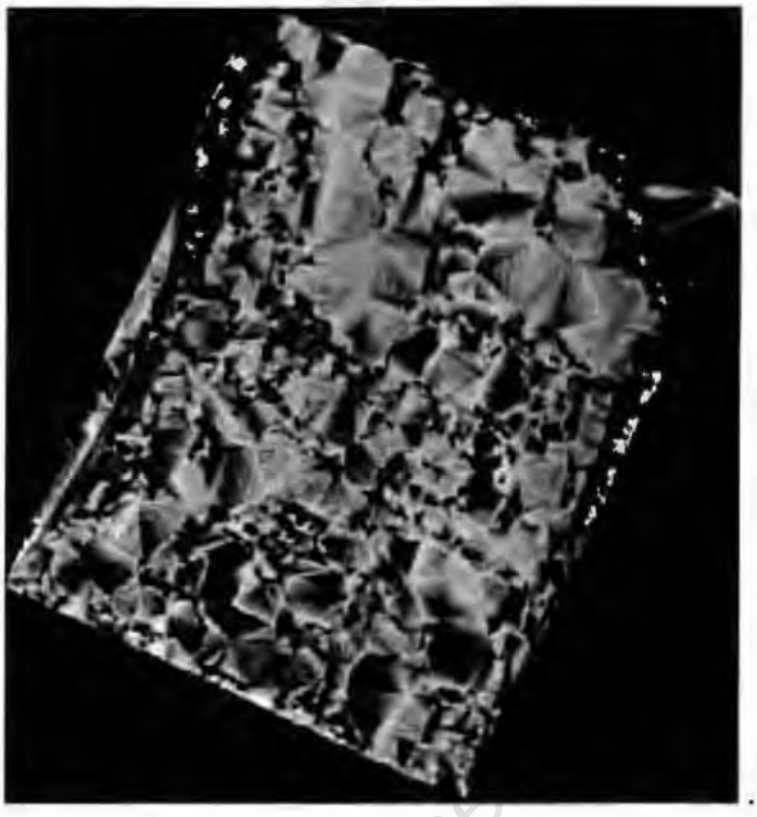


Fig.3.133 Derived DTM

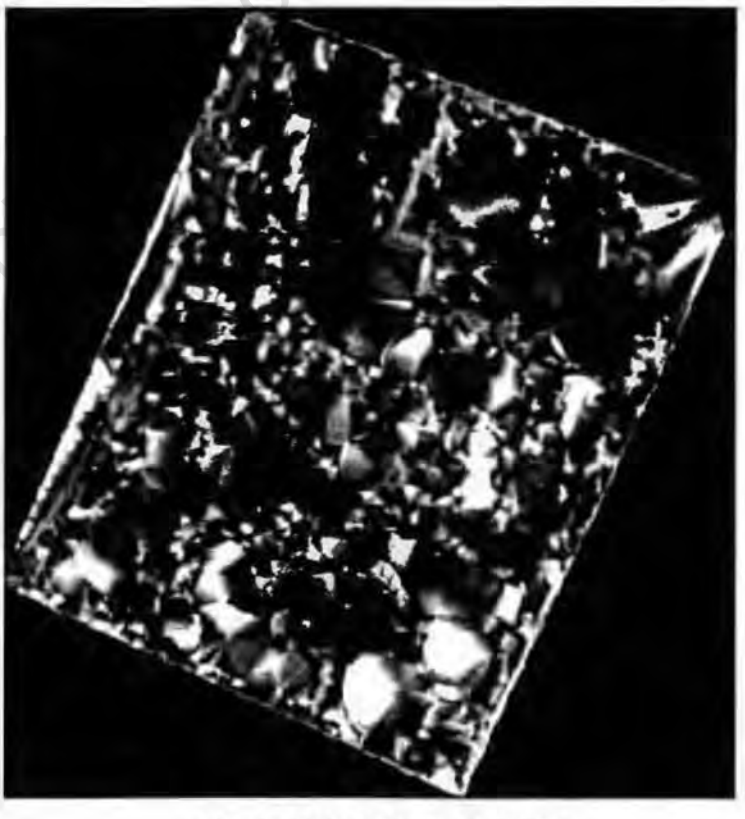


Fig. 3.134 Difference DTM

A histogram of the difference DTM was derived to quantify the differences between the derived DTM and the original DTM. Figure 3.135 represents a histogram of the difference DTM.

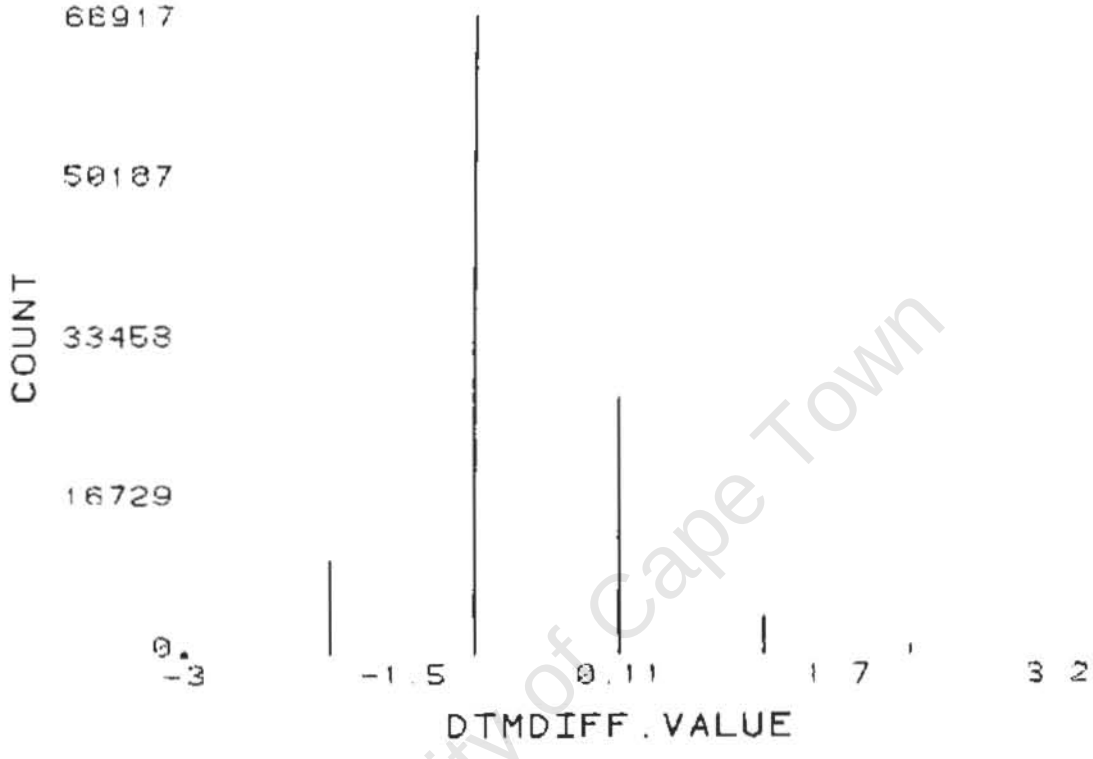


Fig. 3.135 Histogram of DTM - derived DTM

The areas with black holes in the difference DTM data corresponds to the areas where there was no data after removing the blobs. These are the areas with the high difference values in the histogram in figure 3.135. These errors are as large as 2m and are associated with the derived DTM due to the interpolation for values in the areas with no data.

3.6.1.2 Vector based approach

This approach is based on a continuous co-ordinate space which is not quantised. The DSM is represented by points, lines and polygons in vector mode. This approach is based on direct detection of shacks from the DSM. As stated in the raster based approach the shacks in the DSM are connected due to the presence of shadows and limitations of the image matching. The DSM data set is transformed into a grey-value

grid. In the grey-value grid the shack blobs appear as bright clusters. This grid is then converted into contours. We use an interval which is the same as the resolution of the DSM to avoid loss of buildings. In this project the DSM resolution used is 0.25m. Such a dense grid spacing is also necessary to distinguish buildings that are close to each other and to avoid strong smoothing of discontinuities.

Shack blobs in the grey-value grid are characterised by closed contours, at least in theory. To detect these shack blobs we employ context rules of area, perimeter and elongation. However, problems arise at the building detection stage because the shacks are merged together in the contours. This results in one contour enclosing a large number of shacks and makes it impossible to extract individual shacks. The shack boundaries inside an enclosing contour as described above are not closed as shown in the figure 3.15. The figure clearly illustrates that individual shacks cannot be extracted because their boundaries are not closed.

Fig.3.15 Shack polygons

Another approach is to use the results of the raster approach together with the vector approach (Hybrid method). Here the detected blobs are transformed into contours. The goal is to get a result with only the shacks. Figure 3.16 depicts the results of the hybrid method.

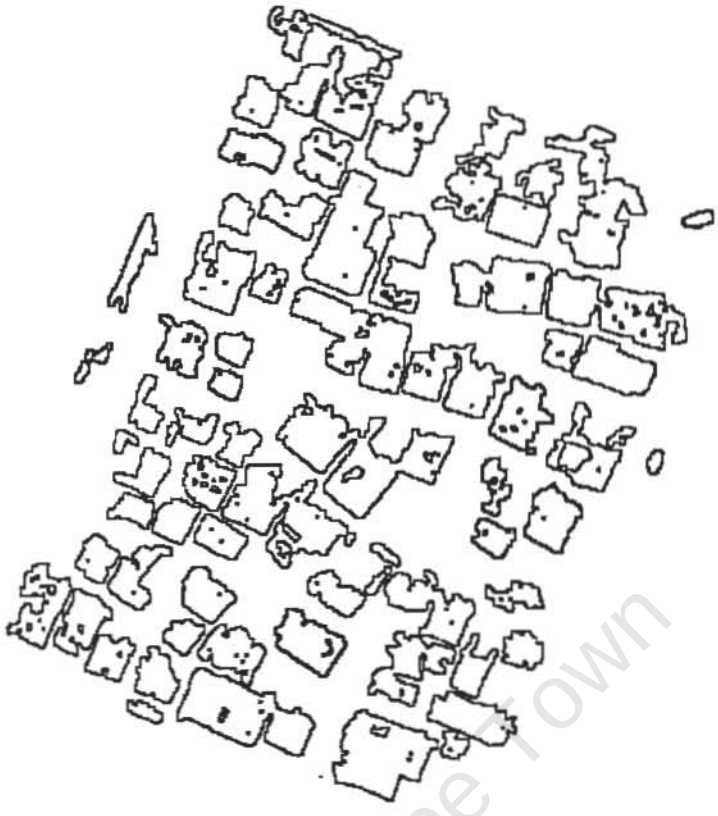


Fig. 3.16 Polygons of detected shacks

Context rules were applied on the blob contours to clean up the results. As stated above the detected car(4 from bottom right) has the same value as some of the detected shacks. Two values were assigned to the DSM-DTM threshold data, namely 1 for non shadow regions and 0 for shadow regions. This was done to facilitate the subtraction of the shadow data which had a value of 0. Also most of the noise in the above figure has the same perimeter as some of the detected shacks. This means that getting rid of the car and the noise implies loosing most of the detected shacks. Most of the detected blobs appear connected in the blob contours. This is due to the fact that they are close to each other and are connected as a result of the limitations of the image matching algorithm.

3.6.2 Avenches project

In this approach only the vector hased approach was implemented.

Avenches Strategy

The main features of the Avenches strategy are illustrated in figure 3.17.

- A DSM is transformed into contour lines.
- Polygons (closed contour lines) are extracted from the contour lines. Buildings and other objects such as trees and cars are modelled as closed structures hence the extraction of polygons.
- Context rules are used to isolate buildings from other non-building objects.
- Height statistics are calculated to identify which outer polygons represent buildings.
- The outer polygons are overlaid on the original imagery. Image information, such as gradient orientations, inside the polygons is used to verify if the polygons are buildings.

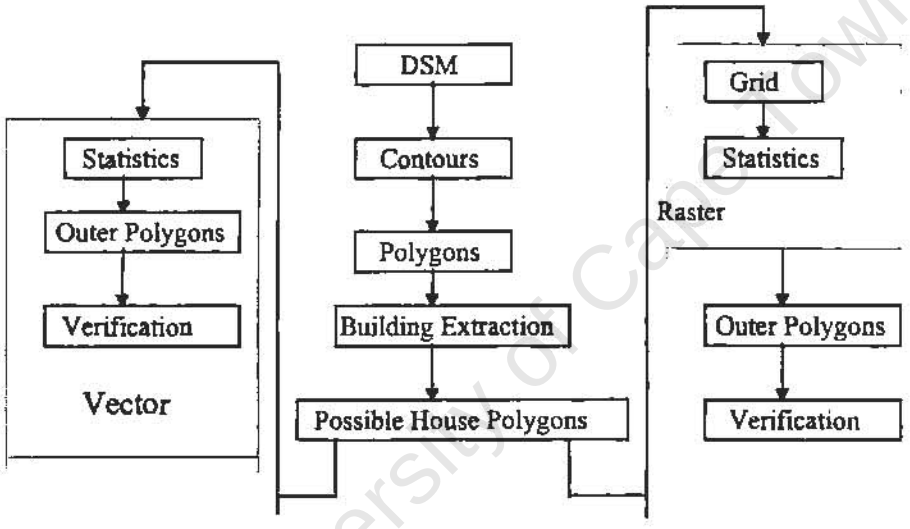


Figure3.17 Strategy for building extraction in Avenches project.

Figure 3.18 shows contour lines generated from the DSM. The contours have to be at a selected contour interval, which should not be too large or too small as this results in too much or too little detail respectively. The optimal contour interval is one which best models the DSM. In this project a 1m interval was used. Buildings are closed structures, hence the need to eliminate all unclosed contour lines. Figure 3.19 is a representation of all closed contours. The following Arc/Info commands were used to extract closed polygons:

Latticecontour *ingrid outgrid interval* (1)

ef polys (2)

sel all (3)

put *name* (4)

In equation 1 ingrid is the name of the input grid, outgrid is the output grid and interval is the contour interval. Command 1 generates a contours from the input grid. The second and third commands select all closed polygons which are then save in command four.

Knowledge about buildings is used to segment buildings from other man-made or natural structures, such as cars and trees, which are also represented as closed structures. This knowledge is derived from context rules such as area, perimeter, elongation and height. A tree and a building will have different areas, perimeters and elongation though they may have the same height.

Sel perimeter > 40m (5)

sel area > 20 square meters (6)

sel elongation < 0.6m (7)

Equations 5-6 extract polygons with perimeter greater than 40m, area greater than 20 square meters and elongation less than 0.6m.

Figure 3.20 depicts polygons likely to represent buildings. Height statistics are used to determine the height differences between the outer polygon and the inner ones. The result of this statistical analysis is used to extract the outer polygons which are likely building outlines as shown in figure 3.21. The following Arc/Info commands are used to derive statistics and extract outer polygons:

dissolve dsm dsm1 dsm-id (1)

identity dsm dsm1 dsm2 line (2)

frequency dsm2.aat dsm.frq (3)

statistics dsm.frq dsm.stat dsm1-id (4)

sel lpoly# = 1 (5)

Command 1 merges adjacent polygons or lines in **dsm** which have the same value for a specified item**(dsm-id)** and outputs **dsm1**. In command 2 the inputs are **dsm** and **dsm1** while **dsm2** is the output. The command computes the geometric intersection of two coverages. All features of the input coverage, as well as the those of the identity coverage that overlap the in coverage, are preserved in the output coverage. The third command produces a list of the unique code occurrences and their frequencies.

Command 4 generates summary statistics while command 5 selects the outer polygons. Figure 3.22 shows the blobs overlaid on the DSM.

Fig. 3.18 Contours of a section of fig. 3.1(stretched)

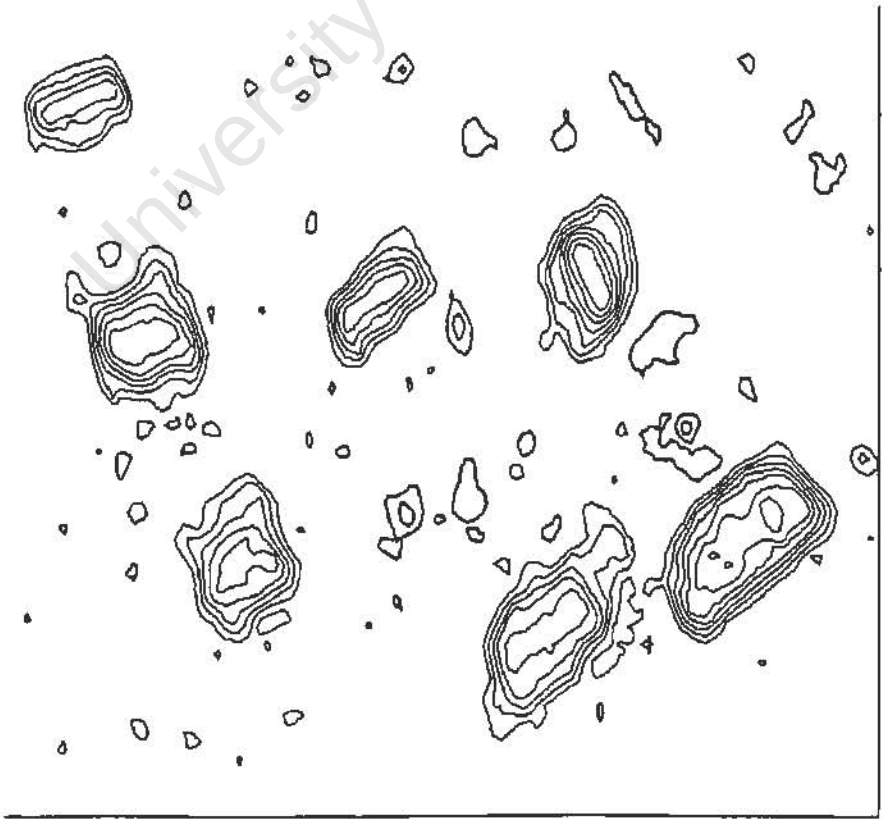


Fig. 3.19 Polygons of fig.3.18

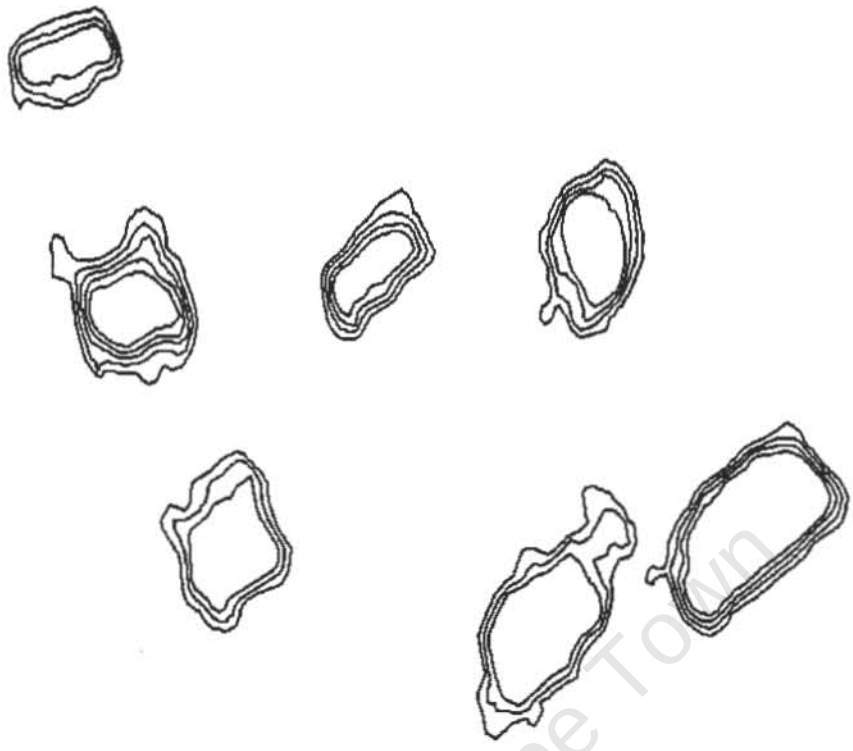


Fig. 3.20 Extracted blobs

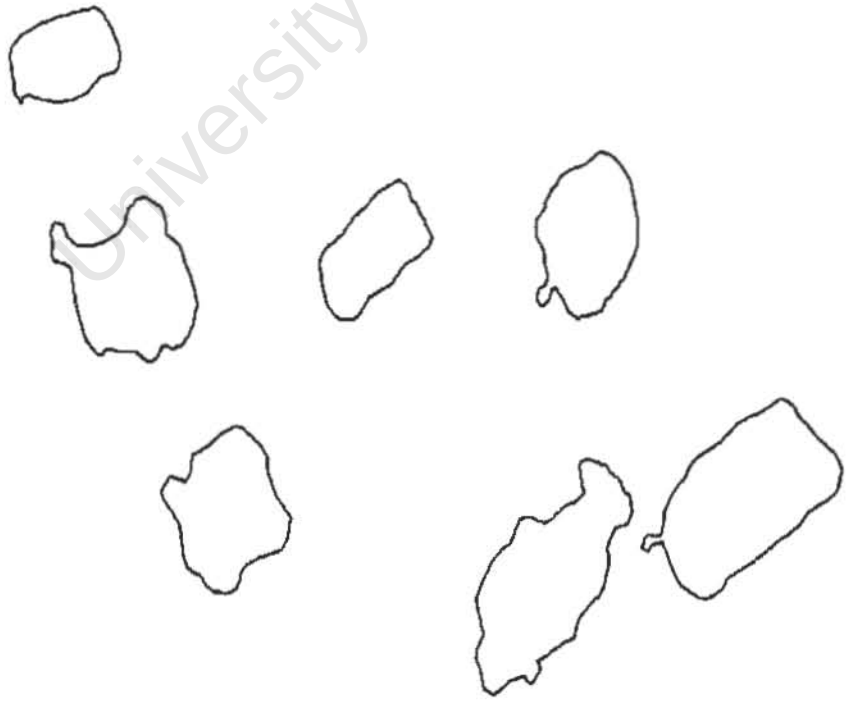


Fig. 3.21 Blob outlines

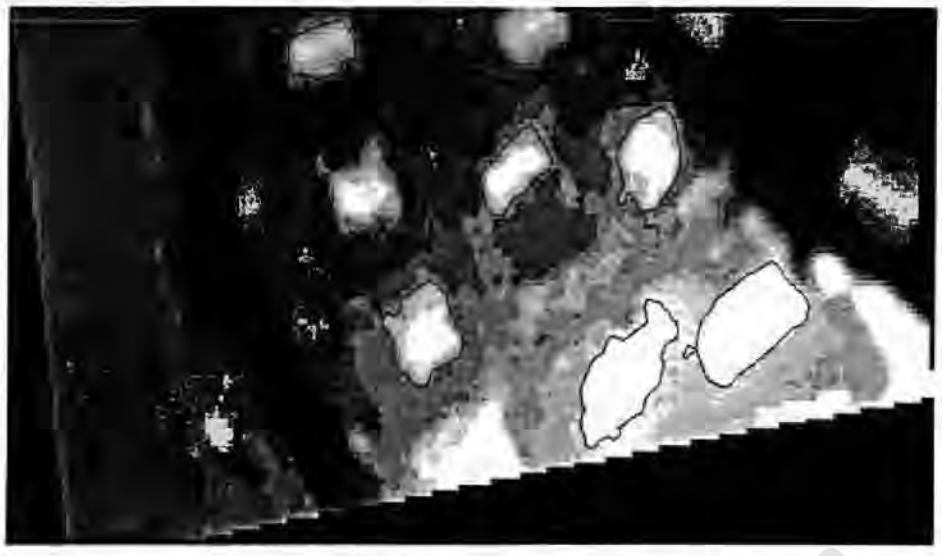
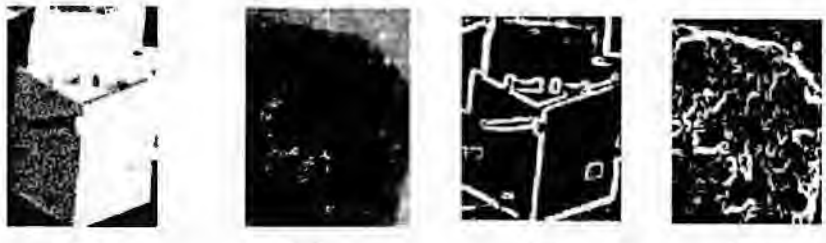


Fig 3.22 blob outlines overlaid on section of DSM

Figure 3.22 shows some blobs not detected. This is because these blobs are not closed and we are extracting closed polygons. The blob outlines detected in a DSM (fig 3.21) can be attributed to many different 3D objects. Context knowledge, such as shape bounds (area, elongation, etc.) on buildings in a given area, has already been applied to preclassify the blobs and we rely now on information in the images. A histogram of the gradient orientations in the range 0 to 360 degrees at all pixels belonging to strong gradients in an image region associated with a DSM blob outline is computed to distinguish buildings from trees. The histograms are generated by the Arc/Info **histogram** command which takes a grid of gradient orientations as its input. Histograms belonging to buildings contain significant peaks 90 degrees apart for regularly shaped buildings with additional peaks being detected for more angular buildings. On the contrary, histograms belonging to trees are random with few observable peaks (Baltsavias *et al,* 1995).



(a) **(b)** **(c)** **(d)**

Fig.3.23 (a)formal building, (b) tree, (c) gradient of building and (d) gradient of tree

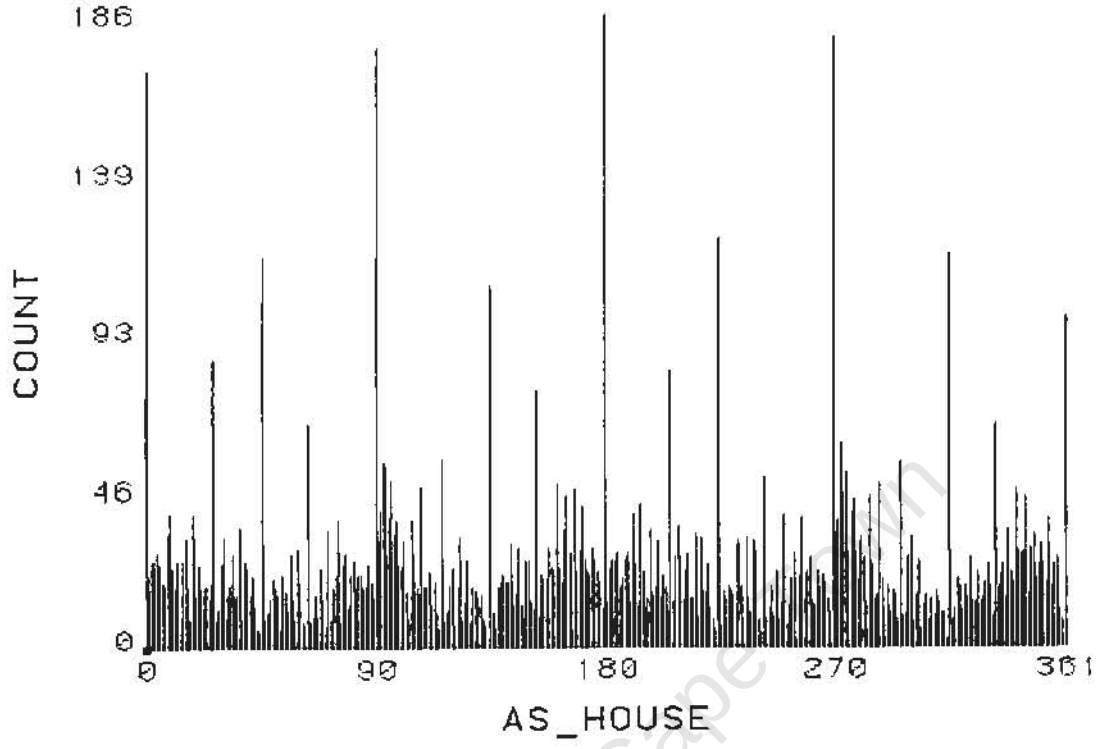


Fig.3.24 Histogram of gradient orientations of building

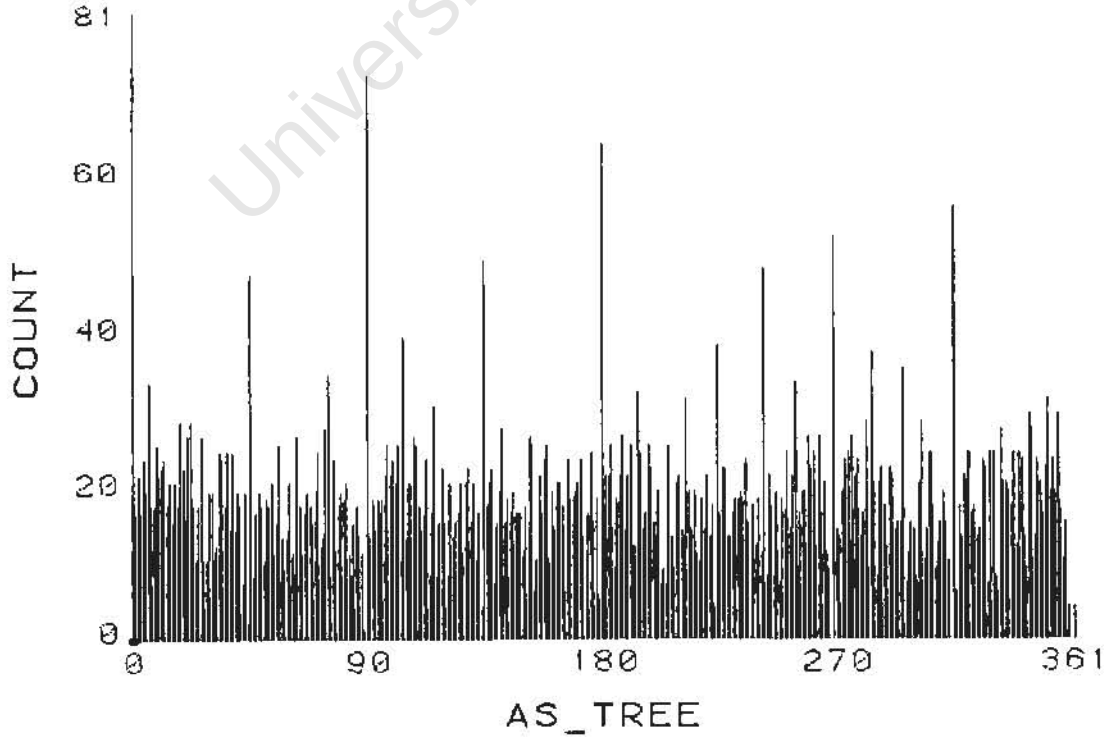


Fig.3.25 Histogram of gradient orientations of tree

The histogram in figure3.24 has four distinct peaks 90 degrees apart and the gradients are stronger than the gradients of figure 3.25 which has rather smooth peaks. As it has been stated above, the use of a DSM in formal settlements, where buildings have similar charecteristics, to detect buildings is expected to work well. This expectations arises from the fact that the buildings in formal settlements are separated and thus can be segmented. This implies that the image matching has worked well. Figure 3.18 shows groups of polygons for each individual building. These groups of polygons are so detached that one can already discern the buildings. The processing of these groups of polygons to remove the contours representing the terrain and other man-made features results in figure 3.20. Using height statistics on the detected building blobs enables us to extract the outer polygons which represent buildings as depicted in figure 3.21. With such results one can conclude that the use of a DSM to detect buildings has indeed worked for this particular settlement as verified in figure 3.22 and by the histograms of gradient orientations. The buildings that are not detected as shown in the overlay are not complete buildings in the DSM, so their polygons are not closed and thus cannot be extracted. This approach is limited by the results of the image matching algorithm. If the results are good, that buildings are seperated in the DSM, then the approach is expected to work well.

3.8 Chapter Summary

This chapter describes the uses of both the DSM and the DTM. The uses of a DTM are further analysed in the chapter on hydrological analysis (chapter 5). This chapter begins by analysing the different building detection strategies employed in formal settlements. In general, two strategies are presented. One is the detection of buildings directly from DSM while the other uses both the DSM and the DTM. The work presented in this chapter employs both strategies in informal settlements and only the vector approach in the formal settlement. Analysis of the results reveal that the vector approach is applicable in the formal settlement only. The raster approach was implemented in the informal settlement only and results revealed that it is applicable. The errors of omission and commission are also presented in the analysis section.

In the informal settlement the raster approach assumes the presence of a DTM from a previous survey. The first step of the building detection is the subtraction of this DTM from the DSM. This subtraction puts the data on an approximate surface which is a plane and hence the thresholding this difference data set. This ensures that only information above the terrain is obtained. A threshold of 1.5m is used in this study, which is a minimum expected height of the shacks. Buildings cast shadows onto each other. If these building are in close proximity they will be connected due to the presence of shadows. Thus the difference data set is segmented by subtracting the shadows data. The resulting data contains a DSM of segmented shacks.

The vector approach aims to segment the DSM without using the DTM. The DSM data is transformed into contour lines. The resulting contours are not closed and hence the extraction of closed polygons which represent individual buildings is impossible.. In the formal settlement the same approach works because the buildings are properly structured and detached. This factor enables the successful detection of the buildings' outlines.

CHAPTER 4

DSM Quality Analysis

4.1 Parametric statistical test of DSM quality

4.1.1 Introduction

The Digital Surface Model (DSM) is the basic source of information for developing other models, totally or partially dependent on topography such as building models because it models all 3D objects above the topographic surface. The usefulness and validity of the results obtained are intimately associated with the quality of the original DSM model. Quality is measured in terms of the kind and magnitude of its errors. However, DSMs are created, distributed and used very frequently without any reference to the magnitude of the error implied or to the methods applied to its disclosure and correction.

In the case of matrix DSMs (regular grid), the errors are of an attributive type, that is, they imply an incorrect assignment of altitude and they modify the values of the Z-ordinate. These errors commonly appear in the creation process of DSMs, both by automatic and manual procedures. It is therefore necessary to apply systematic methods for their detection, measurement and correction (Felicísimo 1994). The method presented here is based on the evaluation of transition probabilities, that is, that the differences between one Z-value and its neighbours may acquire a determined value.

4.1.2 The need for statistical test

The large errors in regular models are a special case in which some data have clearly underestimated or overestimated values. A regular grid is a grid of continuous data with points selected at specific intervals. These errors may be generated by automatic stereo correlation methods, that may have operative problems as a result of low contrast in images, ambiguities due to the repetition of objects or to periodic patterns.

In practice, the detection of errors in DSMs is carried out in many cases by visual inspection of perspective displays or shaded relief displays, whereas systematic and exhaustive calculation analysis methods should be applied. Generally, the errors tend to be detected by the assumption that land surfaces are essentially continuous and hence that comparison of close values may be a valid tool for the detection of anomalous data. These tests present problems in practice because they have no statistical value in the sense that they do not offer an error probability measurement for the altitude that is being checked.

The method proposed by Felicísimo (1994) has the following properties :

(1) It is an objective method, where error parameters are determined by statistical methods.

(2) We can assess the probability of erroneous points.

(3) The threshold values are deduced from data from the model, and are therefore adapted to the relief characteristics of the data under study.

4.1.3 Parametric test based on deviations of altitude values

The test is based on the analysis of the existing differences between two altitude values for each point : the first one is that collected by the DSM (right or wrong), the second is a value obtained by interpolation from the neighbouring points. The process calculates for each point of the model an altitude value using its closer neighbours. A bilinear interpolation using four closer neighbours may be enough. In this instance, the altitude assigned to a point placed in row i and column j is given by the following expression :

$$\bar{z} = 0.25 * (z_{i,j-1} + z_{i,j+1} + z_{i-1,j} + z_{i+1,j})$$

The difference between the estimated height and that of the DSM is :

$$\delta_{i,j} = \bar{z}_{i,j} - z_{i,j}$$

If the process is applied to the totality of the DSM, we can obtain an arithmetic mean and a standard deviation of the differences :

$$\bar{\delta} = 1/n \sum_{i,j} \delta_{i,j}$$

$$s_\delta = \sqrt{1/n\sum_{i,j}(\delta_{i,j} - \bar{\delta})\hat{}\ 2}$$

The previous values define a Gaussian distribution $n(\bar{\delta},s_\delta)$. Knowing the parameters that define the distribution allows us to perform a meaningful test for the individual values of $\delta_{i,j}$ with which we can validate or reject the hypothesis that the individual deviation value observed may belong to the population of deviations. This is implemented by the Student *t* test :

$$t_{i,j} = (\sigma_{i,j} - \bar{\delta}) / s_\delta$$

This value is considered a standardised deviation, and as the number of data from the model is normally high, its magnitude is compared to the tabulated value $t_{\alpha[\infty]}$ of 3.219 for α = 0.001. The location of points with significantly high $t_{i,j}$ does not necessarily imply an error, but is an excellent alarm sign (Felicísimo 1994).

4.1.4 Algorithms for error detection

The error detection model presented here has been developed by Felicísimo (1994) in the context of the Arc/Info Geographical Information System using the AML (Arc Macro Language) computing language. The program carries the test for the grid data structures (matrix DSMs used by the GIS) within a module of the same name. If the DSM is named, for instance, dem, the basic commands are :

(1) mddif = dem(0,0) - ((dem(0,-1) + dem(0,1) + dem(-1,0) + dem(1,0))/4)
(2) &describe mddif
(3) &set zmean = %grd&mean%
(4) &set zdesv = %grd&stdv%
(5) valts = abs ((%zmean% - mddif)/%zdesv%)
(6) demtmp = setnull (valts > 10, dem)

The first command constructs a regular grid named mddif, being the result of finding the difference between each point in the DSM and the mean altitude of the four closest neighbours. The command is implicitly recursive and it sequentially affects all elements in the DSM. Command 2 executes a statistical routine that calculates the arithmetic mean and standard deviations of mddif values, by assigning them to two system predefined variables (grd&mean and grd&stdv) and executing equations 3 and

4. The arithmetic mean and standard deviation values are stored in zmean and zdesv variables in order to be used in equation 5. Command 5 calculates the absolute values of the statistical Student *t* test. Then command 6 generates a regular grid similar to the original DSM, but where wrong data are, as they are identified by the error condition (valts > 10), eliminated by the command *setnull* by assigning them a special value called *nodata*.. Results of the statistical test are presented in figures 4.1-4.4.

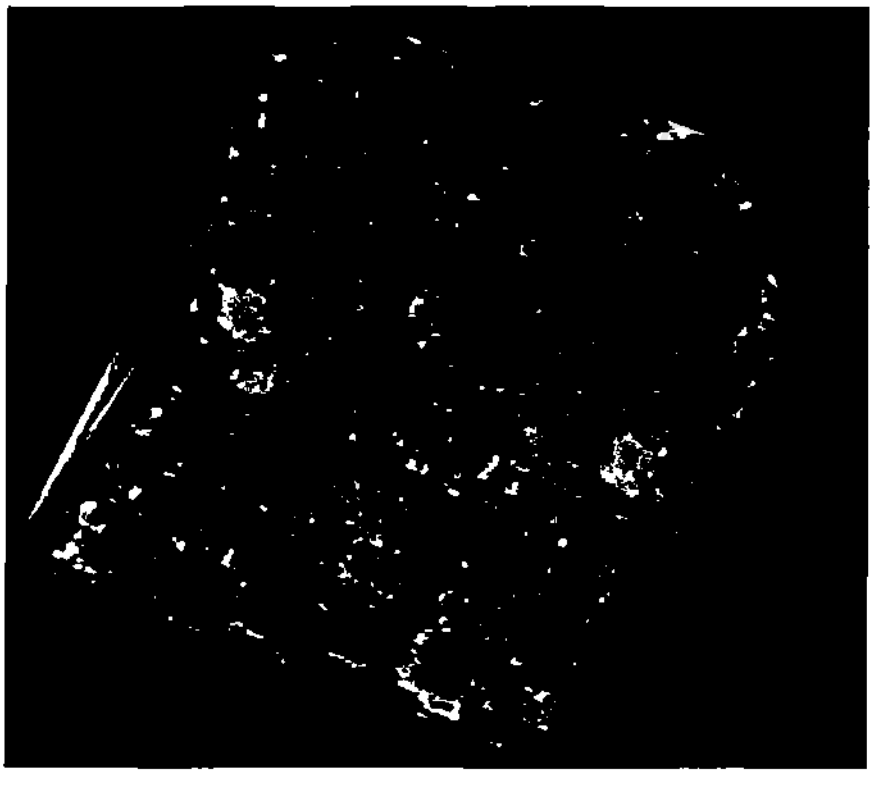


Fig. 4.1 Wallis with nodata values

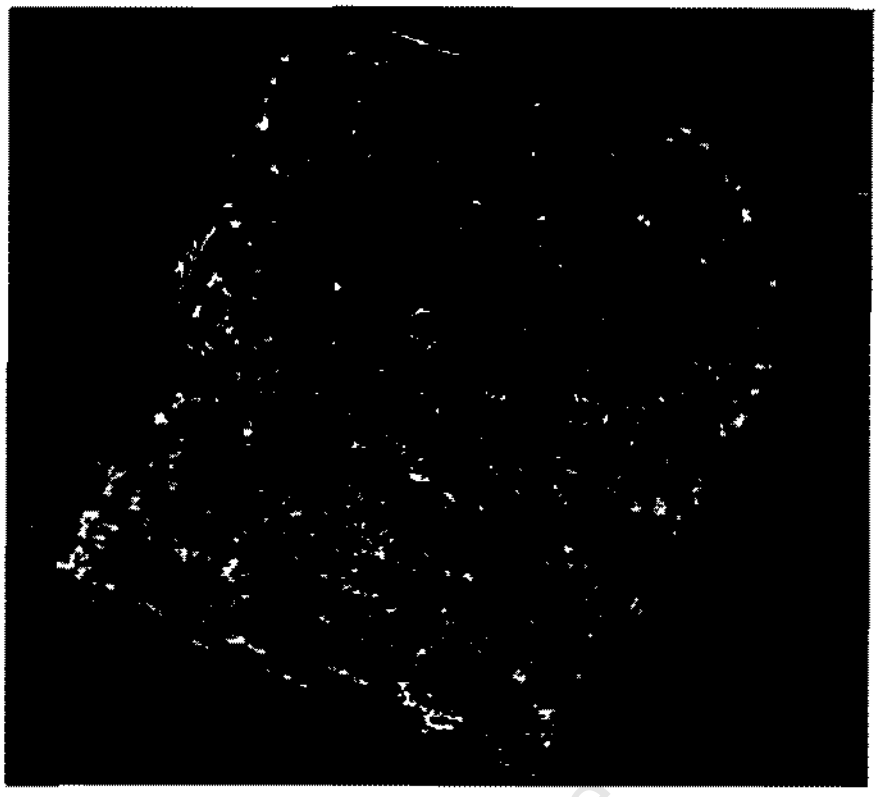


Fig.4.2 Median with nodata values

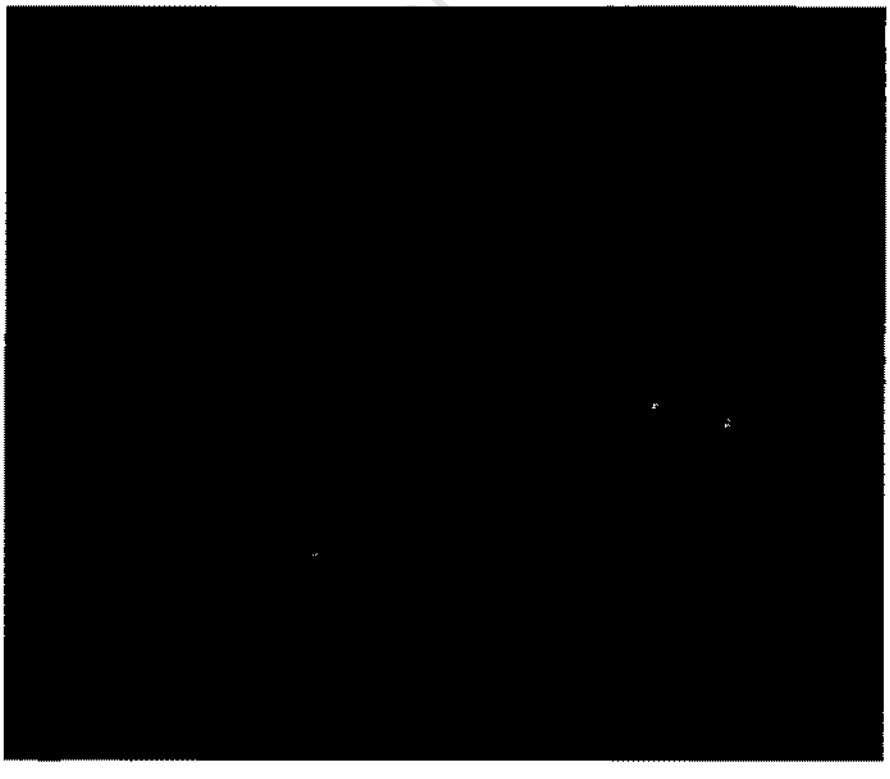


Fig.4.3 Errors in Wallis DSM

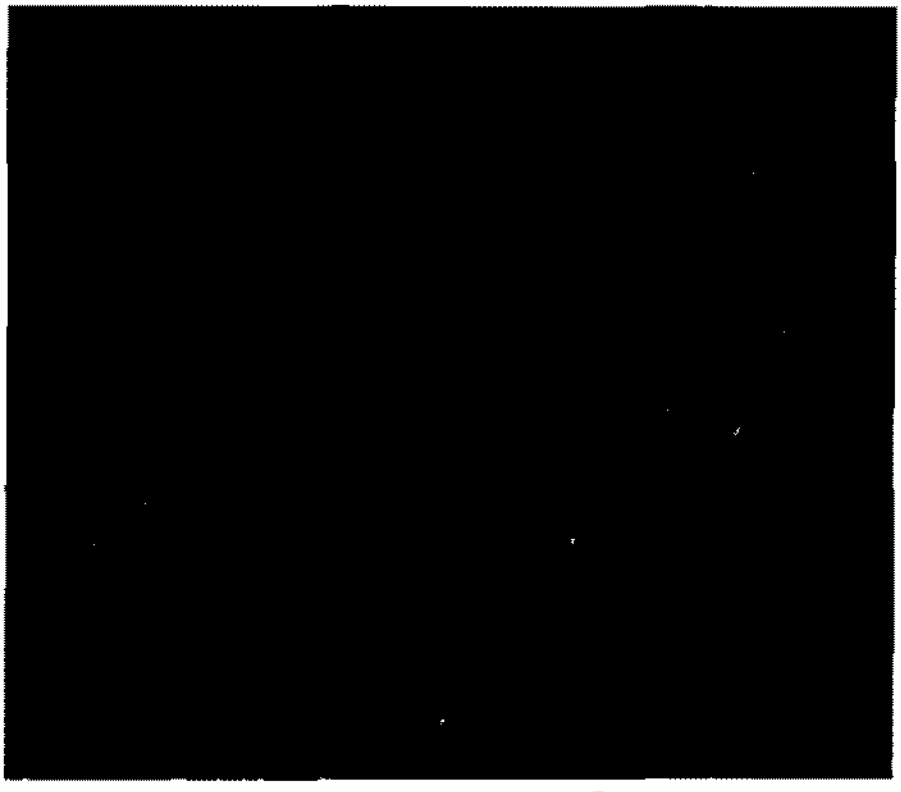


Fig. 4.4 Errors in Median DSM

4.2 Visual Quality Analysis

Two grids were generated by image matching with the same exterior orientation parameters. The two grids are the Wallis and Median enhanced grids in figures 2.2 and 2.3 respectively. The enhancement was performed prior to the image matching to reduce noise and to smooth the original DSM grid. Ground truth vector data was also generated from an orthophoto of the original grid. Mismatches can result from a variety of conditions, including sensor noise, low contrast in portions of the image, relief-induced distortions between the images and the presence of ambiguities due to identical objects or highly periodic textures on the terrain (Hannah 1981).

The shacks detected from the Wallis and the Median enhanced grids in building detection were overlaid on the ground truth data (see fig. 4.5 and 4.6). Visual inspection, of perspective displays, for the deviations of the Wallis and Median enhanced grid shacks from original shack contours was performed. From this inspection no diffences can be detected. Another visual approach is to derive histograms from the ground truth (fig.4.7), Wallis enhanced image (fig.4.8) and

Median enhanced image (fig.4.9). Also scattergrams of the Wallis-ground truth (fig.4.10) and Median-ground truth (fig.4.11) are presented.

Fig.4.5 Wallis and ground truth

Fig.4.6 Median and ground truth

The overlays in figure 4.5 and 4.6 show no visual differences.

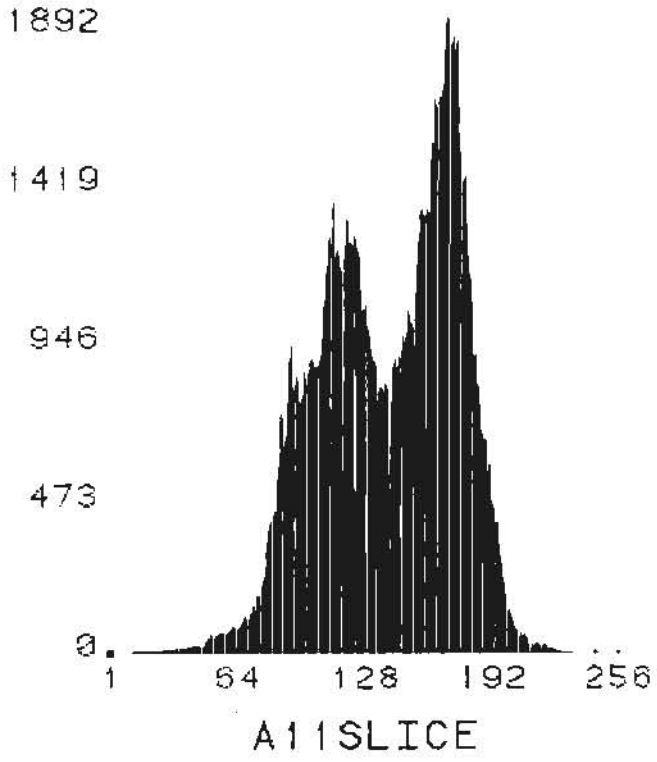


Figure 4.7

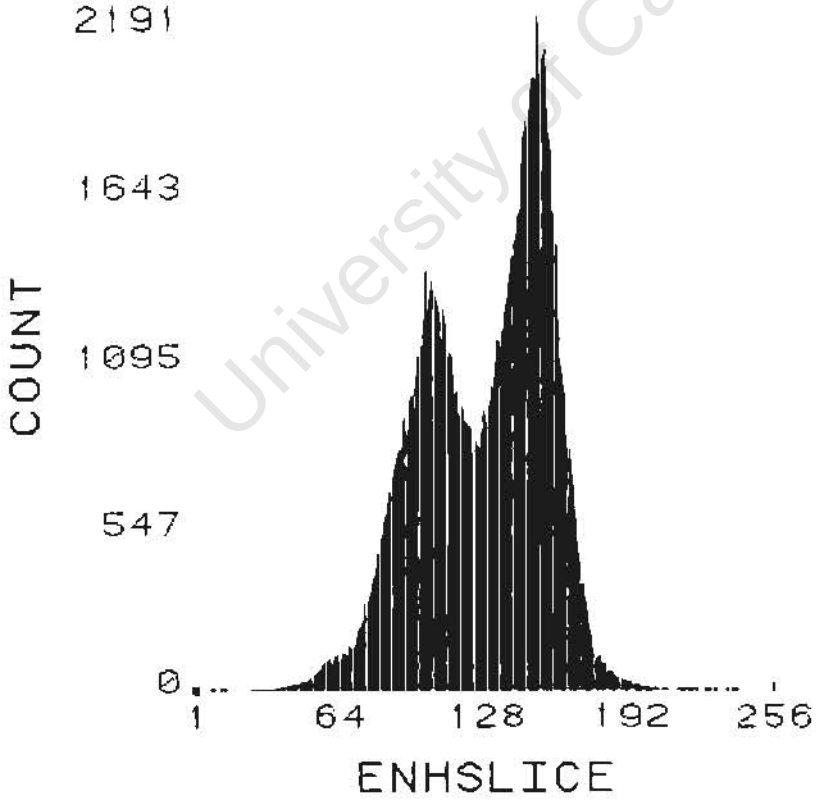


Figure 4.8

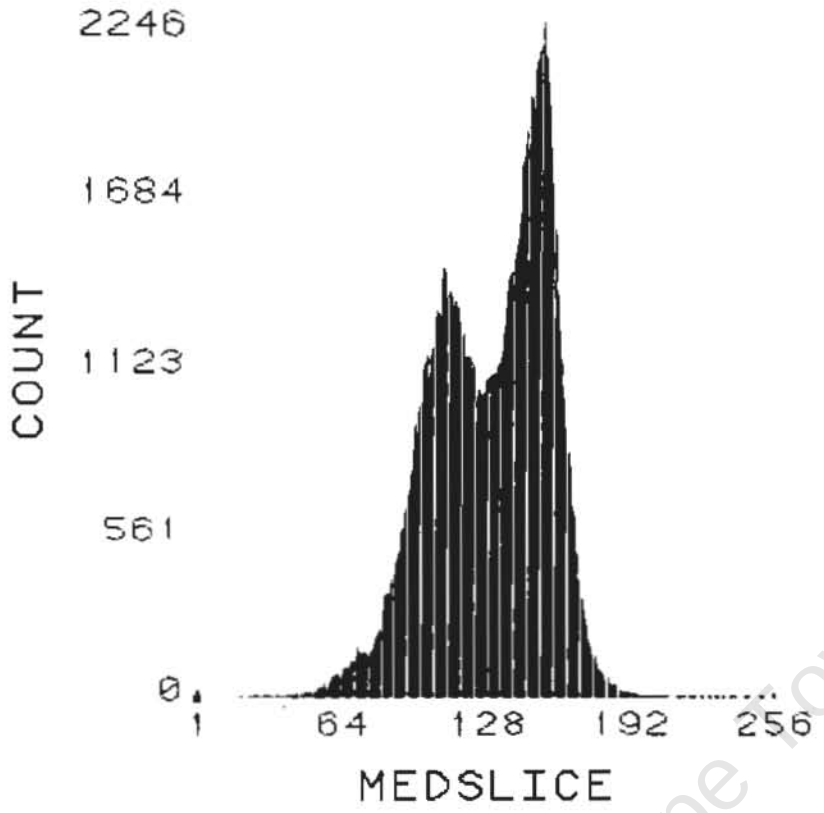


Figure 4.9

The histograms in figures 4.7 to 4.9 were derived using the Arc/Info **histogram** command which takes a grid as input. The Wallis filter and the Median filter were applied to figure 4.7. The results show no apparent diffences between the Wallis enhanced grid and the Median enhanced grid

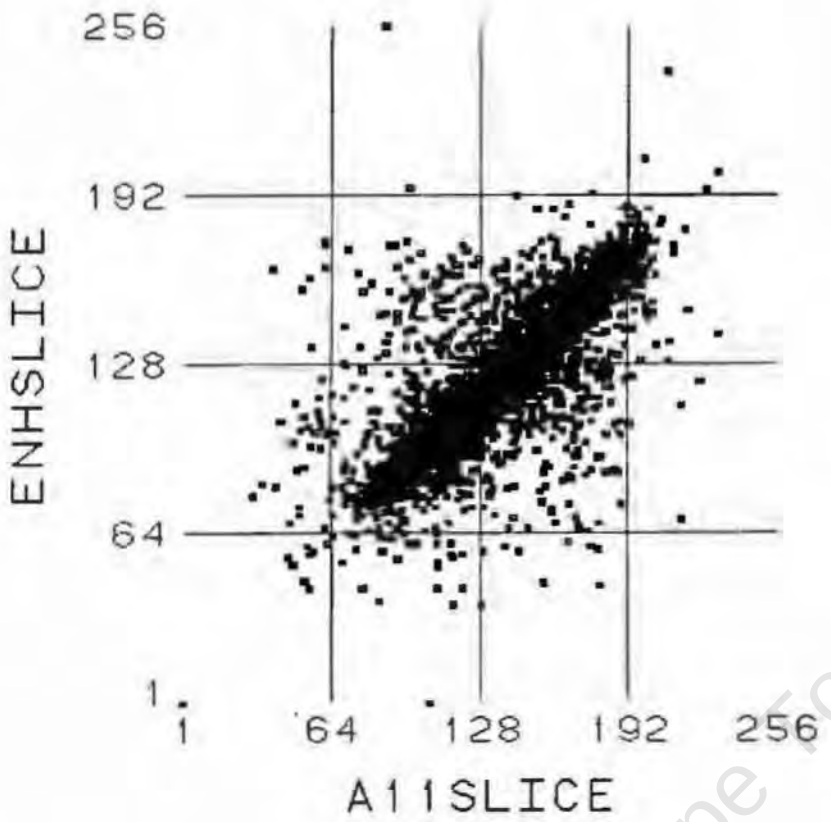


Figure 4.10

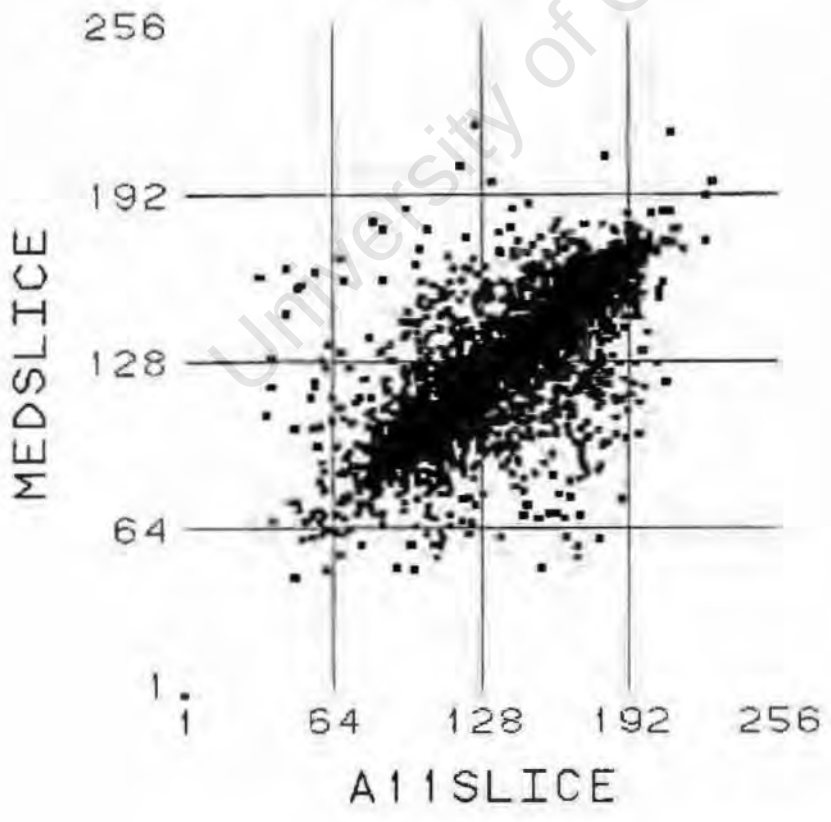


Figure 4.11

Figure 4.10 and 4.11 shows a 2D relationship between the Wallis enhanced grid and the Median enhanced grid with the original grid respectively. Because the Wallis filter better enhances than the Median filter, the dots in figure 4.11 are more concentrated in the centre than in figure 4.10. This implies that the Wallis filter enhanced grid is better enhanced than the Median filter enhanced grid.

4.3 Analysis of results

Figure 4.1 and 4.2 represents a DSM similar to the original Median and Wallis DSM but with the erroneous data replaced by *nodata* values. Figures 4.3 and 4.4 give the errors for the values of the student t-test greater than 3.219. The figure shows that the Wallis enhanced DSM has fewer erroneous data than the Median enhanced DSM. The errors are represented by the white dots in figures 4.3 and 4.4.

In figures 4.5 and 4.6 of the visual inspection the Wallis and the Median enhanced DSMs overlaid on the CAD data from the ground truth show no visual differences. Therefore the overlay is not a useful analysis procedure in this case. A promising approach is to use histograms and scattergrams as shown in figures 4.7 to 4.11. Histograms in fig.4.7, 4.8 and 4.9 for the original shacks, Wallis and Median enhanced DSM respectively, show no differences between the latter two. The scattergrams in fig.4.10 and 4.11 show the relationship between Wallis-ground truth and Median-ground truth respectively. The Wallis- ground truth relationship has more outliers than the median-ground truth relationship. This is because the Wallis filter enhances the original grid more than the Median filter. Most of the data in both relationships is between the pixel values of 64 and 192 in the 8-bit quantized DSM representation as depicted in the scattergrams. However, if the noise range is large the 8-bit representation results in only 256 grey-values and thus the quantization becomes coarse.

4.4 Chapter summary

This chapter describes two approaches to DSM quality analysis. One is based on statistical analysis using the student t-test while the other is based on visual analysis. The student t-test showed that the Wallis enhanced DSM is superior to the median enhanced DSM. In the visual analysis, overlaying of data is not an efficient tool because differences cannot be observed between the two data sets. On the other hand the histograms and the scattergrams clearly demonstrate that the Wallis enhanced DSM is superior to the Median enhanced DSM , the Wallis filter enhances the original data more efficiently.

The statistical test used in this study is based on analysing differences between two elevation values for each point. One value is from the DSM while the other is an approximation of that value calculated by the bilinear interpolation. The test generates an arithmetic mean and a standard deviation of the differences which are then used in the student t-test. The student t-test generates a standardised standard deviation. Because the data in the DSM is normally high the standardised standard deviation from the t-test is compared to the tabulated value of 3.219 for each point. Values that are higher than the tabulated value are a good signal for erroneous data. An algorithm for the student t-test is presented in section 4.1.4.

The visual analysis is based on the assumption that the data is of a continuous nature. The analysis employs the overlaying of data and the production of histograms and scattergrams. The overlay provides a means to check how each data set deviates from the ground truth. The histograms show the distribution data value across the DSM. Comparison of histograms is also used to determine the deviations of these data values from the shacks. Scattergram are a two-dimensional comparison of two data sets and the determination of the deviations between the two. The scattergrams show that the Wallis filter is superior to the Median filter. The visual analysis is problematic because it has no statistical value, i.e. it offers no error probability measurement for the elevation that is being checked.

CHAPTER 5

HYDROLOGICAL ANALYSIS

5.1 Introduction

Hydrological analysis is one the important applications of digital terrain models. The provision of infrastructure such as dams, hydrologic stations, drainage systems as well as roads and bridges is important in upgrading the quality of life. Town planners and architects rely on slope and aspect of the terrain in planning sustainable houses. Civil engineers depend on slope information to design roads and bridges. Analysis of the terrain is vital in the placement of dams and drainage systems. Drainage systems are important for the provision of water and sewer. A digital terrain model is a primary input in the hydrologic analysis of a surface.

The shape of a surface will determine how water will flow across it. The hydrological tools in Arc/Info GIS GRID module provide a method to describe the physical characteristics of a surface. Using a DTM (digital terrain model) as input, it is possible to delineate a drainage system. Watersheds created from DTMs using GRID can be used to determine the height, timing and inundation of floods (ESRI 1990). An understanding of the shape of the earth's surface is useful for many fields such as urban and regional planning, agriculture and forestry. These fields often require an understanding of how water flows across an area, and how changes in that area may affect that flow in both formal and informal settlements. The informal settlements around Cape Town have a problem of flooding, hence the importance of analysing an existing DTM to determine which areas are most likely to be affected.

Digital Terrain Models (DTM) can be used to derive a wealth of information about the morphology of a land surface. The algorithms traditionally included in most raster processing systems use neighbourhood operations to calculate slope, aspect, and shaded relief. While watersheds are closely related to slope and aspect information they also present non-neighbourhood problems such as determining direction of flow in the interior of a large flat area. Software has been developed that uses neighbourhood techniques as well as iterative spatial techniques that can best be visualised as region-growing procedures. They also provide an analyst with the ability

to extract from DTMs information on morphologic features especially topographic depressions and flow directions that may be further processed to delineate watersheds (Jenson *et al.*, 1988).

A cell that is equal in elevation to all its neighbours' elevations can be classified as belonging to a flat area. Previous research has recognised that depressions, areas surrounded by higher elevation values, in the DTM are the nemesis of determining hydrological flow directions (Jenson *et al*, 1988). This is because the depressions must fill before the flow can continue. Some depressions, however, are data errors introduced in the surface generation process while others are real topographic features such as natural potholes or quarries. Researchers have attempted to remove the depressions by smoothing the DTM but smoothing removes shallow depressions and deeper ones remain. Another approach is to 'fill' the depressions by increasing the values of cells in each depression to the value of the cell with the lowest value on the depressions boundary.

An initial conditioning phase produces three data sets that are of general utility for all subsequent steps. The three data sets are a depressionless DTM, a flow direction DTM and a flow accumulation DTM. Figure 5.1 shows the original DTM of Marconi Beam.

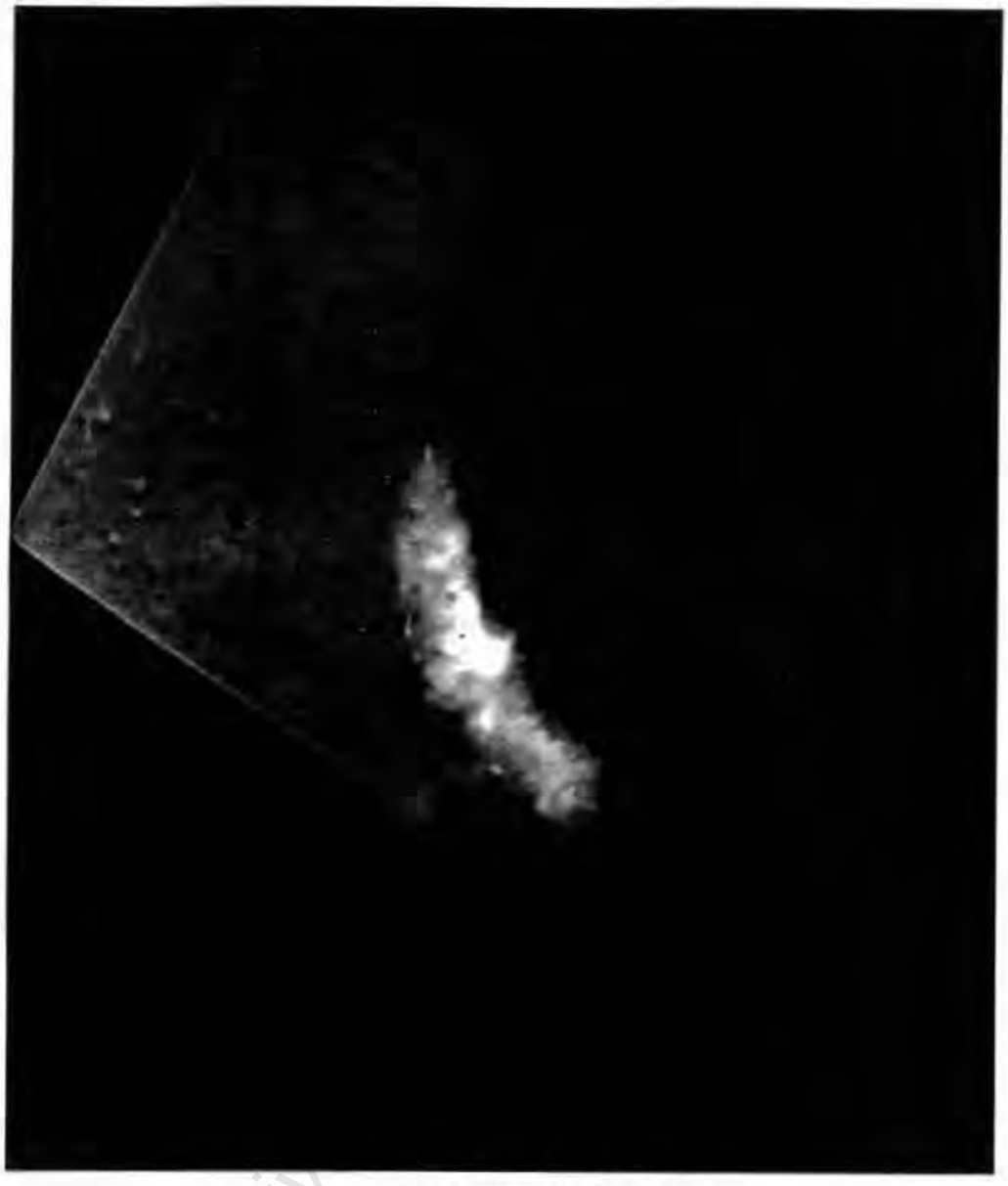


Fig.5.1 Original DTM

5.2 Surface characteristics from DTM

ESRI (Environmental Sciences Research Institute, 1990) presented the following diagram (fig.5.2) for a successful surface hydrological analysis.

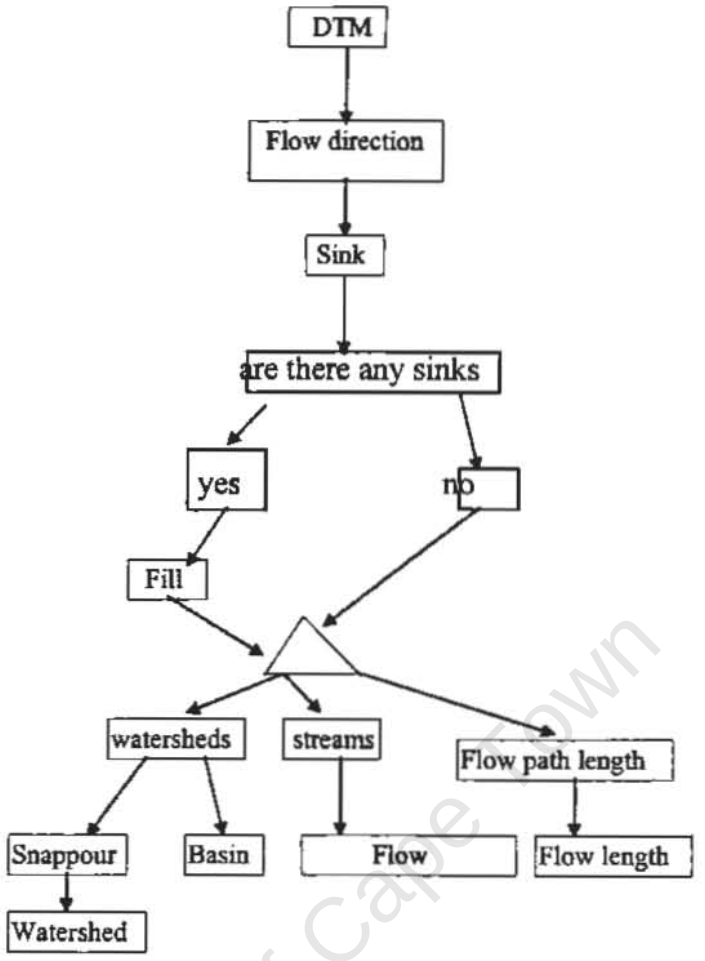
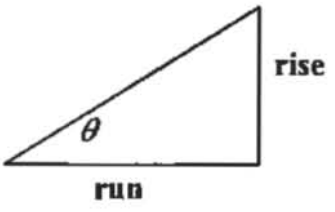


Fig.5.2 Hydrological analysis

A DTM can also be used to derive other surface grids such as slope and aspect grids (fig. 5.3 and 5.4 respectively) for an understanding of the surface. Slope identifies the maximum rate of change in value from each cell to its neighbours. An output slope grid can be calculated as percent slope or degree of slope. In this study we use the degree of slope.



$\tan\theta$ = rise / run

degree of slope = θ

percent of slope = (rise / run) * 100

e.g. degree of slope = 30 degrees

percent of slope = 58 percent

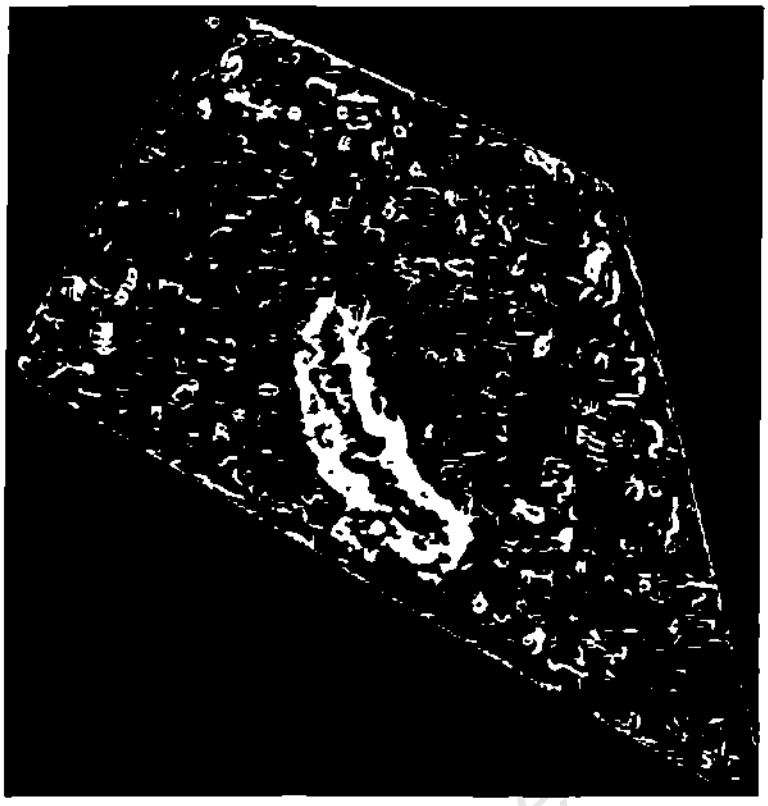
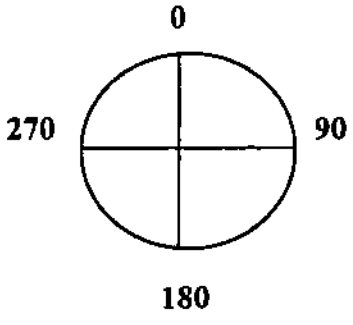


Fig. 5.3 Slope grid

The slope grid of Fig. 5.3 shows that the brighter parts of the grid have higher slope values than the darker parts of the grid. Thus the darker parts are flatter than the bright parts. The bright parts have slope values higher than 10 degrees while the darker parts are lower than 10 degrees.

Aspect identifies the down-slope direction of the maximum rate of change in value from each cell to its neighbours. Aspect can also be thought of as the slope direction. The values of the output grid will be the compass direction of the aspect.



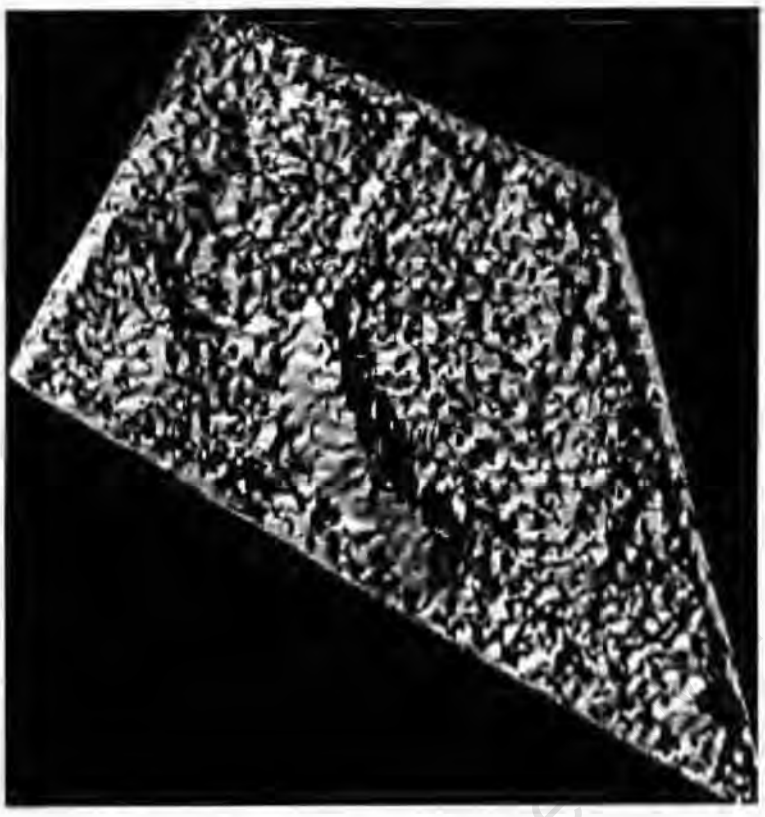


Fig. 5.4 Aspect grid

In Fig. 5.4 the bright parts are in the range between 180 to 360 degrees while the darker parts range from 0 to 180 degrees. Thus the bright parts are in the third and fourth quadrant while the dark parts are in the first and third quadrant.

5.3 Filling depressions in a DTM

DTMs almost always contain depressions that hinder the flow routing. The aim of the this step is to create a depressionless DTM in which the cells contained in depressions are raised to the lowest elevation value on the boundary of the depression. Each cell in the depressionless DTM will then be part of a monotonically decreasing path of cells leading to an edge of the data set.

A DTM free of sinks, a depressionless DTM (fig.5.5), is the desired input to the FLOWDIRECTION function in Arc/Info. The presence of sinks may result in an erroneous flow-direction grid. Since FLOWDIRECTION is the first step in deriving hydrological information about a surface, its input should be as accurate as possible. When a sink is filled using the FILL command in Arc/Info, it is filled to its pour

point, the minimum elevation along its watershed boundary. The FILL command takes a grid as its input.

Fig.5.5 Depressionless DTM

5.4 Flow directions

One of the keys to deriving hydrological characteristics about a surface is the ability to determine the direction of flow from every cell in the grid. This is done with the FLOWDIRECTION function. This function takes a surface grid as input and outputs a grid showing the direction of flow out of each cell. There are eight valid output directions relating to the eight adjacent cells into which flow could travel as shown in table 1 :

32	64	128
16	x	1
8	4	2

Table.1 Flow directions

In table 1 x is the processing cell. The north and south directions are represented by 64 and 4 respectively while 16 and 1 represent west and east directions respectively. Also north-west and south-east directions are represented by 32 and 2 respectively.

The direction of flow is determined by finding the direction of steepest descent or maximum drop, from each cell. This is calculated as

Maximum drop = change in Z-value / distance

The distance is determined between cell centres. When the direction of steepest descent is found, the output cells are coded with the value representing that direction.

There are four possible conditions to consider in determining flow direction. Condition 1 occurs when all eight neighbouring cells have values higher than the processing cell (cell x). The flow direction will be encoded as negative for such a cell indicating an undefined flow direction. Condition 1 represents single-cell depressions. These single- cell depressions will not be present after the sinks have been filled. Condition 2 is the case where the distance-weighted drop from the centre cell is higher for one cell in the neighbourhood over all of the other seven and the flow direction is assigned to that cell. Most cell are condition 2 cells. When two or more cells are equal in having the highest distance-weighted drop, the flow direction is assigned logically using a table look-up operation. For example, if three cells along the one edge of the neighbourhood have equal drops, the centre cell is logically assigned the flow direction. If two cells on opposite sides have equal drops one is arbitrarily chosen Condition 4 occurs when all cells are higher or equal to the centre cell. In this case the cell is located in a flat area and the flow direction is not known. Then the flow direction is calculated as the sum of the lowest multiple directions (Jenson *et al.*, 1988). The flow directions and the number of cells flowing in that direction resulting from processing the Marconi Beam DTM are presented in table 2 and a flow direction diagram in fig. 5.6. :

RECORD	VALUE	COUNT
1	1	242015
2	2	121857
3	4	206427
4	8	124029
5	16	222359
6	32	129658
7	64	284255
8	128	189034

Table.2 Flow direction and count for a depressionless DTM

Table 2 lists the results of the flow direction data set. The value determines the direction of flow while the count gives the number of cells flowing in that direction. This table is used to determine the direction of flow of water across a surface. From table 2 it can be observed that most cells have a North flow direction (value 64), followed by cells in the East direction (value 1) while the least cells flow in the South West direction (value 8).

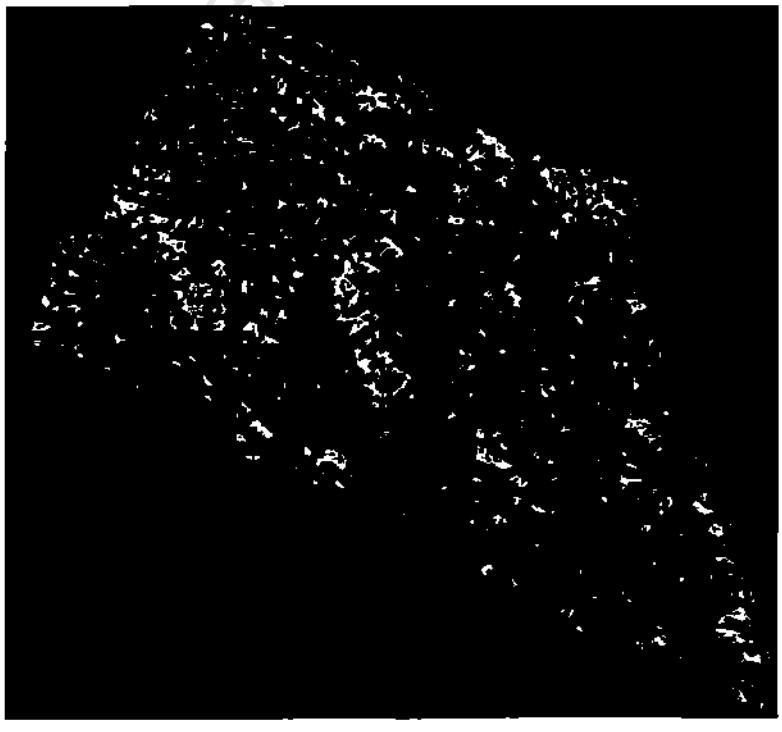


Fig.5.6 Flow direction diagram

5.5 Flow Accumulation

This is the third procedure in the conditioning phase and makes use of the flow direction data to create a flow accumulation data set. Here each cell is assigned a value equal to the number of cells that flow to it. Cells having a flow accumulation value of zero, to which no other cell flows, correspond to ridges. Because cells in a depressionless DTM have a path to the data set edge, the pattern formed by highlighting cells with values higher than some threshold value delineate a fully connected drainage network (see fig.5.7). Because we want to see how water behaves in this DTM surface, and not trying to delineate specific features such as dams, no threshold was set in this work.

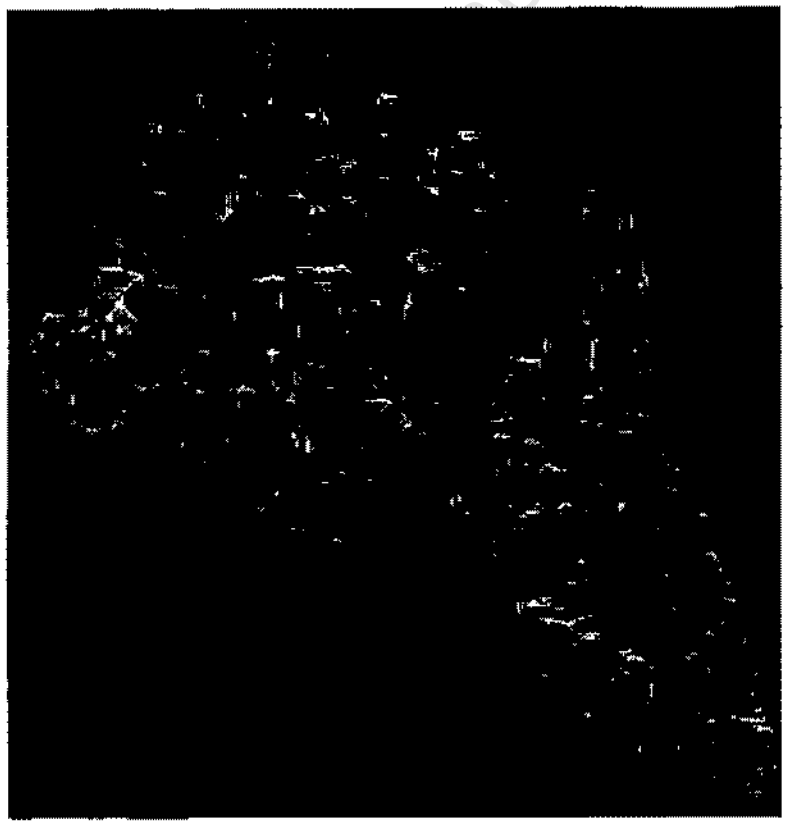


Fig.5.7 Flow Accumulation

The flow accumulation grid delineates a drainage network. The brighter parts in the depressionless DTM (fig.5.5) are higher than their surrounding areas. Hence these

higher parts in fig.5.5 flow into their neighbouring cells in the flow accumulation DTM (fig.5.7). The brighter parts in fig.5.7 delineate a drainage network.

5.6 Watersheds

This is an important application of hydrological analysis. A watershed is the up slope contributing flow to a given location. It is also referred to as a basin, catchment, sub-watershed or contributing area. It is an area that drains water and other substances to a common outlet as concentrated drainage. The boundary between watersheds is referred to as a drainage divide. Watersheds together with a drainage network are important in determining areas which are likely to flood and the design of drainage systems etc. The determined catchment areas are illustrated in fig.5.8.

Fig.5.8 Catchment areas

The watersheds in fig. 5.8 delineate areas that drain water to a common outlet. The higher the area the less likely it is to flood. In fig.5.8 the brighter areas are higher than the darker areas. The dark areas are therefore the ones most likely to flood. To determine the magnitude of flooding one needs a rainfall data.

5.7 Chapter Summary

This chapter investigates the use of DTM in hydrological analysis. It starts by presenting a flow diagram for the hydrological analysis. Then the steps required for successful analysis are presented. The first step is to fill the DTM because the depressions in the DTM hinder flow routing. Flow direction data is then computed to show the direction of flow from every cell in the grid. Then a flow accumulation data set is produced to determine the number of cells flowing into each cell in the grid. A watershed data set determines areas which are most likely to flood.

The flow of water across a surface is determined by the shape of the surface it flows on. Drainage networks and height, timing and inundation of floods can be determined using flow accumulation and watersheds respectively. For hydrological analysis to succeed, the depressions in a DTM must be filled. This is because depressions in a DTM hinder the flow of water across its surface. Slope and aspect grids also provide valuable insight into the nature of the surface. Slope determines the rate of change in elevation while aspect determines the slope direction.

Three grids are important in the hydrological analysis, namely a depressionless DTM, a flow direction grid and a flow accumulation grid. In a depressionless DTM, the depressions are filled to allow for a continuous flow of water across the surface of the DTM. The flow direction grid determines the direction of flow of water across the DTM and outputs the number of cells that flow into a specific direction. There are eight directions of flow as depicted in table 1. The flow accumulation grid determines the number of cells flowing into a processing cell. This grid also enables the delineation of drainage networks. Watersheds derived from the DTM determine flood zones. These zones determine the areas that draw water to a common output.

CHAPTER 6

CONCLUSIONS AND FUTURE OUTLOOK

In this thesis I have presented a method for automatic detection of shacks from a DSM in an informal settlement as well as for automatic detection of buildings from a DSM in a formal settlement. In the informal settlement I have focused on the use of an available DTM and shadow data to aid the shack detection. Also the use of a DSM to directly detect buildings/shacks in both settlements has been presented. I have demonstrated that shack detection in informal settlements can be achieved by using a DSM, a DTM and shadow data. Also a DSM suffices in the detection of buildings in a formal settlement.

The thesis is structured into four main parts, namely, image matching to derive the data sets, building detection to delineate buildings from other man-made objects in DSM (digital surface model), DSM quality analysis to determine the reliability of the data, hydrological analysis to determine flood zones as an additional example of DTM application.

Two data sets were used to test the strategies for building detection, namely a formal and an informal settlement.

6.1 Informal settlement

6.1.1 Raster approach

The raster approach is based on the automatic extraction of simple shacks, e.g., 4-sided shacks using multiple cues. These cues are edge contours, DSM, shadows and a DTM. The first step in this approach is the normalisation of the DSM. This means the subtraction of a DTM from the DSM. Normalisation puts the DSM on a reference surface which is approximately a plane. Thresholding the DSM is done to ensure that we get only information above the terrain. Due to the limitations of image matching the shacks are connected in the DSM, hence the need to use shadows to segment the DSM blobs. The boundaries of the shack blobs are further segmented by filtering and then the shack outlines are extracted.

The detected blobs have to be verified as shacks. This is achieved by using the histogram of gradient orientations. For a building the histogram of gradient orientations will contain peaks which are 90 degrees apart (see fig. 3.12), while for a tree they are not as significant. In this informal settlement there are no trees so we verify only the shacks. Then a decision is taken as to whether the detected blob is indeed a shack. An overlay of the detected blobs on manually measured CAD data also verifies the blobs (fig. 3.13).

6.1.2 Vector approach

This approach uses a DSM to directly detect shacks. The DSM is first transformed into contours. Buildings/shacks are characterised by closed contours - polygons. Thus, the polygons have to extracted. Figure 3.15 depicts the shack polygons extracted from the contours. Based on the results it is clear that individual shacks cannot be extracted because the polygons are not closed.

Another approach is to use the rasher approach's result in a vector approach (hybrid method). The detected shacks are transformed into polygons as shown in figure 3.16. From the figure one can see that the shacks are merged together and hence groups of shacks can be detected. This is because the shacks are close to each other and contouring fuses them together.

Based on the discussion above we conclude that the raster approach works well while the vector approach does not work for this settlement.

6.2 Formal settlement

The strategy used in this work is similar to the vector approach above. The DSM is converted into contours and polygons are extracted. Figure 3.18 depicts the polygons of a formal settlement. These polygons are closed. So we can use context rules to detect building as described in section 3.6.2. Figure 3.21 depicts the blob outlines which need to be verified.

Verification, as described above, is achieved by histograms of gradient orientations. Figure 3.22 shows the blob outlines overlaid on the DSM. The blob outlines fit the areas where buildings are expected. Figure 3.24 shows the histogram for a formal settlement. The histogram shows strong gradients 90 degrees apart. The gradient for a tree are weaker and most seem to be uniform as shown in figure 3.25. This approach works well for this settlement.

6.3 Quality analysis

The quality analysis employed in this work was achieved by a parametric statistical test and a visual analysis. The statistical test is based on a student t-test. The test takes the standard deviation and the arithmetic mean of the differences between the original values and the estimated values of the DSM and computes a standardised standard deviation. This standardised deviation is compared to the tabulated value of the student t-test (3.219). Values higher than this value are considered erroneous. The Wallis enhanced DSM in figure 4.3 has less erroneous data than the Median enhanced DSM in figure 4.4.

Visual analysis employs histograms , scattergrams and overlays. Comparison of the histograms reveals that the Wallis DSM histogram in figure 4.8 is smoother than the Median DSM histogram in figure 4.9. The scattergram also reveals that the Wallis-ground truth distribution of values has less outliers than the Median-ground truth. The overlays are not useful because the two DSMs have the same visual appearance.

Based on the visual and statistical tests one can conclude that the Wallis filter is a better smoothing function than the Median filter. Thus the Wallis enhanced DSM is preferred to the Median enhanced DSM.

6.4 Hydrological analysis

- The shape of a surface determines how water will flow across it. This valuable information can be obtained from the slope and aspect data derived from the original DTM. These two data sets show the rate of change in elevation and the slope
  direction.
- A depressionless DTM is essential because depressions in a DTM hinder the flow of water. This DTM is used as a first input in hydrological analysis.
- A flow direction grid determines the number of cells that flow into a specific direction. A flow accumulation grid determines the number of cells that flow into a processing cell, hence it delineates a drainage network. Watersheds are used to determine areas that drain water into a common outlet and thus areas which are likely to flood. To determine the magnitude of the flood we need to have rainfall data.

Possible future work could be undertaken to achieve the following :

(1) The extent to which building density patterns can be used to predict the presence of other buildings using probabilistic techniques.

(2) A method for identifying densification changes in informal settlements based on subtracting two DSMs derived from different epochs and classifying topographic differences.

(3) Where the surface modelling of a building by a DSM is sufficiently accurate, coarse classification of building geometry should be investigated with the view of automating building geometry modelling from DSMs.

BIBLIOGRAPHY

Baltsavias, E. P., Mason, S. and Stallmann, D., 1995. Use of DTM/DSMs Orthoimages to Support Building Extraction. In Grun, A., Kubler, O. and Agoris, P.(eds) : *Automatic Extraction of Man-Made objects from Aerial and Space Images*, Birkhauser Verlag, Basel :199-210.

Baltsavias,E. P., 1991. Multiphoto Geometrically Constrained Matching. PhD. Thesis, Institut für Geodäsie und Photogrammetrie, nr 49, Zurich.

Bishop, L., Mason, S. 1997. Emerging Technologies for Development of an Interative Decision Environment, *5th intl.conf.on Computers in Urban Planning and Urban Management*,Bombay,12 pages.

Bolstad, P.V., Stowe, T. 1994. An Evaluation of DEM Accuracy: Elavation, Slope, and Aspect. *Photogrammetric Engineering and Remote Sensing,* 60(11) :1327-1331.

Burrough, P.A., 1986. Principle of GIS for Land Resource Assessment. Oxford University Press, New York.

Collins, S.H., and Moon, G.C., 1981. Algorithms for Dense Digital Terrain Models. *Photogrammetric Engineering and Remote Sensing*, 47(1) :71-76.

Eckstein, W., 1996. Segmentation and Texture Analysis. *Proceedings of the 18th ISPRS Congress, Commission 3*, Vienna: 165-173.

ESRI, 1990. Environmental Systems Research Institute. Redlands, California.

Fayek, R.E., 1996. Preserving Topography in 3D Data Compression for Shape Recognition. *Proceedings of the 18th ISPRS Congress, Commission 3*, Vienna: 186-191.

Felicísimo, A.M., 1994. Parametric Statistical Method for Error Detection in Digital Elevation Models. *ISPRS Journal of Photogrammetry and Remote Sensing,* 49(4):29-33.

Förstner, W. and Weidner, U., 1995. Towards Automatic Building Extraction from High Resolution DEM, *ISPRS Journal of Photogrammetry and Remote Sensing,* 50(4): 38-49.

Haala, N. and Anders, K., 1996. Fusion of 2D-GIS and Image Data for 3D Reconstruction. *Proceedings of the 18th ISPRS Congress, Commission 3,* Vienna: 285-290.

Haala, N. and Hahn, M., 1995. Data Fusion for the Detection and Reconstruction of Buildings. In Grun, A., Kubler, O. and Agoris, P.(eds) : *Automatic Extraction of Man-Made objects from Aerial and Space Images,* Birkhauser Verlag, Basel: 211-220.

Hannah, M.J., 1981. Error detection and Correction in Digital Terrain Models. *Photogrammetric Engineering and Remote Sensing,* 47(1) : 63-69.

Hartl, Ph. and Cheng, F., 1995. Delimiting the building heights in a city from the shadow on a panchromatic SPOT-image:part2. *International Journal of Remote Sensing,*16(15):2829-2842.

Hearnshaw, H.M., and Urwin, D-J., 1994 Visualization in GIS. Wiley and Sons, New York.

Heipke, C., 1996. Overview of Image Matching Algorithms. In Kolbl, O.(eds): *Proceedings of the OEEPE-workshop on applications of Digital Photogrammetric Workstations,* Lausanne: 173-189.

Henricson, O., Bignone, F., Willuhn, W., Ade, F., Kübler, O.,Baltsavias, E., Mason, S., Grün, A., 1996. Project Amobe, Strategies, Current Status, and Future Work. *Proceedings of the 18th ISPRS Congress, Commission 3,* Vienna:321-330.

Jenson, S.K., and Dominique, J.O., 1988. Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis. *Photogrammetric Engineering and Remote Sensing*, 54(11): 1593-1600.

Kienzle, S.W., 1994. The Application of DTMs for Slope Analysis and Improved Soil Delineation. *In Terrain Evaluation and Data Storage including a Seminar on Remote Sensing Techniques " 4th Symposium"*, Escom Conference Centre-Midrand, Johannesburg.

Keates, J. K., 1973. Cartographic Design and Production. Lowe and Brydone Printers Limited, Thetford, UK.

Kim, T. and Muller, J-P., 1995. Effects of Image Resolution on an Automated Building Extraction System. *Proceedings of the 21st Annual Conference of the Remote Sensing Society on Remote Sensing in Action(RSS95)*,Southampton,UK:11-14.

Lee, D-C., and Shenk, T.,1992. Image Segmentation From Texture Measurement. *Proceedings of the 17th ISPRS congress, Commission 3*, Washington:75-80.

Li, J., Mason, S. and Rüther, H., 1997. Experiences in Automated Shack Extraction using Multispectral Image Classification, iKusasa/Consas'97, August 24-28, Durban:10p.

Li, Z., 1994. A Comparative Study of the Accuracy of Digital Terrain Models based on Various Data Models. *ISPRS. J. of Photogrammetry and Remote Sensing*, 49(1); 2-13.

Liang, T. and Heipke, C., 1996. Automatic Relative Orientation of Aerial Images. *Photogrammetric Engineering and Remote Sensing*,62(1):47-55.

Mason, S. 1996. 3D building reconstruction Using Composites of Surface Primitives: *Concept, International Archives of Photogrammetry and Remote Sensing*, 31B(3): 517-522.

Mason, S. 1996. Automation in Image-based Mensuration: Advances in Digital Photogrammetry, SA. J. Science, 92:408-410.

Mason, S. 1997. Digital Still Video Mapping of Informal Settlements for Urban Upgrading, Africa GIS ' 97, Gabarone.

Mason, S.,Barry, M., 1997. Managing Informal Settlements Spatially: Experiences, Africa GIS ' 97, Gabarone.

Mason, S., Baltsavias, E., 1997. Image-Based Reconstruction of Informal Settlements, in Grün, A., Baltsavias, E., Henricson, O. (eds), *Automatic Extraction of Man-Made objects from Aerial and Space Images 2*, Birkhaeuser Verlag, Basel,(to be published).

Mason, S., Baltsavias, E, and Stallman,D., 1994. High Precision Photogrammetric Data Set for Building Reconstruction and Terrain Modelling. *Internal Report, Institute of Photogrammetry and Geodesy*, Zurich.

Mason, S., Rüther, H., 1997. Managing Informal Settlements Spatially:, *Proc. ASPRS Annual Meeting*, Seattle.

Mason, S. and Streilein, A., 1996. Photogrammetric Reconstruction of Buildings for 3D City Models. *SA. J. of Surveying and Mapping*, 23(5): 1-14.

Mitchell, C., 1991. Terrain Evaluation. Wiley and Sons, New York.

Novak, K., and Sperry, S.L., 1992. Intergration of Digital Photogrammetry and Raster GIS. *Proceedings of the 17th ISPRS congress, Commission 4*, washington:894-898.

Paško, M., and Gruber, M., 1996. Fusion of 2D GIS Data and Aerial Images for 3D Building Reconstruction. *Proceedings of the 18th ISPRS Congress, Commission 3,* Vienna: 257-260.

Polidori, L., Chorowicz, J., Guillande, R., 1991. Description of Terrain as a Fractal Surface, and Application to Digital Elevation Model Quality Assessment. *Photogrammetric Engineering and Remote Sensing*,57(10):1329-1332.

Schenk, T. and Wang, Z., 1992. 3D Urban Area Surface Analysis. In *Research Activities in Digital Photogrammetry at The Ohio State University*: 65-70.

Sinning-Meister, M., Gruen, A. and Dan, H., 1996. 3D City Models for CAAD-Supported Analysis and Design of Urban Areas. *ISPRS. J. of Photogrammetry and Remote Sensing*, 51(4) : 196-208.

Sonka, M., Hlavac, V. and Boyle, R., 1993. Image Processing, Analysis and Machine Vision. International Thomson Computer Press, London.

Tomlin, C.D., 1990. Geographical Information Systems and Cartographic Modelling. Prentice Hall, Englewood Cliffs, USA.

APPENDIX A

AML FOR AMOBE PROJECT

```
latticecontour dsm_gd cont1 1
build cont1 poly
arclabel cont1 cont1 right 0.5 contour
build cont1 poly
describe cont1
ae
mape cont1
edit cont1
ef polys
sel all
put contpoly1
q
build contpoly1 poly
ae
mape contpoly1
edit contpoly1
ef polys
sel all
select contpoly1-id = 0
selectput arc
ef arcs
delete
build
save
ef polys
sel all
additem elongation 10 10 n 5
save
ef polys
sel all
calculate elongation = area / perimeter
save
ef polys
sel all
sel perimeter > 40
put peri1
save
ef polys
sel all
sel area > 20
ef polys
put out
q
build out poly
ae
mape out
edit out
ef polys
sel all
sel elongation < 0.6
ef poly
put out2
```

```
q
build out2 polys
ae
mape out2
edit out2
ef polys
sel elongation = 0
ef poly
selectput arcs
ef arcs
delete
save
q
build out2 poly
ae
mape out2
edit out2
ef labels
sel all
list
calc out2-id = 1
save
q
build out2 poly
dissolve out2 out2dis out2-id
build out2 line
identity out2 out2dis out2ids line
frequency out2ids.aat out2e.frq
contour
out2dis-id
~
y
y
~
y
n
~
statistics out2e.frq out2e.stat out2dis-id
maximum contour
minimum contour
~
n
n
ae
mape out2
ef arcs
sel all
sel lpoly# = 1
put blobs
q
build blobs poly
ae
mape blobs
edit blobs
q
&watch &off
```

APPENDIX B
AML FOR MARCONI BEAM PROJECT

```
grid
sewindow a11enhg
setwindow
setwindow a11enhg
dtm25g = dtm25enh
dtmenhg = dtm25g
mape dtmenhg
image dtmenhg
clear
enhdifg = a11enhg - dtmenhg
mape enhdifg
image enhdifg
cellvalue enhdifg *
select
enh15g = select (enhdifg , 'value > 1.5')
clear
image enh15g
cellvalue enh15g *
clear
image a11bwshadg
cellvvalue a11bwshadg *
cellvalue a11bwshadg *
a11bwsel = select (a11bwshadg , 'value = 1')
clear
image a11bwsel
cellvalue a11bwsel *
clear
image enh15g
cellvalue enh15g *
outenhg = con(enh15g > 0 , 1)
clear
image outenhg
cellvalue outenhg *
a11bwnd = con(isnull(a11bwsel),0,a11bwsel)
clear
image a11bwnd
cellvalue a11bwnd *
clear
outenhnd = con(isnull(outenhg),0,outenhg)
image outenhnd
cellvalue outenhnd *
clear
outblobs = outenhnd - a11bwnd
image outblobs
cellvalue outblobs *
boundaryclean
try 1 = boundaryclean (outblobs,descend,twoway)
try1 = boundaryclean (outblobs,descend,twoway)
clear
image try1
clear
try = boundaryclean (outblobs,descend,oneway)
image try
```

```
kill try1
clear
image try
cellvalue try *
shacks = select (try,'value = 1')
clear
image shacks
gridpoly
shackpoly = gridpoly (shacks)
linecolor red
arcs shackpoly
clear
q
ae
edit shakpoly
drawe labels off
drawe nodes off
draw
edit shackpoly
drawe labels off
drawe nodes off
draw
q
describe shackply
describe shackpoly
ae
edit shackpoly
drawe labels off
drawe nodes off
draw
ef poly
list
sel all
list
sel many
list
unsel all
ef poly
sel area < 17
list
ef poly
unsel all
sel many
list
unsel all
sel area < 26
ef poly
unsel all
sel perimeter < 50
ef poly
unsel all
sel area < 17
ef poly
selectput arcs
ef arcs
delete
build
```

```
q
n
build poly
build shackpoly poly
ae
edit shackpoly
drawe labels off
drawe nodes off
draw
ef polys
sel area < 17
selectput arc
ef arc
delete
save
q
build shackpoly poly
ae
edit shackpoly
drawe labels off
drawe nodes off
draw
q
end
quiet
end.
&watch&off
&watch &off
```