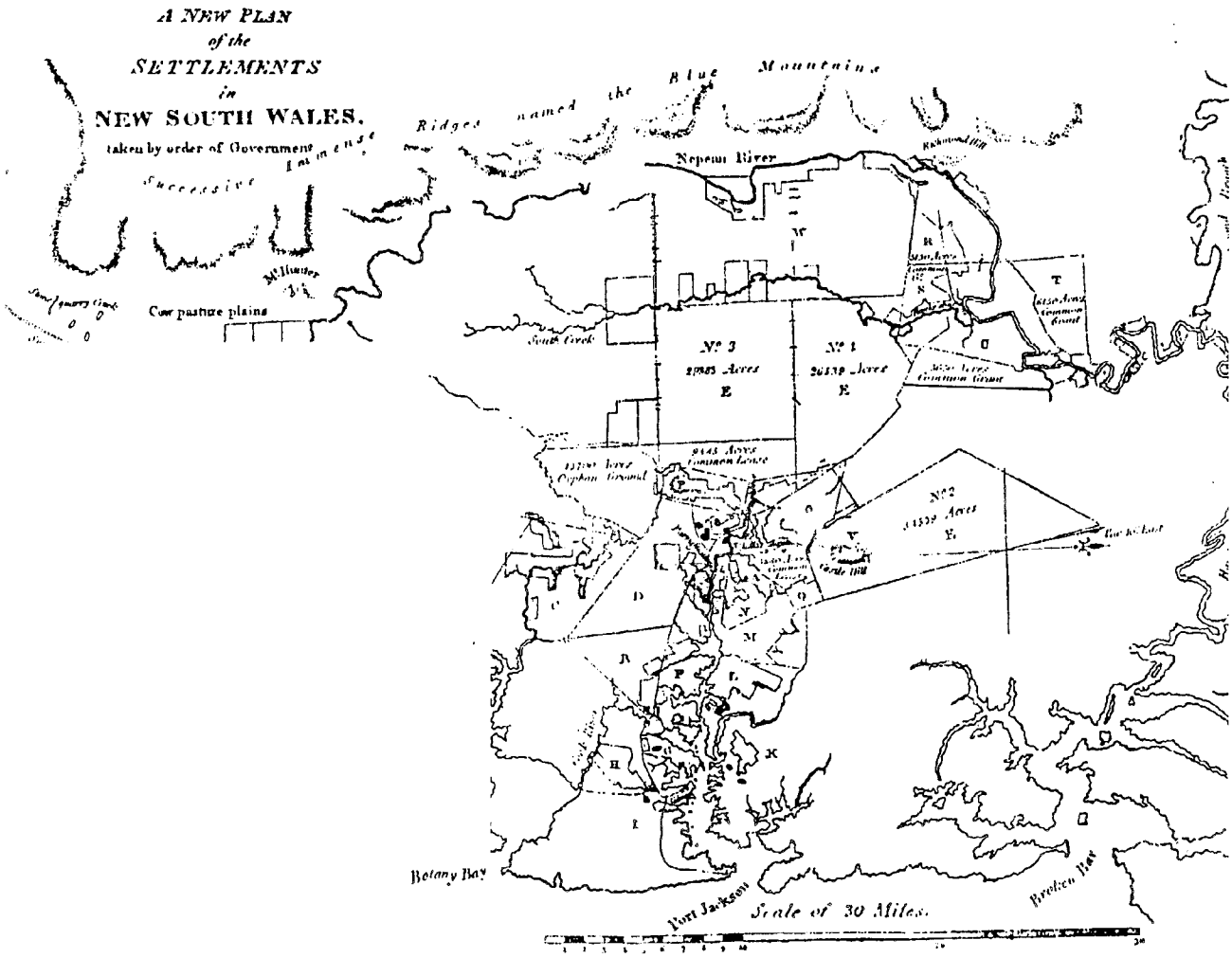
AUTOMATIC BUILDING EXTRACTION FOR 3D TERRAIN RECONSTRUCTION USING IMAGE INTERPRETATION TECHNIQUES

YI HUI LU



UNISURV S-75, 2004

Reports from

SCHOOL OF SURVEYING
AND SPATIAL INFORMATION SYSTEMS

THE UNIVERSITY OF NEW SOUTH WALES UNSW SYDNEY NSW 2052 AUSTRALIA

UNISURV REPORT S-75, 2004

AUTOMATIC BUILDING EXTRACTION FOR 3D TERRAIN RECONSTRUCTION USING IMAGE INTERPRETATION TECHNIQUES

YI HUI LU

Received: May 2004
Accepted: June 2004

SCHOOL OF SURVEYING AND SPATIAL INFORMATION SYSTEMS
(formerly GEOMATIC ENGINEERING)
UNIVERSITY OF NEW SOUTH WALES
UNSW SYDNEY NSW 2052
AUSTRALIA

UNISURV REPORTS

Series Editor: Dr. J. M. Rüeger

National Library of Australia

Card No. and ISBN 0 - 7334 - 2141 - 5

Foreword

The determination of elevations from overlapping aerial images by image matching is an operational procedure in most digital photogrammetric workstations. However, the computed elevations are usually subject to errors caused by matching on objects other than the terrain surface, such as trees and buildings. While software is available for modelling some of these errors, most of them must be corrected manually. This can be a time-consuming process which reduces the efficiency of digital photogrammetry. There is a need to be able to interpret the content of the images so that terrain cover can be determined, and if this cover is elevated above the terrain surface, to correct the computed elevations so that bare earth elevations can be derived. Therefore, an intermediate step in the process of determining bare earth elevations involves extracting features in the images such as buildings and trees. This thesis describes the process of extracting buildings. The School of Surveying and Spatial Information Systems at UNSW has undertaken research on the determination of elevations from overlapping space and aerial images, since the early 1990s. The research described in this thesis has been partly supported by an ARC Large ARC Grants A49700856 in 1997-1999 and A10007004 in 2000-2002, in collaboration with Kurt Kubik, from the University of Queensland. Future work in this research involves the fusion of elevation information extracted from laser scanners and aerial images

This thesis presents a number of components in the extraction of buildings from aerial images. They include the determination of a dense digital surface model (DSM) by stereo image matching, a multi-spectral image classification, and Normalised Difference Vegetation Index (NDVI) computation on the images to determination vegetation areas. A shape modelling algorithm based, on the level set formulation of curve and surface motion, has then been developed to precisely delineate the building boundaries. These processes have derived three types of information which describe the locations of buildings, namely, the DSM, areas determined by the classifications, and the building boundaries. These three sets of information were then combined by a data fusion approach using the Dempster-Shafer algorithm, to extract the most likely buildings in the

images. The Dempster-Shafer algorithm is a statistical approach providing a theoretical basis for evaluating the reliability of the extracted buildings from the combination of the different data sources.

The methods have been tested on a series of large scale colour aerial photographs in Australia. Because large scale images covering the infrared region were not available, the computation of the NDVI was replaced by the VVI (visible vegetation index), based on the colour images. The tests demonstrate accuracies of extracting buildings of the order of 90% for a range of areas covered by the aerial images. While the methods achieve satisfactory results on most of the samples tested, there are some examples where the algorithms were below expectations. Further refinements of the methods may lead to better results and overcome some of the deficiencies.

The original thesis was presented with many colour images, which cannot be reproduced in this volume because of the cost. These images are produced in black-and-white in the body of the volume, but the colour versions of the images are contained on the CD stored at the back of this volume. The work of Mr Brian Donnelly from the School of Surveying and SIS in producing the CD is acknowledged.

Emeritus Professor John C. Trinder
June 2004

ACKNOWLEDGMENTS

My first thanks goes to Professor John Trinder for his supervision of this work. His advice, support and encouragement throughout the course of this are greatly appreciated. I am also grateful to his careful reading and correction of my thesis.

I wish to thank Professor Kurt Kubik for his supervision and advice when I studied in University of Queensland. I also wish to thank Mr. Brian Donnelly and the administrative staff in the School of Surveying and Spatial Information Systems for their assistance.

The financial support provided by the Australia Postgraduate Awards and Australia Research Council (ARC) Fund is gratefully acknowledged.

Finally, I wish to thank my colleagues and other postgraduates for their friendship and help.

Abstract

The production of Digital Terrain Models (DTM) from images has been an active research topic in computer vision and digital photogrammetry. DTM are often used as a component in complex Geographic Information Systems (GIS) modelling, and also used for cartography, urban and town planning. Conventional image matching techniques are unable to overcome the disparity discontinuities in the stereo model caused by man-made structures. As well they supply a Digital Surface Model (DSM), because the matching occurs on the top of man-made objects such as buildings, or on the top of the vegetation rather than the terrain surface. In order to produce a more accurate DTM of the bare earth, the characteristics of the terrain cover, such as buildings and trees must to be determined, to reduce the elevations derived from image matching to the terrain surface.

An automatic approach and strategy for extracting building information from aerial images using combined image analysis and understanding techniques is described in this thesis. For a more accurate DTM, the approach involves the interpretation of the images to provide greater reliability in eliminating the errors due to ground cover.

The thesis begins with the discussion on deriving building interest areas from complex scenes in urban areas using unsupervised classification, stereo image matching and texture segmentation methods. Based on the building interest areas, shape modelling using level set formulation of curve and surface motion is used in order to precisely delineate the boundary of building.

In order to provide more reliability in the automatic extracting building system, the formal treatment of the combination of multiple data sources and of the uncertainty in the image interpretation is desired. Data fusion in remote sensing is undertaken to combine several image data sets for the purpose of information extraction. The Dempster-Shafer approach is a statistically-based classification algorithm used for data fusion. It can be used to combine the information from several data sources of the same region, and to find the intersection of propositions on extracted information derived from these datasets,

together with the associated probabilities. An uncertainty estimation, based on Dempster-Shafer evidence theory, is presented which combines evidence from three data sources to effectively guide the segmentation to a suitable solution. A number of test areas have been investigated, the results of the study for extracting buildings are very encouraging and can result in the determination of more accurate elevations of the terrain surface.

Contents

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Background

One of the most challenging problems in the fields of computer vision and digital photogrammetry is 3D reconstruction of the terrain surface from complex aerial images in urban or suburban areas where buildings, roads, trees and vegetations are intermingled in an intricate and complex fashion. Digital Terrain Models (DTM) produced by stereo image matching algorithms have been one of the primary goals of cartography for many years. Recently, interest in the area has been stimulated by the need for digital orthophotos, 3D city models, 3D building reconstruction, production and management of 3D databases for urban and town planning and Geographic Information Systems (GIS) modelling. DTM in this thesis is defined as the elevations of the bare earth, which excludes the effects of man-made structures and vegetation, whereas a Digital Surface Model (DSM) represents the elevations of the visible surface, including man made structures.

Stereo image matching involves the matching of corresponding pixels or features in two overlapping images and is the fundamental to digital photogrammetry for elevation determination. Difficulties in stereo image matching in urban areas are caused by the density of man-made structures and vegetation, leading to many hidden areas or occlusions in the images. Therefore, corresponding points often may not exist in the two images and/or the geometry of the images may vary significantly. Difficulties are also caused by the fact that many scene features, such as walls or roads, have a similar appearance in the images, often roof areas are homogeneous, and moving cars may disrupt the images. These characteristics of aerial images of urban scenes cause the matching process to become ambiguous. Because disparity discontinuities are common in the images, the normal methods of ambiguity checking, such as geometric constraints, and order constraints must be used carefully.

Conventional image matching techniques cannot overcome these matching problems and therefore only supply a DSM. This means that matching occurs on the top of man-made objects such as buildings, or on the top of the vegetation rather than the terrain surface and hence the computed elevations are therefore not on the terrain surface. In order to provide a more accurate DTM, the characteristics of the terrain cover, such as buildings and trees must to be determined to predict errors in the computed elevations. Although there are many features on images that may lead to errors in the DTM, the most common features are buildings and trees. There are many types of buildings in urban areas. This study will concentrate on houses because they represent the majority of structures in many urban areas. The final outcome of the extraction of buildings is to derive an accurate DTM of the bare terrain surface.

The problem of building extraction from aerial images can be approached using different methodologies depending on the purpose of the extraction and the available data. Although many automatic building extraction algorithms have been proposed by researchers, there are still no commonly accepted algorithms because each method focuses on a particular application and is dependent on the types of features extracted in

the study. A more detail summary for building extraction methods can be found in Chapter 2.

The purpose of building extraction in this thesis is to interpret the image and identify and locate building areas to guide terrain reconstruction. It is proposed that the methods developed in this study will enable the determination of more accurate elevations of the bare terrain surface. In order to realize these aims, an automatic approach and strategy for extracting building information from aerial images is needed. The contribution of this research is that it integrates compound image analysis and interpreting techniques into a building extraction system. Once building areas are extracted, errors in DSM due to the ground cover can then be eliminated.

1.2 Objectives of The Research

The main objectives of the research are:

- To improve the process of DTM determination by recognizing types of terrain cover and correction of their effects on the computed DSM.

- To develop robust building extraction procedures that include multiple data sources and combine image analysis and interpretation techniques.

- To establish practical and effective modules in the building extraction process including:
    1) an efficient matching algorithm based on least squares and robust estimation, which can achieve greater accuracy in the computation of disparities in overlapping images for areas where discontinuities caused by buildings and trees exist.
    2) algorithms for differentiating buildings from other objects and revealing building areas based on the integration of multiple information sources derived from different image interpretation techniques.

3) an algorithm for extracting building boundaries based on level set shape modelling for curve propagating interfaces. The method is highly robust and accurate for tracking building boundaries in complicated urban areas.
4) the use of multiple information data in automatic building extraction and a mechanism for integration of multiple knowledge. A data fusion method is desired to provide more reliability in the automatic extracting building system.

The scientific significance of the research lies in the treatment of the process of 3D surface reconstruction from stereo images, as an operation of an image understanding and interpretation procedure. This procedure is required to acquire knowledge of the content of the images, since the 3D information cannot be extracted accurately unless the characteristics on the terrain surface are known.

The results of the image interpretation approach can lead to improved 3D reconstruction over current approaches in most existing software packages, which are based on the determination of elevations of regularly spaced points on the terrain. Since these elevation points are usually derived in isolation from one another, no account is taken of the fact that neighbouring elevation points represent a terrain surface that has a particular, but yet to be determined shape. Therefore, if an elevation point occurs on the top of a man-made structure or a tree, there is normally no means of checking whether the height of that point is consistent with the height of the neighbouring bare earth surface. The errors caused by elevations being determined on the tops of buildings and trees are included in DSM. The study aims to provide an accurate DTM of the bare earth, by recognising the characteristics of the terrain cover, such as trees, vegetation, river, road and buildings, and thus reducing the elevations derived from image matching to the terrain surface.

1.3 Organization of the Thesis

The thesis consists of eight chapters and references.

Chapter 2 gives a general overview of the stereo image matching for 3D reconstruction and building extraction techniques. After reviewing and comparing the building reconstruction algorithms, the framework of this research is introduced, comprising the proposed models for the steps in the process of identifying and extraction of building areas.

Chapter 3 introduces a robust hybrid-based image matching algorithm and a least square matching algorithm using robust estimation for detecting breaklines in the images. Using these breaklines will enable greater accuracy in the computation of disparities in the areas comprising discontinuous surfaces and occlusions.

Chapter 4 describes the module of low-level image analysis and interpretation. In this module, the significant features from images are extracted and classified into the meaningful regions. An effective compound image analysis technique, which analyses the relationships between spatial analysis and spectral properties, is proposed to reveal the building interest areas.

In Chapter 5 the automatic extraction of boundaries of building is addressed. The proposed method is based on image segmentation that implements mathematical and numerical shape modelling using level set formulation of curve and surface motion. It offers a highly robust and accurate method for tracking curve interfaces moving under complex motions.

In Chapter 6 the uncertainty estimation is presented. The Dempster-Shafer data fusion technique provides the theoretical basis for evaluating the reliability of the extracted buildings from the combination of the different data sources by a statistically-based classification. This procedure uses the evidence from different data sources to effectively guide the segmentation to a suitable solution.

Results of experiments using these methods on aerial images are given in Chapter 7, for a number of test areas, which include buildings with different sizes, shapes and roof

colour. The tests results are encouraging and demonstrate that the modules of the system are effective procedures for building extraction, and the determination of more accurate elevations of the terrain surface. The test data available is limited to colour aerial photography. Multispectral images with 2 metre and 1 metre resolution were evaluated for the proposed automatic building extraction system. After an analysis, these images proved to be inadequate as it was not possible to generate high accuracy DSM and supply large enough buildings. It was more practical and appropriate to use higher resolution colour aerial images for the system.

The conclusions and perspectives for future studies are given in Chapter 8.

CHAPTER 2

Review of Building Extraction from Stereo Images Matching

In this chapter, stereo image matching algorithms and building extraction methods will be reviewed. After comparing the wide variety of methods, an approach to automatic building extraction for 3D terrain reconstruction using interpretation techniques will be proposed.

2.1 Stereo Image Matching for 3D Reconstruction

Methods of automatically generating digital surface models from overlapping aerial photographs use stereo image matching techniques. Stereo image matching, also known as the correspondence problem, requires the location and matching of corresponding pixels or features in two overlapping images. The 3D world coordinates of the object points represented by the pairs of corresponding pixels in two images can then be determined by triangulation, using the image coordinates and known orientation of the

images (Marr and Poggio 1979, Wolf 1983, Gonzalez and Woods 1992). Many techniques have been explored to solve the stereo image matching problem. Most techniques are optimized for a particular application, such as in manufacturing, robotics, surveying and vehicle guidance.

Image matching techniques may be broadly classified as area-based matching and feature-based matching algorithms (Alwan and Naji 1996, Atkinson 1996, Aschwanden and Guggenbuhl 1993, Hahn 1993, Baltsavias 1991, Gruen and Stallmann 1991, Li 1991, Allision et al. 1991, Brockelband and Tam 1991, Tu and Dubuisson 1990, Hoff and Ahuja 1989, Zhang and Zhou 1989, Levy 1988, Mcintosh and Match 1988, Eastman and Waxman 1987, Gruen and Baltsavias 1987, Foerstner and Guelch 1987, Foerstner 1986, Grimson 1985, 1983, 1981, Ohta and Kanade 1985, Medioni and Nevatia 1985, Haralick 1984, Medioni and Nevatia 1984, Rosenfeld 1984, 1982, Hummel and Zucker 1983). Combinations of these two have also been proposed. A detailed review and comparison of image matching techniques can be found in Mei (2001), Baillard and Dissard (2000), Lang and Foerstner (1995), Brown (1992), Doorn et al. (1990), Day and Muller (1989), Dhond and Aggarwal (1989), Guelch 1988, Hannah (1988, 1989) and Barnard and Fischler (1982).

Each of the following sections describes the techniques developed by various researchers within the particular fields, according to the above three classes. Area-based matching algorithms are mostly used in photogrammetry for smooth terrain surface reconstruction, while feature-based and hybrid matching algorithms are widely applied in computer vision, close range photogrammetry and 3D building reconstruction. Some commercial software systems have been successfully implemented based on the image matching techniques. Area-based matching and relaxation correlation have been used in the LH systems Socet set and VirtuoZo systems. Feature-based matching techniques have been used in the Match-T system, originally implemented by Inpho company.

2.2 Area-based Methods

In area-based image matching techniques, a sample window comprising a rectangular grid of pixels is selected in one image, let us assume it is the left image, and the window is moved over the right image until a match, based on the intensity values of the two windows, is found. This process continues until as many as possible areas in the left image are matched with the corresponding areas in the right image. The matching windows are usually chosen to be square and with a size selected according to data content in the images.

2.2.1 Correlation matching

The most common area-based matching technique used is known to as cross-correlation. This method has been in use for many years as a means of establishing matching areas of two images (Gonzalez and Woods 1992, Jain 1989, Gonzalez and Wintz 1987, Hannah 1974). If *s(x, y)* is a sample area in the left image and *t(x, y)* is a test area in the right image, the two dimensional correlation equation describing correspondence between the windows is:

$$R(m,n) = \frac{1}{XY} \sum_{y=0}^{Y-1} \sum_{x=0}^{X-1} s(x,y)t(x+m, y+n) \qquad (2.1)$$

where

R(m, n) is the array of correlation coefficient values produced by the correlation, with values between 0 and 1,

X is the horizontal size of the sample area,

Y is the vertical size of the sample area,

M is the horizontal search range in the right image,

N is the vertical search range in the right image,

m ranges from $\frac{-M}{2}$ to $\frac{M}{2}$,

n ranges from $\frac{-N}{2}$ to $\frac{N}{2}$,

x, y is current position of window.

The location of right window when the correlation value *R(m, n)* is a maximum defines the position of the best match with the window on the left image. The method needs an estimation of the approximate matched location before matching is commenced to speed-up the process and ensures that erroneous matches do not occur. Equation (2.1) is simple and easy to program. Since several maxima may occur in the region of a matched point, other auxiliary methods such as probability relaxation and constraints are needed to ensure that the best estimate of the matching point is determined. A comparison of the correlation formulas can be found in (Aschwanden and Guggenbuhl 1993).

2.2.2 Least squares correlation

The high precision least squares approach to image matching, developed by Foerstner (1982) and Ackermann (1984), determines the optimum match between two windows by minimizing the grey value differences of corresponding pixels in windows on two images.

For the two dimensional correlation case, the grey values of the left and right image windows are given by $\bar{g}_1(x,y)$ and $\bar{g}_2(x,y)$. They represent the unknown image functions $g_1(x,y)$ and $g_2(x,y)$ which are perturbed by noise $n_1(x,y)$ and $n_2(x,y)$.

$$\begin{aligned} \bar{g}_1(x,y) &= g_1(x,y) + n_1(x,y) \\ \bar{g}_2(x,y) &= g_2(x,y) + n_2(x,y) \end{aligned} \tag{2.2}$$

Because $g_1(x,y)$ and $g_2(x,y)$ are images of the same object, $g_2(x,y)$ can be described by $g_1(x,y)$ subject to geometric and radiometric transformations as follows:

$$g_2(x,y) = h_0 + h_1 \cdot g_1(x',y') \tag{2.3}$$

where

h_0 represents a zero level shift of the grey level values

h_1 represents a brightness scale factor of the grey level values

Six parameters $(a_0, a_1, a_2, b_0, b_1, b_2)$ may be used to apply affine linear transformation of one window to the geometry of the other, as follows:

$$x' = a_0 + a_1 x + a_2 y$$
$$y' = b_0 + b_1 x + b_2 y \tag{2.4}$$

Expanding Equation (2.3) to linear increments, the following equation for discrete coordinate (x_i, y_j) can be obtained.

$$\begin{aligned} g_2(x_i, y_j) = g_1(x_i, y_j) &+ \dot{g}_x(x_i, y_j)da_0 + x_i\dot{g}_x(x_i, y_j)da_1 + y_j\dot{g}_x(x_i, y_j)da_2 \\ &+ \dot{g}_y(x_i, y_j)db_o + x_i\dot{g}_y(x_i, y_j)db_1 + y_j\dot{g}_y(x_i, y_j)db_2 \\ &+ dh_0 + g_1(x_i, y_j)dh_1 \end{aligned} \tag{2.5}$$

Where

$\dot{g}_x(x_i, y_j) = \partial g_1(x_i, y_j)/\partial x$ and $\dot{g}_y(x_i, y_j) = \partial g_1(x_i, y_j)/\partial y$ are the discrete values of gradient function of $g_1(x_i, y_j)$ in x and y directions respectively.

Considering the difference between the left and right images, Equation (2.6) can be obtained.

$$\begin{aligned} \overline{\Delta g}(x_i, y_j) &= \overline{g}_2(x_i, y_j) - \overline{g}_1(x_i, y_j) \\ &= g_2(x_i, y_j) - g_1(x_i, y_j) + n_2(x_i, y_j) - n_1(x_i, y_j) \end{aligned} \tag{2.6}$$

Substituting $g_2(x_i, y_j)$ with Equation (2.5) and let $v(x_i, y_j) = n_2(x_i, y_j) - n_1(x_i, y_j)$, the final observation equation is in (2.7).

$$\begin{aligned} \overline{\Delta g}(x_i, y_j) + v(x_i, y_j) = &\dot{g}_x(x_i, y_j)da_0 + x_i\dot{g}_x(x_i, y_j)da_1 + y_j\dot{g}_x(x_i, y_j)da_2 \\ &+ \dot{g}_y(x_i, y_j)db_o + x_i\dot{g}_y(x_i, y_j)db_1 + y_j\dot{g}_y(x_i, y_j)db_2 \\ &+ dh_0 + g_1(x_i, y_j)dh_1 \end{aligned} \tag{2.7}$$

A least squares solution minimizes the sum of squares of residual differences $v(x_i, y_j)$ and produces the transformation required to determine the optimum window geometry for matching. The window in the image $g_1(x', y')$ is required to be resampled to the

geometry of the optimum window and the procedure of minimizing the residual differences is iterated until the accuracy requirements are satisfied. The method determines the matching with sub-pixel accuracy.

Least squares matching has been extended by other researchers and developed into different forms, such as the adaptive least squares matching proposed by Gruen (1985). These algorithms are applied in different matching situations. It has been demonstrated that least squares matching makes it possible to fully exploit the accuracy potential of images and systems. It provides a measurement accuracy which is beyond the capabilities of a human operator (Atkinson, 1996). Least squares matching method has been applied in the transform domain by Reeves and Kubik (1998). Object space least squares correlation was proposed by Wrobel (1987), Helava (1988) and Heipke(1992). Due to its complexity with respect to implementation and handling, object space least squares matching has been used only occasionally and under laboratory conditions.

2.3 Feature-based Methods

Feature-based methods match primitives of features within stereo pairs. Firstly, these primitives must be extracted by appropriate image processing methods, then they are used as the basis for matching. The extraction of features is achieved through a variety of techniques, as will be described below.

2.3.1 Interest point matching using labelling relaxation

Scene labelling by relaxation operations was proposed by Rosenfeld (Rosenfeld et al. 1976). In this approach a set of objects in a scene is labelled by a set of possible labels. In the beginning, these labels are ambiguous, so relationships among the objects are used to reduce or eliminate the ambiguity. This algorithm has been extended by many researchers for use in matching algorithms (Marr et al. 1978, Barnard and Thompson 1980, Drumheller and Poggio 1986, Kim and Aggarwal 1987, Lo and Mulder 1992). A detailed comparison of the labelling relaxation can be found in Price (1985).

Based on relaxation labelling, an interest point matching algorithm was proposed by Barnard and Thompson (1980). Matchable points should be small discrete local features such as spots and corners which are likely to be detectable in both images. They should be the centres of highly variable areas. This means the variance should be high in all directions. Since the Moravec interest operator (Moravec 1977) calculates the variance of a window in four directions: horizontal, vertical, and the two diagonals, Barnard and Thompson selected interest points using this operator.

The selected interest points are then matched using the probability relaxation process. Initial match probabilities are assigned to each interest point in the left image for each possible match in the right image. These probabilities are then iteratively refined according to the consistency property of disparity in a local area. If the neighbouring points surrounding a current interest point have similar disparities and high probabilities, the probability of this point having that particular disparity is increased. The iterations continue until each point in the left image has a match in the right image, or the point is assigned no match.

A feature-based correspondence algorithm for image matching was also proposed by Foerstner (1986). The procedures consist of three steps. Firstly, distinct points are selected in the images separately using the Foerstner operator. Then, a preliminary list of candidate pairs of corresponding points is built, which is based on the similarity of the points using the correlation coefficient. The final list of corresponding point pairs can be derived using consistency measure algorithms. This feature-based matching is superior to image correlation in speed and versatility and is also superior to least squares matching in convergence, speed and versatility, but it usually has lower accuracy than area-based matching methods.

The above approach is the basis of the Match-T. Algorithm of Match-T (Ackermann, 1991) is claimed to be a complete and highly automated DSM generation environment. The thorough optimization of the core matching algorithms is reflected in the higher quality of results, and a considerable speed gain is achieved by performing feature point

extraction and epipolar computation. Match-T demonstrates the advantages of feature-based matching techniques.

2.3.2 Edge matching based on polygon approximation and $\psi - s$ representation

Edge matching is the process to determine, among the candidates in the two images, which edges correspond. There are many operators for finding edge elements in images. When the edges have been extracted from the images, edge matching can be implemented based on the attributes of edge elements, such as end point coordinates, local orientations and local intensity profile. An edge feature-based matching algorithm was successfully applied to stereo matching in photogrammetry by Greenfeld (Greenfeld, 1987). Later, this algorithm was improved by several researchers (Greenfeld and Schenk, 1989, Schenk at el. 1991). Two edges matching methods based on different representations are introduced here.

Edges matching based on polygon approximation representation

In Greenfeld's algorithm, the images are firstly convolved with the LOG (Laplacian of Gaussian) operator and the edges defined by zero crossings. Since the chain code is a simple and very useful function of a digital curve, the curves obtained from zero crossing are then converted to the chain code representation.

The polygon approximation process involves breaking the chain code curve into segments. The break points are chosen at points of maximum local curvature, in order to ensure that the structure and shape characteristics of the curve are maintained. The curvature at each point can be computed by the first and second differences of the chain code curve.

When all edges have been extracted and converted to the polygon approximation representation, each edge in the first image is compared with all the edges in the other image. The matching is a selection process in which edges are considered corresponding, according to measures of similarity and consistency.

First, an initial list of matches is established using similarities in the characteristics of the approximating polygon. This polygon matching algorithm is based on vertices rather than on polygon sides. After the initial list of pairs of similar vertices are obtained, the consistency check is used to eliminate the incorrect matches and resolve the ambiguities. The consistency check uses voting schemes to study the geometric pattern and the inter-relationships between the neighboring correct matches.

Edges matching based on ψ-s representation

In the algorithm of edge matching based on polygon approximation, some edge segments which appear to be very similar can only be partially matched. The ψ-s representation method (Greenfeld 1989, Schenk et al. 1991) is proposed to overcome this drawback.

The method commences in the same way as the polygon approximation. Edges are obtained from zero crossings of images convolved with the LOG operator and converted to chain codes. When the chain code representations of the edges are described in the $\psi - s$ domain, s is the index of the pixel in the curve, and ψ is the sum of the differences between all neighboring chain codes up to position s. The formula is as follows:

$$\psi_i = \sum_{j=1}^{i} (f_{j+1} - f_j) \qquad (2.8)$$

$$s_i = i$$

where ψ_i and s_i are the $\psi - s$ representation at point i along the curve and f_j is an integer value corresponding to the chain code value of point j. The $\psi - s$ representation has many advantages in edge matching, such as invariance to the original position of the edge in the image and vertices in the $\psi - s$ domain corresponding to changes of curvature in the original curve.

Edge matching in $\psi - s$ domain is the same as the matching in the polygon approximation method. Since vertices in the $\psi - s$ domain correspond to changes in the

curvature of the original curves and edges of similar shape have similar vertices, the matching process requires finding the $\psi - s$ representation with similar vertices in two steps. In the first step, a set of possible matching candidates is determined by comparing the angles and orientations at each vertex in the left image. In the next matching step, the possible matching points can be refined by the criteria of maximum line consistency, which checks for consistency with neighbouring vertices on the edge. The $\psi - s$ method offers a simple and robust matching procedure.

2.3.3 Relational matching

Relational matching plays an important role in computer vision. A thorough introduction to relational matching techniques is given in Ballard and Brown (1982) and Vosselman (1992). The relational technique is not new, but it can be thought of as an extension of feature-based matching. It is particularly important in image understanding. Its main characteristic is that it collects the detected features into groups of lines or curves depending on the particular application.

Edge pixels can be determined by a variety of edge detection operators which also provide information on edge properties, such as gradient magnitude and direction. Popular edge detection operators include those of Roberts, Sobel, Prewitt, Marr-Hildreth and Canny. The type of data under consideration will determine the appropriate technique to be applied (Russ 1995). The detected edge pixels are then vectorized, by connecting adjoining pixels, resulting in mathematical representations of the groups of edges, rather than physical ones. This forms a more robust technique than feature matching, since relational matching is a high level vision process which involves matching relational descriptions of objects.

A critical step in applying relational matching is the proper representation of the features and their relations. Shapiro (1980) presented an example of useful feature representations. Shapiro and Haralick (1981) formulated structural description as a relational representation of a 2-D or 3-D entity consisting of a set of primitives each having its own attributes and named relations. Successfully extracting the structural

description of the objects tends to be highly dependent on the application. Relational matching then involves comparing the two relations using the primitives of the structural descriptions (Zilberstein, 1992).

2.4 Hybrid Methods

Hybrid methods combine area-based and feature-based matching methods to improve the overall matching.

2.4.1 A stereo image matching system based on intelligent vision

A stereo image matching system based on intelligent vision was introduced by Greenfeld (Greenfeld 1991). Instead of interpreting the matching problem as one of simply finding similar points, he proposed that it as an intelligent vision problem should be capable of performing stereo matching automatically.

In Greenfeld's system, the Central Monitoring System (CMS) is the heart of the matching system and is similar to the engine of an expert system. The goals and functions of the CMS are to interpret the user's request and efficiently control the feature matching, resolve difficult matching problems and supply the products needed by the user. CMS assigns the tasks to different modules and provides the necessary data to each process. In turn, the various modules provide information and feedback to the CMS in order to assist the solution of the problem areas. For example, information from the object recognition and image interpretation can be used to guide the edge matching.

The system also includes image preprocessing, edge extraction, symbolic representation and edge matching. The digital terrain model is generated by means of correlation methods. Area-based correlation is performed following the edge matching, so that edges can be used to break up images into smaller regions and constrain the correlation process. The matched edges serve as a guide for selecting corresponding image patches. If point positioning is required, the interest point and interpolation module is started by CMS.

For poor textured content areas, a surface interpolation can be employed instead of correlation.

The above system cannot be fully realized because it lacks knowledge about the cognitive and intelligent process. Greefeld only implemented those tasks that can be realized, such as interest point matching, DTM generation by correlation and automatic relative orientation, but he stressed the importance of a knowledge based reasoning process for selection of the features is stressed.

Although low level image processing has achieved some success for information extraction from images, intelligent image understanding and interpretation have become more important. Researchers in photogrammetric computer vision and artificial intelligence are working on those high level image processes. The detailed review for some algorithms used in knowledge based representation and modelling for machine vision can be found in (Sowmya and Trinder, 2000).

2.4.2 Area-based correlation with edge information

Area-based correlation with an edge information algorithm was proposed by Cochran and Medioni(1992). First, the original images are preprocessed to find the epipolar images. Next, a pyramid of three image pairs is formed by convolving the epipolar images with a Gaussian operator and subsampling by a factor of two in both directions. The area and feature based processes are done independently to produce a set of edge features and a dense disparity estimate. These are then combined to refine the disparity map.

- Area-based matching

  Area-based matching uses a measure of the local texture in the image pair to produce a dense disparity map by the cross correlation formula. Lack of texture causes the matching of some selected features to be ineffective. In this algorithm, the local variation of the image texture is computed. When there is little or no matchable texture, a small value is returned. This is used to mark points for which no match should be generated during the initial disparity estimation. After the above

processing, the initial disparities can be obtained. Then a set of constraints such as the left and right consistency checking, order constraint and isolated pixels are applied to identify and remove unlikely matches. Order constraint ensures two points in the left image have the same order as their corresponded points in the right image.

- Disparity map refinement

The ambiguous points in the disparity map can be refined using edge information. Instead of using matched edges by independent feature-based matching, the algorithm uses the edges from one image and associates them with the disparity obtained from the area-based matching. The advantage of this combination is that even those edges with an orientation close to epipolar line still play an active role, but they are discarded in all other feature-based matching approaches.

2.4.3 A stereo matching algorithm for urban digital elevation models

This method was proposed by (Baillard and Dissard 2000) and takes advantage of both feature-based and area-based matching methods. It takes occlusions into account through a dynamic programming optimization process. The novelty of the approach lies in the use of several successive steps, each of them is useful for certain kinds of pixels. The system includes the following steps:

- Edge matching

Edges are extracted from each image with the Canny filter followed by thresholding. A local elementary similarity cost function is computed for each edge pair. This cost function defines potential edge pairs and potential disparity values. An analysis of the distribution of potential disparity values along edge chains leads to the computation of the final matching cost function. The matching is then achieved by dynamic programming (Ohta and Kanade, 1985) for each epipolar line pair. Finally, a post processing is performed to remove local errors and guarantee the consistency of the 3D chains.

- Area-based matching

In order to produce dense elevation data, area-based matching is performed between the matched edge points. Dynamic programming is used again, which aims to match the pairs of intervals defined by the matched edges in the first step. Using matched edges as anchor points reduces the search space for corresponding pixels, therefore decreasing the number of false matches and the computation time. During the matching, two constraints have been used:

- A strong radiometric similarity constraint is applied in order to produce only reliable pairs.
- A looser radiometric but a stronger geometric constraint is used to complete the matching on unmatched areas.

This hierarchical strategy relies on the assumption that local extremes of depth along epipolar lines are recovered as reliable pairs during the radiometrically constrained processing. Some information about image radiometry and shadows is also computed and exploited during the matching. The algorithm provides dense elevation values and preserves height discontinuities. Its main drawback is the relative planimetric inaccuracy in the neighborhood of height discontinuities.

2.5 Summary of the Image Matching Methods

Comparing the matching methods described above, both area-based and feature-based matching techniques have inherent advantages and disadvantages.

Feature-based matching provides generally good and highly reliable results because of the distinctive properties of features. This method is less expensive, but usually produces fewer disparity values because the feature extraction process disregards some of the information in the images. In some regions, few points may be matched, which leads to large areas being subject to inaccurate interpolation on calculation (Cochran and Medioni 1992). The problem is overcome in Match-T, since advanced techniques are used, such as interpolation from an extremely dense cloud of derived 3D points, elimination of outliers using robust finite element interpolation and automatic adaptive DTM grid

spacing depending on the surface curvature. High reliability and accuracy have been achieved.

Area-based techniques are simple and have the potential to produce dense disparity maps, but tend to be computationally more expensive. Area-based global image matching can be designed to match at regular grid points in image space or object space, and the denser the grid points are, the more object details can be described. Area-based methods are intolerant to distortions caused by terrain slopes and are more suited to undulating or smooth terrain. Area-based methods work well only if there is sufficient texture in the image. Hence there must be adequate feature information available in the images.

In general, matching algorithms work well if there is adequate detail in the images and fail if the content is repetitive or inadequate. After reviewing and comparing the matching algorithms, it is clear that most matching algorithms involve the initial determination of corresponding points, then refinement of the matches to discard the ambiguous points. The refining process includes probability estimation and global continuity checking and ordering constraints. These constraints are based on the assumption that disparity changes smoothly.

The area-based matching algorithms are normally used in photogrammetry for smooth and undulating terrain reconstruction. The feature-based and hybrid matching algorithms are usually used in computer vision, close range photogrammetry and 3D building reconstruction. In urban areas, since many discontinuities exist in the disparity map, a refinement process such as those suggested above, cannot be used. Neither area-based matching nor feature-based matching can supply sufficient accurate information for reconstruction of an accurate DTM in such areas. DTM modelling therefore needs to integrate multiple information cues and image understanding techniques, in addition to geometry constraints. A DSM which preserves the discontinuities over urban areas is an important information cue and is needed to assist automatic building extraction in urban areas. In this thesis, a hybrid matching method for DSM generation is proposed.

2.6 Building Extraction

Accurate 3D building representation is required for many purposes, by professionals such as urban planners, architects and telecommunication and environment engineers. Manual 3D processing of aerial images is time consuming and requires experienced people and expensive equipment. A method of interpreting and accurately processing digital aerial images for building reconstruction and integrating the results into spatial information system is urgently needed.

Until the early 1990s, most researchers focused on detecting the buildings and extracting the outline of the roof using a single grey-valued image together with simplified object models. Even though the assumption of simple roof models is used, the analysis of single image for building extraction is a very difficult task because feature correspondences cannot be used to infer 3D and ambiguities are hard to resolve (Huertas and Nevatua 1988, Shufelt and McKeown 1993, Lin and Nevatia 1998). Over the past decade, researchers have presented new and encouraging results from automated 3D building reconstruction methods. Stereo matching has been proved to be an important information cue for building extraction.

Based on the early vision studies and stereo image matching algorithms introduced in sections 2.1-2.4, a large number of automated structure extraction methods have been proposed (Huertas and Nevatia 1988, Liow and Pavlidis 1990, Mckeown 1991, Shufel and McKeown 1993, Haala and Hahn 1995, Baltsavias et al. 1995, Henricsson et al. 1996, Lammi 1996, Henricsson and Baltsvias 1997, Gabet, et al. 1997, Jaynes et al. 1997, Haala et al. 1998 , Baillard et al. 1998, Kim and Muller 1998, Vosselman 1999, Sahar and Krupnik 1999, Chen and Hsu 2000, Masaharu and Hasegawa 2000, Mcintosh et al. 2000, Lu et al. 2002, 2003,). Shadow analysis based algorithms have been used (Liow and Pavlidis 1990, Nevatia et al. 1999). Information fusion based systems have been reported (Mckeown 1991, Haala and Hahn 1995). Methods supported by DTM and orthoimages have also been reported (Baltsavias at al. 1995, Horiguchi et al. 2000, Straub and Heipke 2001, Brunn 2001). Considering the different shapes, environments and image intensity

for different buildings, together with the occurrence of occlusions and shadow effects, the automation of building extraction is a complicated and difficult procedure (Sahar and Krupnik 1999). In addition to developing better schemes, the inclusion of more information is an essential direction for the research. Henricsson (1998), Chen and Hsu (2000) and Niederost (2001) used colour images to improve the system performance for roof determination and edge extraction. Spreeuwears (1997) and Gabet (1997) used multi-view images to reduce the effect of occlusions. Multi-image 3D feature and DSM extraction for building change detection were proposed by Paparoditis (1998, 2001). Laser scanner data is used by (Mcintosh et al. 2000, Masaharu and Hasegawa 2000 , Haala et al. 1998). Multi-resolution analysis of wavelets for house extraction has been proposed by Shi and Shibasaki (1995). A detailed review and comparison of building reconstruction techniques can be found in (Noronha and Nevatia, 2001, Henricsson 1998, Collins et al. 1998).

The purpose of this research is to extract the building boundaries for more accurate DTM reconstruction rather than extract the buildings themselves. The extracted buildings are used to guide the DTM reconstruction in the urban areas. In the following sections, the methods of automatic 3D building reconstruction are classified and reviewed in four classes.

2.7 Semi-Automatic Building Reconstruction

Semi-automatic building reconstruction systems normally make use of a human operator to choose corner points or lines in a flat building roof. Quiguer (1996) introduced a rectangular building reconstruction method which involves manually selecting a building corner for low level feature extraction. The extraction of high quality low level features greatly assists high level processing. The process follows four steps:

- Building corners are selected manually

  A building corner is chosen as a seed point. In order to compensate for wrong manual selection, a feature point detection technique is used to ensure the seed point is a sharp corner in the image.

- Detection of the first two sides

When the seed point is selected, the building sides can be detected on a single image. The selection criteria for line detection are based on the largest gradient along a line, and on the continuity of the magnitude and the direction of the gradient along that line. The second line is extracted, because it is perpendicular to the first.

- Parallelogram closing

The criterion of parallelism is used to close the parallelogram which constitutes a building. For a set of lines parallel to the first two perpendicular sides, cost function values can be calculated. The lines which have minimum cost function values can be chosen to close the parallelogram. The cost function integrates the homogeneity and discontinuity information.

- 3D reconstruction

When a building is recognized in one image by extracting its sides, the homologous feature in the other image can be found using normalized correlation.

The algorithm has been tested on rectangular building areas with satisfactory results. Because the seed point is selected manually, the errors from low level line detection are avoided, so the algorithms for parallelogram closing and 3D reconstruction of buildings are usually successful. The algorithm cannot extract lines which have poor continuity of gradient direction and poor contrast with the background.

Other semi-automatic building reconstruction methods have been proposed by Gruen (1997) and Fua (1996). The semi-automatic system proposed by Gruen aims at the automatic structuring of manually measured 3D point clouds to generate CAD models of complete buildings. It is a reliable and flexible procedure which has a large potential to be useful for professional practice. Fua optimized the shapes of buildings in three views simultaneously based on rough building outlines entered into the system. This algorithm is suitable for rectangular structures.

2.8 Automatic Building Reconstruction Based on 2D Features

Automated construction and management of 3D man-made structures is the key part of the research and development for image understanding systems for aerial images. Man-made structure modelling enables efficient exploitation of the tremendous volume of information collected by the sensors. The expected benefits are to decrease the heavy and tedious workload for human operators. By utilizing digital image understanding and photogrammetry techniques, measurement accuracy can also be increased. Automatic building reconstructions are designed to extract 2D image features and group 2D features into larger united 2D building structures based on rooftop boundaries. Then the extracted polygons can be matched in two images using different 3D constraints. Because the method supplies the initial roof hypotheses in 2D rather than 3D, the very complex and difficult 2D inference must be solved in a single image. While many methods (Noronha and Nevatia 2001, Sohn and Dowman 2001, Collins et al. 1995, 1992) have been proposed based on 2D features, only two methods will be described here.

2.8.1 Building modelling system using rapid feature extraction in urban areas

3D feature extraction creates many problems including those of segmentation, 3D inference and shape description, Nevatia (1999) proposed a building modelling system based on feature extraction. In urban aerial scenes, segmentation is difficult due to the presence of a large number of objects that may be confused with buildings such as sidewalks, landscaping, trees and shadows near buildings. The objects to be modelled may be partially occluded. Based on multiple sources of data and perceptual grouping, Nevatia (1999) used a combination of reconstruction and reasoning in 3D. Context and domain knowledge guide the applications of these combinations. Context is derived from knowledge of camera parameters, geometry of objects to be detected and illumination conditions, such as the sun position. The system uses hypotheses and verification. It can work with just a pair of panchromatic images, but can also utilize more images and information cues from DTM and hyperspectral sensors.

The method can be used for building modelling where the buildings are rectilinear, relatively simple and are not close to each other. Complex buildings are decomposed into rectangular parts. The system includes:

- Processing of multiple panchromatic images: low level features such as lines and junctions between lines, and sets of parallel lines, are matched from the available views. These features are the basis to form roof hypotheses.
- Hypotheses of potential roofs are derived from a pair of matched parallel lines or three sides of a parallelogram. Closed hypotheses are formed from these features using the best available image lines. If there are no available lines, the closures are synthesized from the ends of the lines.
- Hypotheses verification is used to verify whether the selected hypotheses have additional evidence in support of the object being a building. This evidence is collected from the roof, the walls and shadows. Each evidence parameter is represented by smaller components of the evidence. For example, shadow evidence is composed of smaller evidence elements such as strong junctions, weak junctions and darkness etc. Bayesian reasoning is used to combine these elements of evidence to decide whether a building is present or not, and how much confidence should be placed in it.
- 3D roof hypotheses are inferred by searching 2D flat roof parallelograms in the other images. This method supplies a single height for a flat roof and overcomes the non unique results in line matching stage. Different heights can be estimated by this method for separated parts of a building.
- A methodology for evaluation of 3D geospatial building modelling systems is used. In order to characterize the performance of the system, detection and false alarm rates are computed.

The important part of the system is integration of information from different image sources and sensors. This results in significant improvements in the quality of the building reconstruction.

2.8.2 Automated building reconstruction from multiple aerial images

An automated building reconstruction system developed by Collin et al (1998) from multiple aerial images includes two parts. It uses 2D image features and grouping operators to detect rooftop boundaries, and then matches these polygons under 3D constraints. After the height is computed, a 3D volumetric model is built. The system works successfully for flat roof building extraction. However, a considerable number of false structures are generated due to the complexity of the scene and presence of buildings outside of the class for which the system is designed. In order to solve this problem, the system has been further developed by Hanson et al (2001) which incorporates artificial intelligence techniques for dynamically controlling the image understanding processes. The improved system can utilize a set of algorithms which fuse 2D and 3D information and make use of available information during the reconstruction process, such as SAR and multispectral image data.

The original system included the following components

- Line segment extraction

  A straight line feature extraction algorithm is used to produce a set of symbolic line segments which represent potential building features. Firstly, zero crossings of the Laplacian of the intensity image provide an initial set of local edges. Then, hierarchical grouping proceeds iteratively, using measures of co-linearity and connectedness. At each iteration, edge pairs are linked and replaced by a single longer edge if the end points are close, the perpendicular offset is small, and the orientation and the difference of average intensities across the line are similar.

- 2D polygon extraction

  The aim of this step is to approximately delineate building boundaries that will later be verified in other images by epipolar feature matching and triangulated to create 3D geometric building models. The algorithm groups the line segments into image polygons which correspond to the boundaries of flat, rectilinear rooftops. The processes comprise three levels and each level generates features that are used in the next level.

1. The low level process uses straight line segments and corners derived from previous step. In order to determine a set of relevant corners, if a pair of line segments are spatially proximate, their end points are grouped together into candidate image corner features.
2. The mid level process collates features and groups corners and lines to form a chain. The collated feature chains are created by paths in a feature relation graph.
3. The high level process extracts the polygon hypothesis based on the collated feature chains. Then, polygon hypotheses are derived in order to obtain a final set of non conflicting, high confidence rooftop polygons.

- Epipolar line segment matching

After a potential rooftop is detected in one image, corresponding geometric evidence is sought in other images based on epipolar feature matching. The roof edge matching often obtains multiple potentially matching line segments of the appropriate length and orientation. The way to solve this problem is to simultaneously process multiple images. Because the disparity of each potential match of a segment line in the rooftop polygon supplies a different roof height, heights from each line matching are combined in a one dimensional histogram. Each height is weighted based on the potential line's orientation and length. Finally, the histogram bucket supplies an estimate of the roof height.

- 3D building models

A nonlinear estimation algorithm has been developed for triangulating 3D rooftop polygons from the line segment correspondences determined by epipolar feature matching. The algorithm determines the precise size, shape, and position of a building model in the 3D site coordinate system. After processing, each refined 3D roof polygon is extended vertically down to the terrain to form a volumetric model.

- Projective intensity mapping for buildings

For each 3D building model recovered by the above steps, a planar projective transformation algorithm is developed for acquiring image intensity maps for walls and roof surfaces. The algorithm provides a mathematical description of how the surface structure from planar building facets maps into an image. By inverting this

transformation using known building position and camera geometry, intensity information from each image can be back projected to paint the walls and roof of the 3D building model. Because multiple images are used, intensity information from all faces of the 3D building can be recovered, although not all of them can be seen in a single image.

In order to delete the number of false buildings generated by the system, a knowledge-based building reconstruction has been proposed based on the original system. More details can be found in Hanson et al (2001). The original system works successfully for flat roof building reconstruction. The knowledge-based system focuses on the use of multiple alternative reconstruction strategies from which the most appropriate strategies are selected by the system. The system utilizes a wider set of algorithms that fuse 2D and 3D information. It is a more robust reconstruction of 3D building models than the original one.

2.9 Automatic Building Reconstruction Based on 3D Features

The approach of this automatic building reconstruction method generates the roof hypotheses based on 3D information rather than firstly grouping the 2D features. 3D information can consist of high quality DSM and 3D features extracted at the early stage. The difference with respect to the automatic building reconstruction based on 2D features is that the 2D grouping problem does not have to be solved before inferring 3D information. The choice of the automatic building reconstruction based on 2D features or 3D features, depends on which one can be achieved more easily, and on the available data, algorithms, the complexity of the scene and the building models.

There are many methods based on 3D features (Paparoditis et al. 2001,1998, Straub and Heipke 2001, Horiguchi et al. 2000, Fischer et al. 1998, Haala et al. 1998 ,Haala and Hahn 1995,). Only three methods are described here.

2.9.1 3D building reconstruction from aerial images

Henricsson proposed an algorithm of man-made object reconstruction from aerial images (Henricsson, 1996) as illustrated in Figure 2.1.

A digital surface model (DSM) is a rich information source for building detection, but must have high accuracy and sufficient density. It can be obtained from area-based correlation or hybrid matching methods. Combined analysis of colour images can detect 3D building blobs which are the approximate positions of buildings and supply guidance for 2D feature extraction and feature matching. 2D edges are extracted from a source image. Photometric and chromatic attributes and their similarity relationships of the edges are computed. The extracted edges are matched in the multiple stereo image pairs by maximizing a function of the gradient along the epipolar line. Geometric and photometric constraints are also used to reduce the number of mismatches. After computing the 3D location of these edges, planes can be obtained by grouping 3D edges. In addition, 2D enclosures are extracted and combined with the 3D planes to create 3D roof primitives. All extracted hypotheses of 3D parts are ranked according to their geometric quality. Finally, the best set of 3D parts that are mutually consistent is retained, and can be used to define the reconstructed buildings.

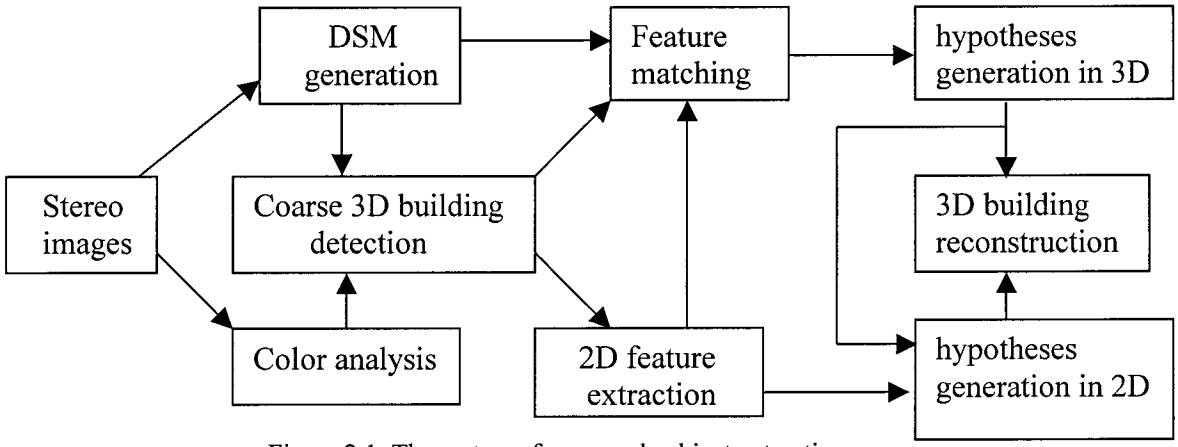


Figure 2.1 The system of man-made object extraction

The majority of research on building reconstruction relies on the use of edge or point based features and their 2D or 3D geometry. Many algorithms rely on the assumption that

man-made objects possess a large amount of geometric regularity, mostly in the form of flat rectilinear roofs. However, these very constrained geometric models may be appropriate for isolated buildings with single flat roof, but not for the majority of suburban and urban buildings. In order to automatically reconstruct a general class of roof types in 3D with a high metric accuracy, Henricsson (1998) further developed his 3D building reconstruction system. The difference between his system and other algorithms is not only the extensive use of colour attributes and similarity relations, but also the far-reaching goal to reconstruct a general class of buildings. The advantages of his system are as following:

- Based on the colour cues, the important role of colour region attributes and similarity grouping in automated 3D building reconstruction was demonstrated.
- A novel framework for grouping contours based on similarity grouping is proposed, using similarity in orientation, position, and photometric and chromatic regional attributes. This general grouping technique efficiently extracts hypotheses of roof parts and is also general enough to extract hypotheses of other man made objects.
- The system shows that although the geometric regularity is important, it cannot serve as the only basis for extracting complex structures for which no generic models exist. The system demonstrates that colour is a very important cue in reconstructing a general class of objects. A general class of object parts can be efficiently extracted by grouping edges and lines by means of similarity, not geometric regularity.
- All available data and information should be retained throughout the entire process and be shared among the individual processing modules as it becomes available. Combining colour region attributes and grouping, the system shows that not only the symbolic features are important but also their attached rich image attributes.

2.9.2 Statistical interpretation of DTM and image data for building extraction

In building extraction, statistical interpretation of data in combination with DTM and image data becomes increasingly important, because many works are based purely on geometry and lack open interfaces to non geometric information. A statistical approach which models and operates on geometric and semantic relations, has been proposed by Brunn (1998, 2000) to give a framework for integration of models of multiple types and of multiple sensors data.

- The statistical detection of buildings in multi-sensor data

A raster DSM has been generated and the DTM is extracted by morphological filtering from the DSM. The normalized DSM is derived by calculating the difference between the DSM and the DTM. A Bayesian Net (BN) is constructed by meta rules, which makes the approach independent of size and resolution of the datasets. Each random variable of the BN is used for classification. It comprises probabilities for the detection of a building object or a non building object. The probabilities for different objects have different evaluated information which are obtained from observation and prior knowledge. The observation information can be the height values of normalized DSM, step edge magnitudes and the variances of DTM surface, as well as colour. For the prior knowledge, the appearance of the objects as a building and as a non-building in all datasets is learned. A user is asked to mark some pixels in one dataset where the objects such as buildings and non-buildings are found. Empirical distributions of both objects on each sensor data type can be obtained. All information features are transformed into probability distributions. The detection procedure based on Bayesian nets calculates the probability distributions of building hypotheses and of other non-buildings, and finally gives a bounding polygon which encloses the complete building.

- Interpretation and reconstruction of buildings using their topology

In order to reconstruct the buildings, an approximate description of the building is generated using a new method called simplicial complexes. It is a graph representation and describes the topology of building surface. The basic elements of

this representation are called simplices. Because the topology is independent of the sensor type used, the topology of the surface can be used as the base representation of the building.

Simplicial complexes consists of 0, 1 and 2 simplices, to represent the topology of the surface of buildings. Each simplex is associated with some appearance attributes, such as geometric or radiometric. The meaning of classified simplices for geometric model is as following:

- 0-simplices are corner points, edge points on the borders of two building planes, and face point inside one plane of the building.
- 1-simplices are breaklines, face edges inside one building plane.
- 2-simplices are vertical, oblique and horizontal faces.

The interpretation is done by classification of the simplices of the building representation using statistical models. During this processing, a set of interactively classified simplices could be evaluated, and an optimum graph representation which includes the attributes is chosen. After this correction of the appearance of the building, different buildings are reconstructed. This method combines geometric and semantic relations, and models multiple building types, but it depends on the prior knowledge learning.

2.9.3 Automated house extraction based on wavelet transform

The physicist Grossmann and Morlet introduced the concepts of "wavelet" and "wavelet transform" in 1984. After a large amount of work of many researchers (Mallat 1989, 1992, Daubechies 1988, 1992, Chui 1992), it has been proved to be practical in both signal analysis and image processing.

In two dimensional wavelet transform cases, for a scale increase, the resolution increases in the spatial domain and decreases in the frequency domain. This is the advantage of the wavelet transform which can be used to "zoom in " and "zoom out" on very sharp changes, such as discontinuities in images. Taking advantage of this, Shi (1995) proposed automated house extraction system based on wavelet transforms. The system includes:

- Multiscale feature extraction based on wavelet analysis

A directional wavelet transform model for images is used, from which local extremes and zero crossings can be found and used to detect the sharp variations of images. The results of wavelet transform show that not only the edges in an image can be detected, but also their orientations. Furthermore, the local maximum can also characterize the local shape of the sharp variations.

By analyzing the shape of wavelet transform of 2^j resolution under the first and second derivative, the modulus of transform near the corners of a man-made structure reaches a peak. It is therefore easy to extract the corners of man-made structures such as houses by searching the local peaks. In addition, the information about the type of corner, either a convex corner or a concave corner, can also be obtained. Such information is significant and useful for the image segmentation and stereo matching, because the corners can be used as the seeds in region growing and the number and types of corners are very important constraints for stereo image matching and pattern recognition.

- Region segmentation based on wavelet transform

A multi-resolution analysis based on wavelet theory is used in image segmentation. Local maxima, zero crossings, contrast sign of zero crossings and corners are extracted, then a seed-based region growing algorithm has been implemented using a coarse-to-fine procedure. The attribute parameters of regions are calculated for region matching (Shi and Shibasaki,1995).

The object which is unable to be extracted at high resolution may be extracted at the low resolutions. The efficiency of a region segmentation algorithm is higher than that of either a split-and-merge algorithm or seed-based region growing methods, because only features such as edges and corners will be taken into account in the procedure, whereas every pixel of an image should be processed in region growing algorithms. The final regions of the image segmentation will be uniform and homogeneous with respect to some characteristic, such as gray level or texture.

- Hierarchical image matching based on multi-features

Two kinds of hierarchical stereo matching schemes corresponding to the above selected features are proposed in this matching system. One is the coarse-to-fine stereo matching scheme which performs the feature-based stereo matching in different scales. The other is the hierarchical stereo matching scheme in which matching based on regions, lines and corners is carried out at each scale. The system is successful on flat roof scenes (Shi and Shibasaki,1995).

2.10 Automatic Building Reconstruction Based on Laser Data

Airborne laser scanning data supplies a new technique for the highly automated generation of digital terrain model and surface models. Laser scanning and photogrammetry are two complementary techniques and their integration can lead to more accurate and complete new products. A new application of laser scanning is the automatic reconstruction buildings in the urban areas for city modeling. (Ackermann, 1999, Baltsavias, 1999). There are many methods in literature on the extraction of buildings from laser scan data (Masaharu and Hasegawa 2000, Vosselman, 1999, Axelsson, 1999, Haala et al. 1998, Hug, 1997). It is beyond the scope of the thesis to review all these methods. Hence, only methods which also use images are reviewed here.

2.10.1 Improvement of automatic DSM generation for automatic building reconstruction using laser scanner data

Laser scanning data is an efficient and accurate method to obtain visible surface information, but the data only provides coordinates. Additional terrain information, such as breaklines, is not available. The location of surface discontinuities may be detected by a planar segmentation method, but the accurate positions cannot be obtained. A method of merging the accurate terrain information from images with the laser scanner data was proposed by Mcintosh, et al (2000). The method can supply more accurate DSM data for automatic building reconstruction than either of the two separate data sets. Surface

matching and data fusion are the key ideas of the method. The algorithm can be implemented using the following components.

- Surface registration

Although photogrammetric data and laser scanner data should theoretically be on the same coordinate system, systematic errors in the laser data may cause a misalignment between the two surfaces. The errors must be eliminated before data fusion can be performed accurately. The algorithms for surface registration have been proposed by Schenk et al, (2000) and Habib and Schenk (1999).

- Image edge extraction

Firstly, the optimal zero-crossings operator (Sarkar and Boyer, 1991) is used for the detection of the edge pixels. Then the connected edge chains are found based on analysis of the edge pixels and straight line segments are determined.

- Line segment matching

Line segments are compared using the correlation coefficient for corresponding points between line segments based on epipolar geometry. Two segments which have the highest average correlation coefficient can be chosen from many potential matching segments and labelled as correctly matched line segments.

- Line segment reconstruction

3D coordinates of the line segments are calculated using the interior and exterior orientation parameters of the aerial image. The accuracy can be checked by comparison with the range of elevations of the laser data.

- Filtering laser data

The laser data is classified as either ground, vegetation or building points based on the spatial frequency of the data. The laser points nearest the breaklines derived from image analysis are checked and determined whether they are associated with a building. The laser data can be used to detect the roofline of buildings and also the ground surface near that roofline.

- Surface discontinuities

The most important edges to be detected are the discontinuities between the roof and the ground. The detection of breaklines is difficult and the results depend on the image and algorithms. When the breaklines representing the roof are detected, they

will be projected vertically onto the ground surface to produce a new breakline in the DSM. The elevation of the ground surface at the location of the added breakline must be determined by searching the laser data in the areas surrounding the breakline.

An integration algorithm to improve the automatic DSM generation for building reconstruction using airborne laser scanner data has been tested (Mcintosh, et al 2000). The results show more accurate DSM can be obtained for automatic building reconstruction.

2.10.2 Automatic reconstruction of buildings in laser scanning data for disaster management

After a strong earthquake, many areas are not accessible because of the blocked streets and the risk of collapsing buildings. Rapid automatic reconstruction of buildings and change detection in buildings using laser scanning data is an important requirement for disaster management. Steinle and Vogtle (2001) proposed a system for this purpose. Because the reconstruction of post-disaster building models should consider missing components of buildings and roofs, changes of object size, shape and structures, which cannot be described by planes, the system is independent of rectangle or parallel edges. The system includes:

- Finding the objects with height above the ground

  In order to find these objects, the normalized DSM is calculated based on the difference between the original DSM and a digital terrain model (DTM) of the same area. Then all 3D objects with significant size and height are segmented by a region growing method. The purpose of this segmentation is to separate 3D objects areas, from all smaller and lower objects.

- Spectral signature

  The spectral image can be used to discriminate buildings from trees and bushes. In the past, an additional sensor had to be used to obtain the spectral image during the same flight. However, new generation of laser sensors involve an integrated system, where not only the DSM may be derived, but also the spectral intensities of the laser pulses can be recorded in the near infrared domain.

- Analysis of the laser data

Laser data has been analyzed using the height texture parameters. One of these parameters, obtained by a Laplace operator, can be used to distinguish between vegetation and buildings. The local curvature, which is the difference of two subsequent gradients, is used to find the areas of significant height texture.

- Normalized vectors

In the planar areas, such as roofs, all normalized vectors have the same direction. However, within tree surfaces, the normalized vectors have different directions. This feature can be further used to support the discrimination of vegetation and buildings.

- Shape and size parameters

In order to calculate the shape and size parameters of the segmented 3D objects with height over the ground, a vectorization has to be carried out to provide 2D contour lines of each segment. All information from the above steps is input to help calculate the features of the regions and separate buildings from all the other objects. The final building hypotheses can then be created.

- Reconstruction and modelling of buildings

Building reconstruction method is strongly influenced by the application. For estimating the changes and damage to buildings caused by a disaster, a fast building reconstruction to support rescue activities is need. It is not necessary to reconstruct the highest level of detail of the buildings. The resulting building models are used not only to overview the damaged situation, but also to find a suitable rescue method. A building reconstruction process is proposed which is strongly based on segmentation of roof planes in a laser height data set. Using the roof segmentation, the neighbouring relations are detected to build up the topology of the roof parts. The intersection of the adjacent planes and building edges can be derived. Intersecting these building edge lines leads to the estimation of building corners and the correct length of the edges.

This automatic 3D building reconstruction is based on dense height data from laser scanner techniques and spectral images. The system is suitable for fast data derivation of

large areas after an earthquake disaster. Reconstructed building models are approximated by planar surface, without the need to form rectangle or parallel edges of buildings.

2.11 Summary of the Building Extraction Techniques

Sections 2.6-2.10 have presented a wide range of building reconstruction techniques which have been described in literature. After reviewing and comparing the building reconstruction algorithms, it can be concluded that most use a bottom-up process by extracting building features and components, then reconstructing the components into building models. One of the objectives of this thesis is to investigate automatic building extraction techniques for 3D terrain reconstruction over urban areas. While complete building reconstruction systems have been described in this review, the locations of building outlines are of primary interest in this thesis rather than the actual complete construction.

Semi-automatic of building reconstruction techniques discussed in this chapter will not usually be suitable because they require manual input. Laser scanner data is a new technique and complements the photogrammetric methods, but it is expensive.

Automatic building reconstruction based on 2D and 3D features using multiple data sources will be the approach taken in this thesis. It is a complex research task which cannot be solved by a single step, and must be solved by a combination of different techniques. Automatic building reconstruction based only on 2D features, involves extracting 2D image features and grouping them into 2D building structures based on rooftop boundaries. Because this method supplies the initial roof hypotheses in 2D rather than 3D, 2D inference of building boundaries must be solved in a single image. On the other hand, the approach of building reconstruction based on 3D features generates the roof hypotheses from the 3D information, but the matching of the features creates problems.

The work in this thesis is concentrated on building boundary extraction for DTM reconstruction. It combines the above two techniques. 3D information such as DSM, with addition information extracted from the 2D images is used to locate building areas. Then 2D shape modelling is used to delineate the precision building boundaries. Because the shape modelling is undertaken over a small area, the complicated and difficult 2D inference based on multiple information datasets can be overcome. Finally, a data fusion method has been used to evaluate the extracted building regions. The approach uses a combined image analysis and understanding technique, which enables greater reliability in the image analysis, and is essential for providing a more accuracy DTM. The details of this method will be given in Chapter 3.

2.12 Architecture of the Proposed Automatic Building Extraction System

In this section, the automatic building extraction system developed in this study will be introduced. The main purpose for developing this system is for DTM reconstruction over urban areas. In order to achieve this aim, the methods chosen for this system should be based on the following principles:

- The methods should have the ability to differentiate buildings from other objects above the ground surface.

  As described in the sections 2.3~2.5, stereo image matching in urban areas results in errors in the elevations derived by conventional image matching techniques, because the matching is made on the top of buildings and trees. In order to provide more accurate DTM, the characteristics of the terrain cover, such as buildings and trees, must to be determined to predict errors in the computed elevations.

- The methods should include integrated image understanding techniques to interpret the content of the aerial image for building extraction.

  In order to differentiate buildings from other objects, the spectral content of the image is needed to classify and interpret the image. However, in urban areas, the

Intensity of an object in an image may vary greatly due to the image noise, its texture, weather conditions, sun angle and slope, etc, leading to incorrect classifications. Therefore, the methods adapted for this system will use classification in addition to other methods of object identification.

In the previous sections, it has been recognized that automatic building extraction cannot be solved satisfactorily by a single step and must be solved by a combination of different image interpretation techniques. The approach taken will depend upon the information that can be extracted from the data. When there are insufficient observed or measured values derived from the images, an ambiguous interpretation may result. Hence the integration of information from multiple data sources has been shown to be an essential procedure for interpretation of image.

- The approach taken in this research involves the extraction of building boundaries to guide the DTM reconstruction.

  In section 2.4.3, a stereo matching algorithm for urban digital elevation models (Baillard and Dissard 2000) is introduced, in which edge features have been used to guide DTM reconstruction. The method shows that such features are important for urban DTM reconstruction. Boundaries of building are superior evidence for feature extraction than isolated pixels and edges, because they are connected edges and present meaningful regions. When building boundaries have been extracted in the left and right images, they indicate the existence of buildings. As a result, in producing dense elevation data, area-based matching can be performed between the matched building regions. Then using matched buildings as seed regions reduces the search space for corresponding pixels on each images, thereby decreasing the number of false matches and the computation time. For each building region, the DTM can be interpolated from the neighbouring terrain points or a building region can be assigned the same elevation value.

- The methods can supply reliable extracted buildings based on evidential evaluation from statistical theory.

The complex urban scene and drawbacks of the algorithms can result in wrongly extracting building regions. The methods are required to evaluate the reliability of the evidence of buildings and delete those buildings that have been wrongly identified.

Figure 2.2 illustrates the architecture of the proposed automatic building extraction system to aid the reconstruction of elevations from overlapping aerial or satellite images over a variety of terrain types and ground cover. It consists of four main modules. Based on the above principles, the following methods for combining multiple data sources have been chosen for the automatic building extraction procedures.

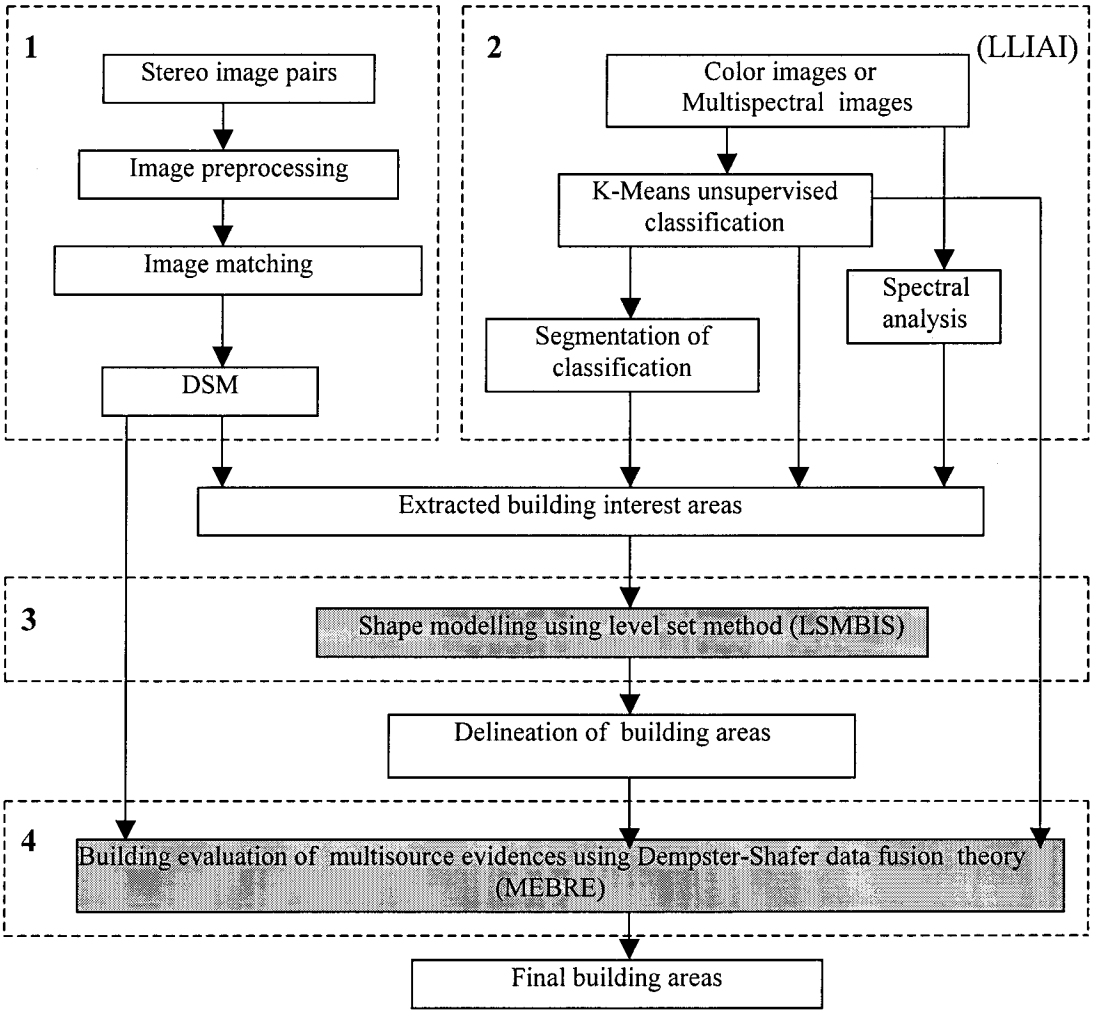


Figure 2.2 The architecture of the proposed automatic building extraction system

The aim of module 1 is to extract a DSM and shown in Figure 2.2, the functions of data acquisition and pre-processing, are to acquire the images in digital form and produce epipolar images from the original left and right images for the subsequent processes by interior orientation and exterior orientation. The detail description of this method of image matching and DSM extraction will be given in Chapter 3.

As shown in section 2.9, Henricsson (1996) used DSM to detect 3D building blocks which can supply evidence for 2D feature extraction and feature matching, while Brunn (1998,2000) statistically detected buildings based on DSM. Therefore, it has been shown that DSM extracted by stereo image matching provides important information cues for automatic building extraction, since 3D information reduces ambiguity in the interpretation of the images. In this study, DSM is required as evidence to differentiate the buildings from other objects. Elevation areas above the terrain surface are likely to be buildings and trees, while objects with elevations similar to the terrain surface are likely to be grassland, cultivated areas or road. In order to supply accurate DSM for subsequent building extraction, a robust hybrid-based image matching and least square matching with robust estimation algorithms have been tested as an approach. The contribution of the methods is to supply greater accuracy in the computation of disparities in the images for areas where discontinuities caused by buildings and trees exist.

The module of low-level image analysis and interpretation (LLIAI) shown as module 2 in Figure 2.2, aims to extract significant features from images and classify the images into the meaningful regions, to reveal the building interest areas using classification and spectral analysis techniques. It includes the following processes and is discussed in Chapter 4:

- Preliminary image analysis using unsupervised classification
- Segmentation of the results of the classification
- Spectral analysis based on Normalised Difference Vegetation Index (NDVI)
- Spatial analysis modelling combining the four data sources to extract building interest areas.

Classification techniques, which include supervised and unsupervised methods, are widely used in remote sensing (Richard and Jia 1999, Chen at al 2000). In this study, an unsupervised K-Mean clustering procedure is used because supervised classification is dependent upon correctly choosing the spectral classes in the regions of interest. For supervised maximum likelihood classification, each class is to be modelled by a normal probability distribution. If a class happens to be multimodal and this is not resolved, then the modelling cannot be effective and some pixels could not be classified correctly. Users of classification techniques can only specify the information classes, but would have little idea of the number of distinct unimodal groups that the data falls into in multispectral space. Clustering method can be used for this purpose (Richard and Jia 1999). It has been applied in many data analysis fields to enable inherent data structures to be determined. K-Means unsupervised classification calculates initial class means evenly distributed in the data space and then iteratively clusters the pixels into the nearest class using a minimum distance technique. The image is segmented into classes representing by class 1, class 2 etc.. Based on the label number, the user can label those classes to corresponding ground cover types. To eliminate the small number of pixels which have not been correctly classified and result in poor segmentation, a basic approach for integrating spatial contextual information with a spectral classifier involves post classification techniques. A majority filter, which examines spatial patterns of class labels within a local neighbourhood of a given pixel, and reclassifies pixels depending on class frequencies within this neighbourhood has been used in this study. Post classification partitions a classified image into meaningful regions and associates the regions with spatial attributes.

Since it is not possible to differentiate buildings from trees by the DSM only, further image classification is required to assist this task. Baltsavias et. al (1995) introduced a method which can distinguish buildings from trees based on analysis of the histogram of the orientations of edges of the regions. This method is only effective for regularly shaped buildings. Henricsson (1996,1998) described a method of colour attributes and similarity grouping in 3D building reconstruction. He proved that colour is a very important cue in reconstructing a general class of buildings. NDVI transforms

multispectral data into a single image band representing vegetation distribution. Therefore, in this study, NDVI method has been used to differentiate the vegetation from buildings based on their spectral response in suburban areas. NDVI and DSM combine to assist in determining the difference between vegetated and non-vegetated objects above the ground. Areas that are above the surface with low NDVI are likely to be buildings, whereas areas above the surface with high NDVI are likely to be trees. Areas with heights similar to the earth's surface and a high NDVI are likely to be grassland or cultivated areas.

In order to extract building interest areas, a spatial analysis modelling is required to combine the above four data sources of DSM, land cover classification, the results of the post classification and NDVI, which represent 2D and 3D information and are indicated by blue in Figure 2.2. This integration is implemented during interpretation and supplies 'building interest' areas for further shape modelling. The spatial analysis model is a multistage decision method and consists of a number of connected decision nodes. Each decision node only performs part of the segmentation task. The advantages of using a multistage decision method in the spatial analysis model are that different data sources, different sets of features and even different algorithms can be used at each decision stage.

Level set modelling based image segmentation (LSMBIS), shown as module 3 in Figure 2.2 is used to precisely delineate the boundaries of buildings. Because of its importance in this research, it will be described in detail in Chapter 5.

The building interest areas detected by the module 1 and module 2 only indicate that buildings may exist, but the method does not supply the boundaries of the buildings. Hence, as explained earlier, an algorithm is required to extract the building boundaries. Although the building boundary extraction methods have been described in sections 2.8~2.9, the difficulties mentioned in section 2.11 still exist. In order to precisely extract the building boundaries, 2D shape modelling based on the level set method is used. The level set method for curve propagating interfaces is based on mathematical and numerical work of curve and surface motion. It offers a highly robust and accurate method for

tracking curve interfaces moving under complex motions based on user input. This method has not previously been applied for complicated urban image analysis. The multiple data sources used for interpretation and integration above, result in the modelling of shapes for extracted building interest areas. This enables level set shape modelling without user input. The method also overcomes the complicated and difficult 2D inference in a single image. Since the evolving function in the level set method is associated with edge information in the image through a speed function, the precise building boundaries can be obtained.

Multisource evidential reasoning based region evaluation (MEBRE) shown as module 4 in Figure 2.2 is based on multisource evidential reasoning and evaluates the regions using the combination of uncertain evidence from multiple information sources. Uncertainty estimation, based on Dempster-Shafer evidence theory, is presented in Chapter 6.

After the building boundaries are extracted, according to the principles on which the methods for the system are based, a method is required to evaluate the reliability of the extracted building regions. The multiple data sources present the same scene, so they are partly redundant. Since they may highlight different characteristics, they are also partly complementary. None of the data sources provides completely decisive and reliable information. In addition, the information is often imprecise and uncertain. The aim of the evaluation is to improve the decision by increasing the global information while decreasing its imprecision and uncertainty, by making use of redundancy and complementarity. Data fusion techniques have the ability to interpret simultaneously information from multiple data sources, in order to obtain correct conclusions, and hence will be used in this study. Three main numerical data fusion techniques are the probabilistic methods, fuzzy set theory and Dempster-Shafer (Hegarat-Mascle et al 1997). The limitation of probabilistic method such as Bayesian inference is that the imprecision of measurement cannot be modelled. Although the fuzzy fusion approach enables the analysis of many combinations, they must be selected by the user. The advantage of Dempster-Shafer evidence theory is that it provides estimations of imprecision and uncertainty of the information derived from different sources. The

Dempster-Shafer data fusion technique has not previously been applied for evaluating the reliability of building areas. It provides the theoretical basis for evaluating the reliability of the extracted buildings by a statistically-based classification. It can be used to combine the knowledge from several data sources of the same region, to find the intersection of propositions of extracted information derived from these datasets, together with the associated probabilities.

CHAPTER 3

Stereo Image Matching and DSM Extraction

3.1 Introduction

In this chapter, the focus will be on stereo image matching and the DSM extraction. The analysis of single images for building extraction is a very difficult task because it often leads to an ambiguous solution. A DSM derived from stereo image matching is an important information cue for building extraction. Processing steps for stereo image matching and DSM extraction are shown in Figure 3.1 and highlighted in green.

Two approaches for producing a better quality DSM are proposed in this chapter. The first is robust hybrid-based image matching algorithm developed in this study, described in section 3.2. The algorithm combines cross correlation with edge information and supplies a disparity map which preserves the discontinuities in the images. The second is a least square matching using robust estimation, described in section 3.3. The algorithm

can detect breaklines in the images and supply greater accuracy in the computation of disparities in the discontinuous areas.

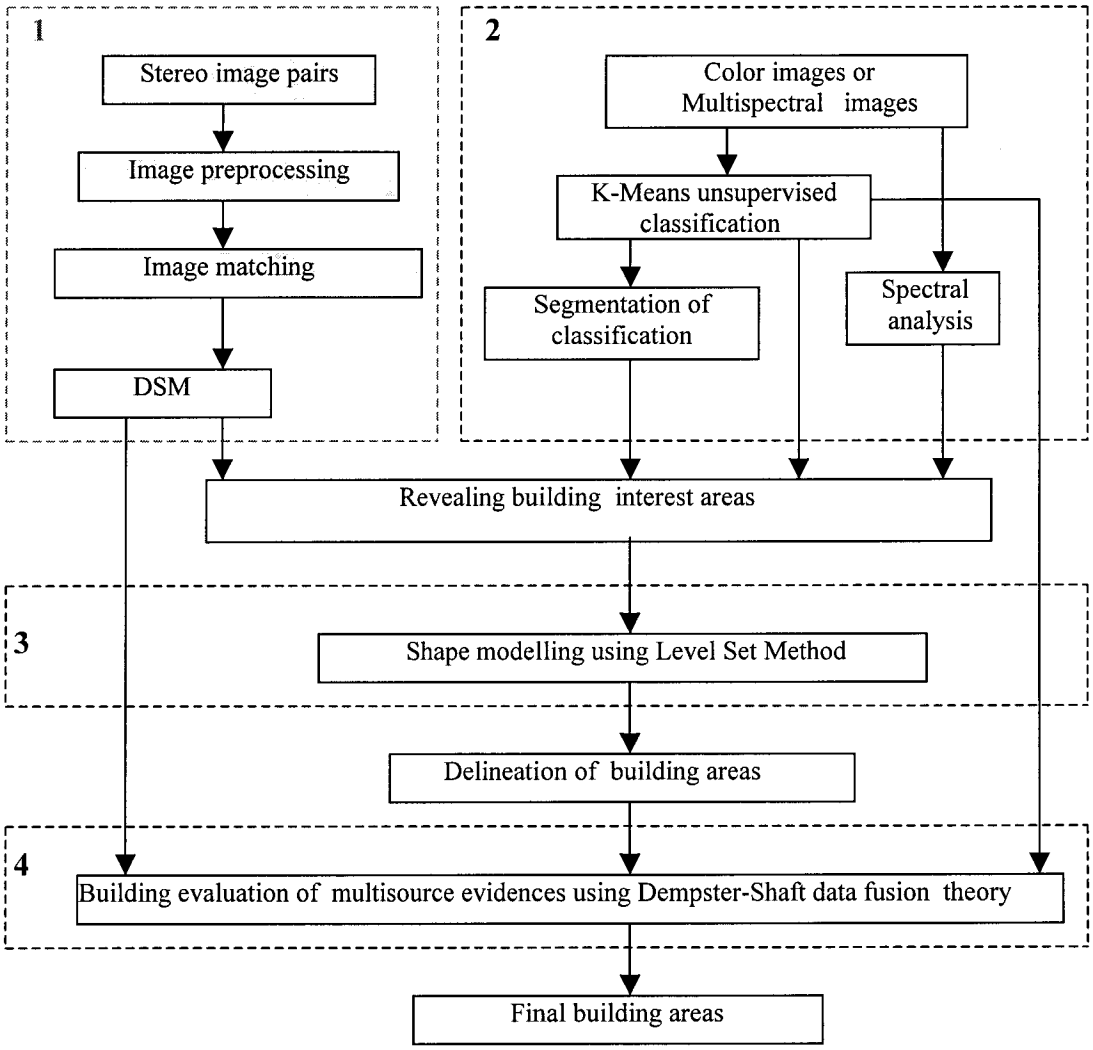


Figure 3.1 DSM extraction in the architecture of the proposed automatic building extraction system

3.2 Robust Hybrid-based Image Matching Algorithm

In order to produce an accurate DSM for subsequent building extraction, a more accurate and robust image matching algorithm has been developed and tested. It considers the discontinuities caused by the man-made structures and trees in the images and supplies

more reliable building areas which would not be recovered if existing area-based matching algorithms were used.

The algorithm uses area-based stereo matching method with information derived from feature-based processing. Figure 3.2 illustrates the overall flow through the new algorithm.

In summary, left and right epipolar images are input to the system. Area-based image matching is used to produce a dense disparity map in which buildings should appear higher than the terrain surface. At the same time, the left image is used to produce an edge map defining the edges of objects. From appropriate start points for region growing and boundary tracking, the object areas such as buildings can been found in the edge map. Then the results from these two parts of Figure 3.2 are combined to produce a dense disparity map which preserves depth discontinuities. This algorithm takes account of the discontinuities in the images and supplies a more accurate disparity map for the further processing.

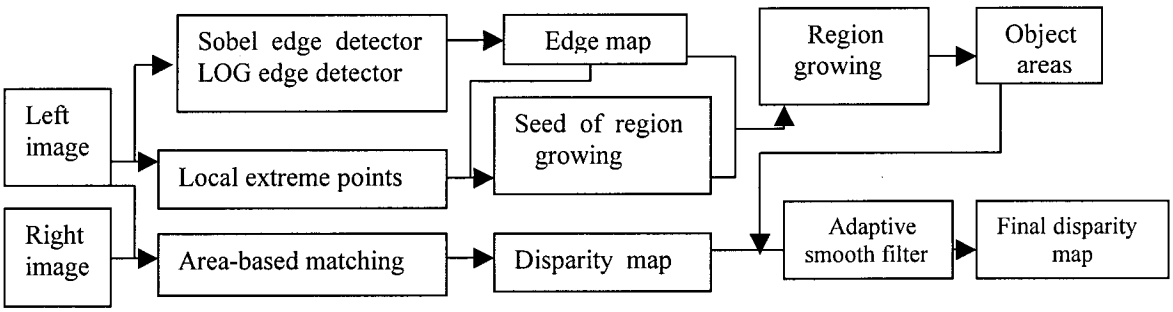


Figure 3.2 Procedure for new image matching algorithm

3.2.1 Area-based image matching

As discussed in Chapter 2.2, area-based matching procedures are based upon the correlation between the pixels in small windows surrounding the corresponding points in two images. There are a number of methods that can be used to find the corresponding points. The correlation equation has been introduced in Section 2.2.1. Since a zero mean normalized correlation coefficient (Wang, 1990) is not affected by brightness difference

of the left and right images, it has been used to determine a dense disparity map for the robust hybrid-based image matching algorithm.

3.2.2 Feature-based processing

The disparity results obtained from area-based matching could be subject to inaccuracies of up to half a window width, due to the averaging effect caused by using a window of significant size for matching (Gonzalez and Woods 1992). In the disparity map, this would be revealed as poorly defined object discontinuities. Matching on buildings and tree tops are also obvious problems. In order to correct the disparity values at the discontinuities, a feature-based extraction procedure has been used to support the image matching.

The Sobel operators have the advantage of providing both a differencing and a smoothing effect. Since derivatives enhance noise, the smoothing effect is a particularly attractive feature of the Sobel operators. Laplacian of Gaussian (LOG) operator combines the two operations of the second derivative and Gaussian filter. As shown in Figure 3.2, the left image has been processed using both Sobel and Laplacian of Gaussian edge detectors. Because the edges can be extracted by thresholding the gradient magnitude derived by Sobel operator, the width of the edges can be more than one pixel. The edge map from LOG operator can include the zero crossings of noise. In this study, the combined edge map from the two edge detectors has been used to keep the true edges. Based on the results of the Sobel edge detection, local maxima of the edge gradient magnitude can be extracted. The points which are in both the combined edge map, derived by two edge operators, and local maxima map were used as the seeds for region growing. Then a seed based pixel to pixel boundary tracking procedure was performed to extract the building boundaries. The method is effective because only the points on the edges will be taken into account in the procedure and it is not necessary to process every pixel in the region.

3.2.3 Adaptive smoothing filter

An adaptive smoothing filter (Saint-Marc et al, 1991) smoothes a signal while preserving discontinuities. This is achieved by repeatedly convolving the signal with a small averaging mask, weighted by a measure of the signal continuity at each point. Let $I^{(0)}(x,y)$ be the image before smoothing. The smoothed image $I^{(t+1)}(x,y)$ at the *(t+1)*th iteration is simply shown as following:

$$I^{(t+1)}(x,y) = \frac{1}{N^{(t)}} \sum_{i=-1}^{+1} \sum_{j=-1}^{+1} I^{(t)}(x+i,y+j) * w^{(t)}(x+i,y+j)$$

$$w^{(t)}(x,y) = e^{-\frac{G_x^2+G_y^2}{2k^2}} \qquad (3.1)$$

$$N^{(t+1)} = \sum_{i=-1}^{+1} \sum_{j=-1}^{+1} w^{(t)}(x+i,y+j)$$

G_x and G_y are gradients in x and y directions. The parameter k determines the magnitude of the edges to be preserved during the smoothing process.

When $w^{(t)}(x,y) = 1$ for all *(x,y)*, the filter smoothes the data equally everywhere, even across the discontinuity areas. If the locations of these discontinuities are known, then the corresponding weights at these points can be set to zero in the filter window, so that no smoothing near discontinuities would occur at these points. Averaging will not occur for points belonging to regions separated by discontinuities since the filter is limited to 3 x 3 in size. For points along discontinuities or points next to edges, the repeated averaging process would force them to be assigned to one of the nearby regions, for example, the region for which they have similar disparity values.

In this procedure, an adaptive smoothing filter is applied to the disparity map derived by area-based matching. Since the area-based matching algorithm has an averaging effect on the disparities and hence causes blur in the disparity map, some discontinuity information may be lost. By only using the disparity map and an adaptive smoothing filter, it is still difficult to recover the discontinuities in the depth information. The weights used for the

adaptive smoothing filter of the disparities are obtained from both the disparity map and edge locations from the feature-based processing. These weights are converted to binary maps and then added, because this is a simple and easy process to implement. After the adaptive smoothing filter with the combined weights is applied to the disparity map, the depth discontinuities should have been enhanced because the weights force the pixels to be assigned to different regions based on the discontinuities. The discontinuities become the boundaries of regions with significantly different height values. The accuracy of the method is dependent on the quality of the disparity values from the area-based matching and the weights derived from edge detection.

Figure 3.3 illustrates some test results of this method. Fig. 3.3(a) and Fig. 3.3(b) are a pair of aerial images with 67×162 pixels in the row and column directions respectively. Fig.3.3(c) illustrates the disparity map obtained from area-based stereo image matching based on correlation coefficients. Fig. 3.3(d) is the result of Sobel edge detection on left image. The result of threshold of Fig. 3.3(d) is Fig. 3.3(e), in which edges are too wide. Fig. 3.3(f) shows too dense edges derived from LOG. Since edges in Fig. 3.3(e) and 3.3(f) are not satisfied, the final edge map from original left image is Fig. 3.3(g), which includes the pixels in both edge maps derived from Sobel in Fig. 3.3(d) and LOG edge detection in Fig. 3.3(f). Fig. 3.3(h) is the left image processed only by adaptive smooth filter. Fig. 3.3(i) is the result of Sobel edge detection for Fig. 3.3(h). Fig. 3.3(j) is the threshold result of Fig. 3.3(i). Fig. 3.3(k) shows the edges from LOG and Fig. 3.3(l) is the final edge map from smoothed left image. Comparing Fig. 3.3(g) with Fig. 3.3(l), the latter shows some false edges deleted and is more suitable for further processing, since this will enable fast and reliable region growing. Fig. 3.3(m) shows the seed points for boundary tracking which are in both the edge map and local maxima map. Fig. 3.3(n) illustrates all the tracked boundaries. There are no seed points on the black edges, so they will not be tracked and can be deleted. Some grey edges are tracked, but they are not closed and can also be deleted. Fig. 3.3(o) shows the final tracked boundaries, on which the gaps have been filled. Fig. 3.3(p) shows the weights derived from Fig. 3.3(c), while Fig. 3.3(q) shows the weights from Fig. 3.3(p) and Fig. 3.3(o). Fig. 3.3(r) is the final disparity map which is related to the heights, and the discontinuities in building areas can be obtained.

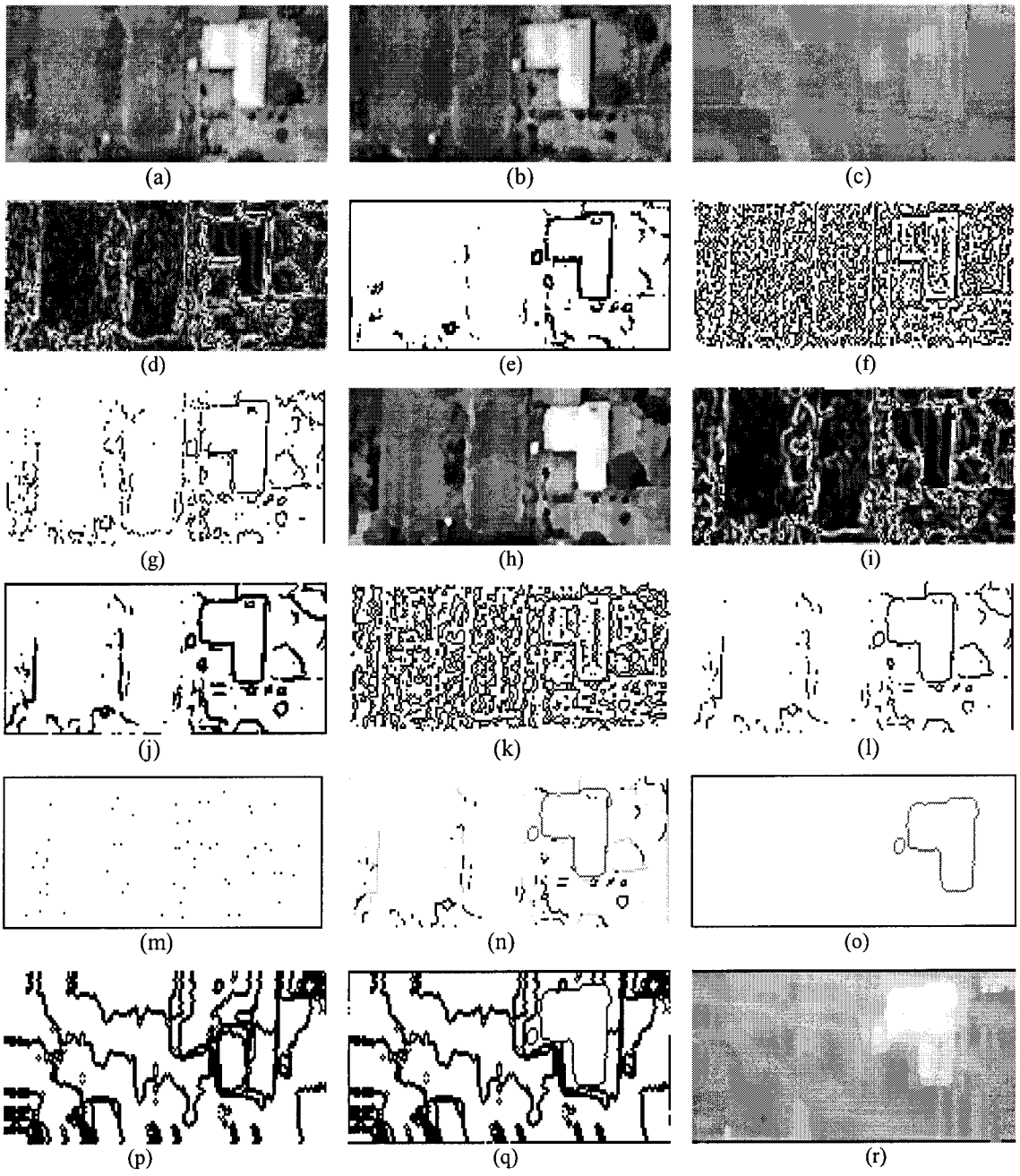


Figure 3.3 Test results of a matching algorithm based on adaptive smoothing

This matching algorithm is a hybrid method which combines cross correlation with edge information. The contribution of the method is that disparity map preserves the discontinuities in the images and can supply more robust DSM for further processing. Investigating the stereo images 3.3(a~b), it can be found that the buildings and trees have the higher elevations. In Figure 3.3 (r), bright grey level represents higher disparity values, while dark represents low disparity values. The disparity values in Figure 3.3(r) prove the

buildings and tree have higher elevations. Comparing Fig. 3.3(c) with Fig. 3.3(r) it is clear that the discontinuities around the building area have been enhanced. Because the building information in the disparity map is important for subsequent processing, robust hybrid-based image matching algorithm can supply more reliable result.

After determining the disparity map, the 3D world coordinates of the object points represented by the pairs of corresponding pixels in two images can be determined by triangulation, using the image coordinates and known orientation of the images. Based on this transformation, the final DSM can be obtained.

3.3 Least Squares Matching Using Robust Estimation

Like the cross-correlation method, the least squares matching algorithm is based on a finite window size and a single disparity value is computed for each window. Hence, it does not work well in the case of windows that include breaklines, or sudden changes in disparity. A further stereo image matching algorithm has been developed based on the least squares formulation, together with robust estimation and weight constraints. It overcomes some of the drawbacks of the least squares method and detects the breaklines and disparity changes during the matching.

3.3.1 Robust estimation theory

A brief introduction to robust estimation is provided first, based on details from (Huber, 1981, Hampel et al., 1986, Rousseeuw et al 1987, Zhang 1995, Zhang et al 1995). Robust estimation techniques enable the location and rejection of outliers in data. Assume r_i is the residual of the i^{th} data value, that is the difference between the i^{th} observation and its true value. The no-weighted least squares method used for computing the most probable values of variables from redundant data minimizes sum of residuals $\sum_i r_i^2$. If there are outliers in the data, the result of the no-weighted least squares will be biased.

The robust estimation reduces the influence of outliers by minimizing the function as follows.

$$\sum_i \rho(r_i) \rightarrow min \tag{3.2}$$

The new function ρ is a function of the residuals, which is defined as a symmetric, positive-definite function with a unique minimum of zero.

Assuming $\boldsymbol{S}=[s_1,...s_j,....s_m]$ is the parameter vector to be estimated. Equation (3.2) is minimized, enforcing the condition of least squares, by taking partial derivatives with respect to each unknown and setting them equal to zero. For each residual r_i, the following equation can be derived:

$$\frac{\partial \rho(r_i)}{\partial s_1} + ... \frac{\partial \rho(r_i)}{\partial s_j} + ... \frac{\partial \rho(r_i)}{\partial s_m} = \sum_j \frac{\partial \rho(r_i)}{\partial s_j} = 0, j = 1,.....m \tag{3.3}$$

The influence function $\varphi(r)$ measures the influence of data on the estimated value of the parameter. Let the influence function $\varphi(r)$ be the derivative of function ρ, $\varphi(r) = d\rho(r)/dr$, then for each residual r_i, the Equation (3.3) can be rewritten in the following form:

$$\sum_j \frac{\partial \rho(r_i)}{\partial s_j} = \sum_j \frac{\partial \rho}{\partial r_i} \frac{\partial r_i}{\partial s_j} = \sum_j \varphi(r_i) \frac{\partial r_i}{\partial s_j} = \varphi(r_i) \sum_j \frac{\partial r_i}{\partial s_j} = 0, j = 1,.....m \tag{3.4}$$

When the weight function is defined as $w(r) = \varphi(r)/r$. Equation (3.4) becomes:

$$\varphi(r_i) \sum_j \frac{\partial r_i}{\partial s_j} = w(r_i) r_i \sum_j \frac{\partial r_i}{\partial s_j} = 0, j = 1,....m \tag{3.5}$$

So Equation (3.5) is the least squares solution of robust estimation, which satisfies robust estimation condition Equation (3.2).

It is well known that the weighted least squares solution attempts to satisfy the following condition through an iterative solution explained earlier by section 2.2.2.

$$\sum_i w(r_i^{(n-1)}) r_i^2 \rightarrow minimum \tag{3.6}$$

Where n is the iterations and $w(r_i^{(n-1)})$ is weight at $(n-1)$ th iteration.

Again, assuming $\boldsymbol{S} = [s_1, ... s_j, ... s_m]$ is the parameter vector to be estimated. By taking partial derivatives with respect to each unknown and setting them equal to zero, for an individual residual r_i, the following sum of gradients must be zero in order to satisfy the least squares condition:

$$w(r_i^{(n-1)}) r_i \sum_j \frac{\partial r_i}{\partial s_j} = 0, j = 1, ... m \tag{3.7}$$

Comparing Equation (3.5) and Equation (3.7), it is clear that the solution of robust estimation is equivalent to the solution of weighted least squares. Therefore, robust estimation technique reduces the influence of outliers in the same manner as weighted least squares method does (Zhang 1995). The comparisons between the results of the robust estimation and the results of weighted least squares can be found in (Harvey,1993). He suggested robust estimation method is not a replacement for weighted least squares method.

If the function $\rho = r^2 / 2$ is chosen, the influence function is $\varphi(r) = r$ and weight function $w(r) = 1$. This means the influence of a data value on the estimate increases linearly with the size of its error. It reveals the non-robustness of the least squares estimate. When an estimator is robust, the influence of a single observation value is insufficient to introduce a significant bias in the results. Many estimators are available. In this study, Huber's estimator has been chosen because it has rarely been shown to be

inferior to other ρ functions (Zhang, 1995). Huber's function, shown in equation (3.8), is a parabola in the vicinity of zero, and increases linearly at a given level $|x|\triangleright k$.

$$Huber\begin{cases} if\ |x| \le k \\ if\ |x| \ge k \end{cases} \quad \rho(x) = \begin{cases} x^2/2 \\ k(|x| - k/2) \end{cases} \quad \varphi(x)\begin{cases} x \\ k\ \mathrm{sgn}(x) \end{cases} \quad w(x) = \begin{cases} 1 \\ k/|x| \end{cases} \tag{3.8}$$

Where x is the residual.

3.3.2 Least squares matching using robust estimation

In the traditional least squares image matching algorithm, corresponding pixels in each window pair form the data for least squares solution, which determines the parameters that transform one window to the other. Therefore, one disparity value is determined for a window. If the window includes breaklines and sudden changes in elevations, the assumption of one disparity value for a window is incorrect.

The algorithm with robust estimation proposed in this study overcomes this limitation by formulating the matching problem without requiring a fixed square window. Instead, it can be based on a whole image. The unknowns are an array of unknown disparity values for every pixel and the parameters of transformation applied to each pixel. In order to detect the discontinuities during the matching, constraints are used. For one dimensional constraints in the x direction, each disparity value must be the similar to its neighbours. A second order constraint is used as follows:

$$-disp_{0(j-1)} + 2disp_{0j} - disp_{0(j+1)} = 0 \tag{3.9}$$

Equation (3.9) requires neighbouring points to have similar disparity values. But in many areas, there are breaklines and discontinuities in the images. To successfully find the discontinuities and supply more accurate disparities, the constraints must be limited by the breakline areas. This is achieved by applying weights to the constraint matrix and the operation of robust estimation as described above.

At the breaklines and discontinuities, the weights of the constraints are set to zero using robust estimation and the corresponding constraints are removed from the solution. The new algorithm can be expressed as follows in one dimension:

$$\begin{aligned} &\text{Observation equation} && \boldsymbol{A}\boldsymbol{x} + \boldsymbol{v}_a = \boldsymbol{L}_a \\ &\text{Constraints equation} && \boldsymbol{C}\boldsymbol{x} + \boldsymbol{v}_c = 0 \end{aligned} \tag{3.10}$$

Where

$$\boldsymbol{x}=\left[da_{00} \quad da_{01} \quad \ldots \quad da_{0N-1} \quad da_1 \quad dh_0 \quad dh_1\right]^T$$

are the unknown increment values of the disparities and transformation parameters which are solved in the matching. Vector $\boldsymbol{v}$ is the residual. Vector $\boldsymbol{L}$ is the vector of observations. $\boldsymbol{A}$ is a coefficient matrix and $\boldsymbol{C}$ is the constraint matrix.

Based on Equation (2.7) in Chapter 2, for a one dimensional case, the coefficient matrix can be shown as follows:

$$\boldsymbol{A}= \begin{bmatrix} \dot{g}_1(0) & 0 & 0 & .. & .. & .. & 0 & 1 & g_1(0) \\ 0 & \dot{g}_1(1) & 0 & .. & .. & .. & \dot{g}_1(1) & 1 & g_1(1) \\ 0 & 0 & \dot{g}_1(2) & .. & .. & .. & 2\dot{g}_1(2) & 1 & g_1(2) \\ : & : & : & : & : & : & : & : & : \\ : & : & : & : & : & : & : & : & : \\ 0 & 0 & 0 & .. & .. & \dot{g}_1(N-1) & N\dot{g}_1(N-1) & 1 & g_1(N-1) \end{bmatrix}_{N,N+3}$$

Where g_1 are the grey levels for pixels and $\dot{g}_1$ are the discrete values of gradient function of g_1 in x direction. $N\dot{g}_1(N-1)$ represents $x_i \dot{g}_x(x_i)$ in Equation (2.7).

$$\boldsymbol{C}= \begin{bmatrix} 0 & 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & .. & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & .. & 0 & 0 & 0 & 0 & 0 \\ : & : & : & : & : & : & : & : & : & : \\ 0 & .. & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & .. & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{N,N+3}$$

$$L_a=[\overline{\Delta g}(0) \quad \overline{\Delta g}(1) \quad .. \quad \overline{\Delta g}(N-1)]^T$$

The two equations in (3.10) can be made into a single equation as following:

$$\begin{bmatrix} A \\ C \end{bmatrix} x + \begin{bmatrix} v_a \\ v_c \end{bmatrix} = \begin{bmatrix} L_a \\ 0 \end{bmatrix} \tag{3.11}$$

The solution of equations can be solved by the method of least squares

$$x=(A^T A+C^T C)^{-1}(A^T L_a) \tag{3.12}$$

With weight $W=\begin{bmatrix} W_a & 0 \\ 0 & W_c \end{bmatrix}$ (3.13)

The solution of equations can be obtained as following

$$x=(A^T W_a A+C^T W_c C)^{-1}(A^T W_a L_a) \tag{3.14}$$

As shown in Equation (3.13), W_a is the weight for least squares equation and W_c is the weight matrix of the constraints. W_a can be set to identity matrix and W_c can be set by Huber weight. If the weight W_c can be forced to zero at the breaklines by robust estimation, the constraints will force the solution to find the discontinuities. Since the neighbouring pixels have different disparities at the breaklines, the residual v of the disparities should be large in the first iteration, smaller weights will be assigned for subsequent iteration. After several iterations, the weights in the breakline areas will tend to zero, while others at pixels which are not the breaklines will increase. Hence, the breaklines and discontinuities will be located.

The above one dimensional constraint system can be extended to two dimensions. For a two dimensional sample grid, the array of unknowns x includes increment values of the disparities denoted as da_{ij} and three transformation parameters da_1, dh_0 and dh_1 for each pixel, with the index i, j denoting the rows and the columns respectively. Because the algorithm is designed for epipolar images and there are no y directional disparities, only three transformation parameters are needed. For an image which has M rows and N columns, there are $M*N+3$ samples in an array of unknown values to be solved for.

By taking the measurement array in row by row order, it can be converted to a column vector, and solved using Equation (3.11). The matrix $\boldsymbol{A}$ will be *(M*N, M*N+3)*. The matrix $\boldsymbol{L_a}$, is also formed by taking the measurements in row by row order. However, the constraint matrix $\boldsymbol{C}$ must be carefully built to accurately indicate the two dimensional case. $\boldsymbol{C}$ will include $\boldsymbol{C_x}$ and $\boldsymbol{C_y}$ constrains in the *x* and *y* directions respectively. The constraints in the *x* direction can be as follows, and as shown in Figure 3.4(a):

$$-disp_{i(j-1)} + 2disp_{ij} - disp_{i(j+1)} = 0 \tag{3.15}$$

The constraints in *y* direction are as follows, and shown in Figure 3.4(b):

$$-disp_{(i-1)j} + 2disp_{ij} - disp_{(i+1)j} = 0 \tag{3.16}$$

Where $0 \le i < M, 0 \le j < N$

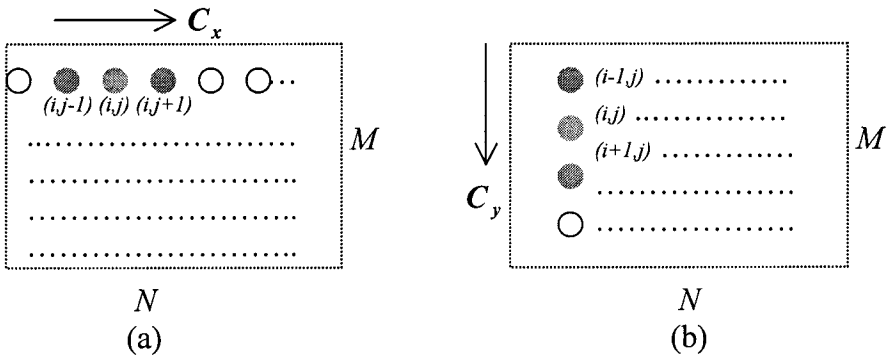


Figure 3.4 Constraints in *x* and *y* directions

For *x* direction constraints in each row of the image, there will be *N-2* constraints, because there are no *x* direction constraints for the first and last columns. This is the same as the one dimensional constraints $\boldsymbol{C}$ in Equation (3.10) and the top and bottom rows are set to 0 in matrix $\boldsymbol{C}$. For *M* rows in the image, there will be *(N-2)*M* *x* direction constraints. Since one constraint requires one row in the constraint matrix, $\boldsymbol{C_x}$ will be a *((N-2)*M, N*M+3)* matrix. To determine the coefficients of the constraint matrix, the *i* and *j* indexes in Equation (3.15) must be converted to an equivalent index corresponding to the unknown vector $\boldsymbol{x}$, which is numbered row by row.

For each row in the matrix $\boldsymbol{C_x}$, three elements in Equation (3.15) are represented by the first, second and third terms in the constraint matrix respectively. All the remaining elements are zero. There are three non zero diagonal elements in the constraint matrix

C_x, which is same as matrix C in Equation (3.10). Since C_x represents two dimensional constraints, the size of C_x is larger than C.

In a similar way to the above, for the y direction constraints in each column of the image, there will be *M-2* constraints, because there are no y direction constraints for the top and bottom rows. C_y will be *((M-2)*N, N*M+3)* matrix as shown below and can be built using Equation (3.16). Three elements in Equation (3.16) are represented by three a_{ij} terms in the constraint matrix respectively.

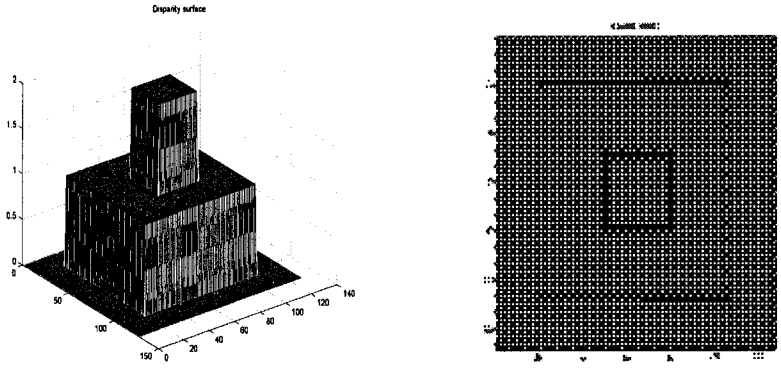
$$C_y = \begin{bmatrix} a_{11} & .. & a_{21} & .. & a_{31} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .. & a_{21} & .. & a_{31} & .. & d_{41} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .. & a_{31} & .. & a_{41} & .. & a_{51} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ : & : & : & : & : & : & : & : & : & : & : & : & : & : & : & : \\ : & : & : & : & : & : & : & : & : & : & : & : & : & : & : & : \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{(m-2)n} & .. & a_{(m-1)n} & .. & a_{mn} & 0 & 0 & 0 \end{bmatrix}$$

After establishing the constraints matrix, Equation (3.14) can be used to calculate the x matrix. In each iteration, the previous residuals are used to obtain the weights for the next iteration and the right image is resampled based on the transformation parameters.

3.3.3 Testing for least squares matching using robust estimation

To test this algorithm, two 128x128 artificial images were used. The grey levels of the left image change linearly and there are no elevation points in the left image. The right image is generated by resampling the left image based on new disparity values. For case 1, there are two blocks, one on top of the other. The bottom block is from 20 to 110 rows and 20 to 110 columns in left image. The top block is from 50 to 80 rows and 50 to 80 columns in left image. The bottom and top blocks have 0.8 and 1.5 pixel disparity values respectively. The right image can be created based on these disparity values. After implementing the least squares matching using robust estimation, the disparity values can be recovered by Equation (3.14). The calculated disparity values are 0.8 and 1.5 pixel, the same as the assumed values. Figure 3.5(a) is the results from stereo image

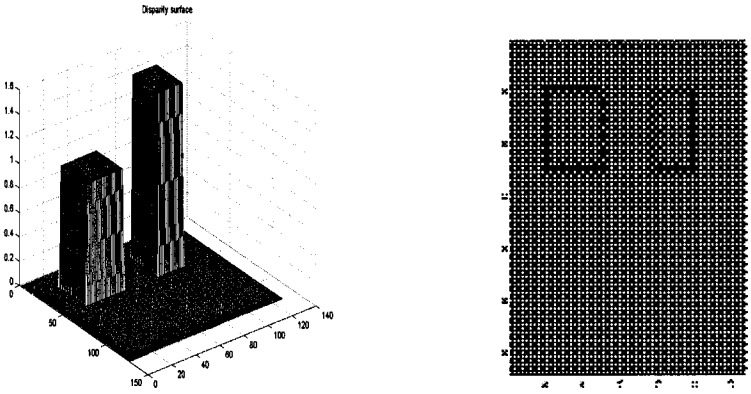
matching, which shows that the correct disparity values corresponding to two blocks have been obtained. Figures 3.5(b) shows the final weights after 20 iterations, corresponding to the breaklines in the images. The blue lines indicate the lowest weight areas which are exactly on the edges of the blocks. The figures show the algorithm can detect the discontinuities during the matching and supply an accurate disparity map.



(a) (b)

Figure 3.5 Disparity surface and weights for case 1

For case 2, there are two blocks beside each other. The first block is from 20 to 50 rows and 20 to 50 columns in left image. The second block is from 20 to 50 rows and 80 to 100 columns in left image. They have 0.8 and 1.3 disparity values respectively. Based on these disparity values, the right image can be created. Figure 3.6(a) shows the 0.8 and 1.3 pixel disparity surface, which have been correctly recovered by Equation (3.14) and



(a) (b)

Figure 3.6 Disparity surface and weights for case 2

correspond to the two blocks. Figure 3.6(b) shows the final weights which correspond to the discontinuities. The blue lines are the lowest constraint weight areas and exactly correspond to the block edges. Although the blocks are close to each other, the algorithm can detect the breaklines and supply an accurate disparity surface.

3.3.4 Summary

The stereo image matching algorithm combines least squares algorithm with the robust estimation of constraint weights. It can detect the breaklines in the images and supply greater accuracy in the computation of disparities in the discontinuous areas. This method will be a robust and efficient matching algorithm for urban environments. This algorithm has limitation compared to traditional least squares method. The algorithm is expansive in computing time as it is 20 times slower to converge due to the computation of the weight image. Because of the limitations, the method is not effective for large real images. The further study is needed to improve the algorithm. For this thesis, this topic is not pursued any further.

The robust hybrid-based image matching algorithm is proposed in Section 3.2. The contribution of the method is that disparity map preserves the discontinuities in the images and can supply more robust DSM for further processing. The accuracy of the method is dependent on the quality of the disparity values from the are-based matching and the weights derived from edge detection.

In sections 3.2 and 3.3, the robust image matching algorithms have been developed. The contribution of the methods is to supply greater accuracy in the computation of disparities in the discontinuities. The DSM derived from the matching algorithms can be used for subsequent building extraction. Since completely implementing the software package for DSM extract is a huge work, including orientation, epipolar image generation, matching and DSM extraction, only the matching algorithms described in this thesis have been tested. The study shows the principle of robust image matching algorithms is useful for automatic building extraction over urban areas. In the future, the potential software package can be set up using all the algorithms.

As mentioned earlier in section 2.12, a high accuracy DSM is important information cue for the proposed automatic building extraction system. In the next chapter, the DSM derived from LH Systems' Socet Set v4.2 has been used for building extraction. This DSM is obtained based on the conventional stereo matching method. If the methods of computing a DSM based on the robust image matching proposed in this thesis are effective, a more accurate DSM should be available.

CHAPTER 4

Low-level Image Analysis and Interpretation

4.1 Introduction

This chapter will focus on the module of low-level image analysis and interpretation. The details of module are described in section 4.2, which includes theory of unsupervised classification and its application, post classification filtering and extraction of spectral properties. In section 4.3, a spatial analysis model is described, which analyses the relationships between spatial analysis and spectral properties, and reveals building interest areas based on spatial attributes, such as height and size, and features related to spectral properties of the image.

4.2 Low-Level Image Analysis and Interpretation

The aim of low-level image analysis and interpretation (LLIAI) in this research is to extract significant features from images and classify them into the meaningful regions. Then the building interest areas are extracted using classification and spectral analysis techniques. The components of the module LLIAI are shown in Figure 4.1 and highlighted in green. The module includes unsupervised classification, post classification, spectral analysis and extracted building interest areas.

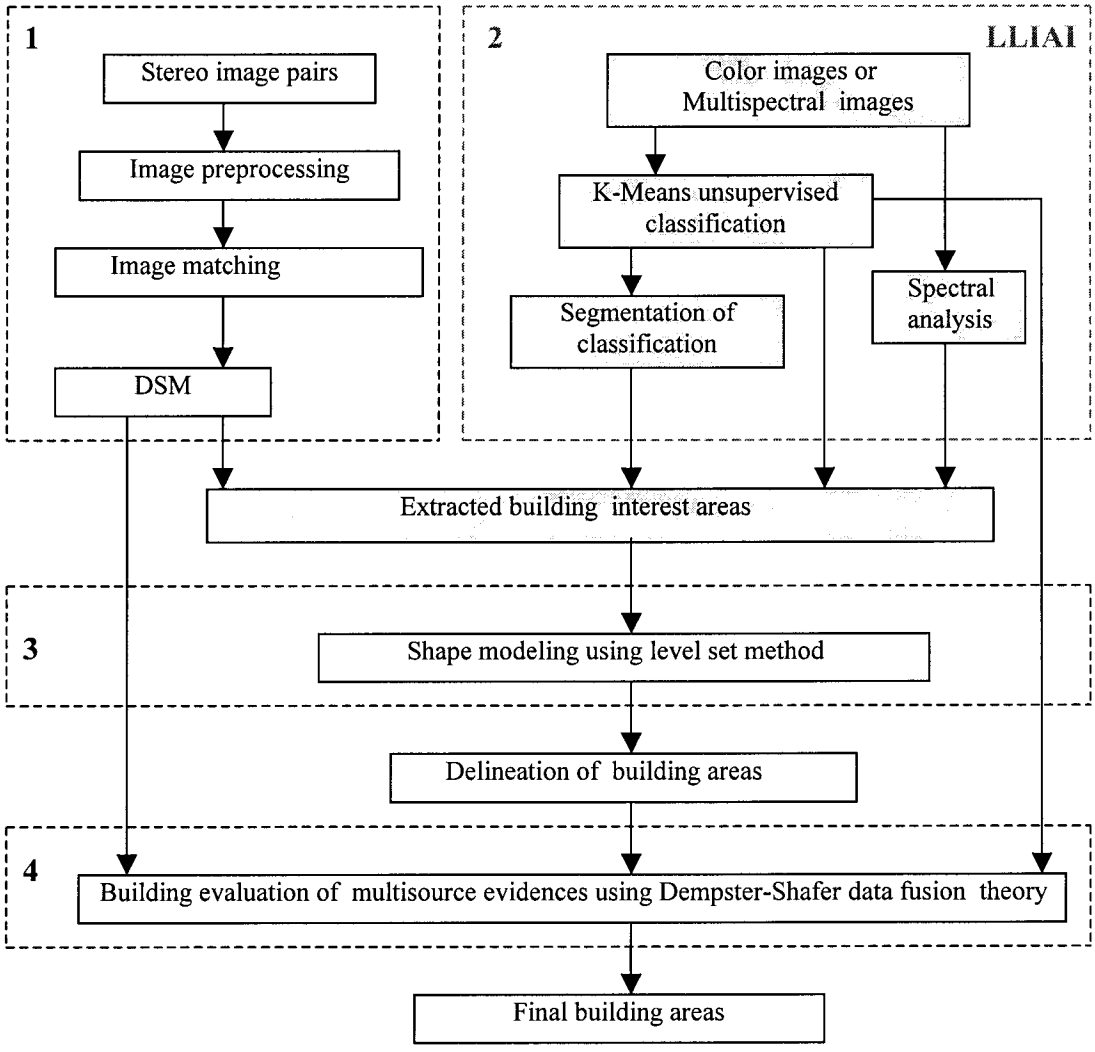


Figure 4.1 LLIAI module in the architecture of the proposed automatic building extraction system

4.2.1 Unsupervised class detection

Multispectral image classification can be used to detect individual object primitives, such as buildings, trees and roads. It is also helpful to reduce the complexity of image content for the next processing step of object structure detection. Many supervised and unsupervised classification algorithms have been developed (Ripley, 1996, Richards and Jia, 1999, Stein et al, 1999 and Chen et al, 2002). As mentioned in section 2.12, image classification is the essential step to differentiate buildings from other objects.

In order to find building interest areas, a K-Means unsupervised classification algorithm is used because it is an automatic procedure and is an easy operation for the user. K-Means unsupervised classification calculates initial class means evenly distributed in the data space and then iteratively clusters the pixels into their nearest class using a minimum distance technique. Each iteration recalculates class means based on all the pixels in one class cluster and reclassifies pixels with respect to the new means. All pixels are classified to the nearest class unless a standard deviation or distance threshold is specified, in which case some pixels may remain unclassified if they do not meet the selected criteria. This process continues until the number of pixels in each class changes by less than the selected pixel change threshold or the maximum number of iterations is reached.

K-Means classification is sensitive to initial values, since the choice of the initial locations of the cluster centres has an influence on the time it takes to reach a final, acceptable clustering. The K-Means technique also requires a pre-specified number of cluster centers. If the initial specification is too high, then post merging and deletions are needed. On the other hand, if the initial number of clusters is too small, multiple class groups will be merged to one class group and the accuracy of classification will be degraded.

The initial locations of the cluster centers are generally chosen uniformly spaced along the multidimensional diagonal of the multispectral pixel space. This is a line from the origin to the point corresponding to the maximum spectral brightness value. From

previously computed histograms, the initial locations of the cluster centers can be refined within the actual range of brightness values. Choice of the initial locations of clusters in the above method is a reasonable and effective way because the centres are well spread over the multispectral space where many spectral classes occur.

The objective of the K-Means algorithm is to minimize the cluster variability. The objective function is the sums of squares of the distance (error) between each pixel and its assigned cluster center.

$$SSD = \sum_{\forall x} [x - C(x)]^2 \tag{4.1}$$

where $C(x)$ is the mean of the cluster that pixel x is assigned to, and x is the grey level of a pixel.

Minimizing the *SSD* is equivalent to minimizing the Mean Squared Error (MSE), which is a measure of the cluster variability.

$$MSE = \frac{\sum_{\forall x} [x - C(x)]^2}{(N - c)b} = \frac{SSD}{(N - c)b} \tag{4.2}$$

Where N is the number of pixels, c indicates the number of clusters and b is the number of spectral bands.

When the clustering is completed, pixels within a given class group are given a symbol to indicate that they belong to the same spectral class. Using these symbols, a cluster map can be produced, which allows a classification to be made. If some pixels with a given label can be identified with a particular ground cover type, then all pixels with the same label can be associated with that class. The classification map interprets original multispectral data into meaningful regions. The classification analysis can be implemented based on the clustering computations, and user defined training fields for every cover type are not needed.

4.2.2 Post classification filtering

Post classification filtering can refine the classification results and remove errors. Based on the labels of classes on the classification map, post classification filtering can extract regions of connected pixels that are contained in the same class. For a class of buildings, post classification filtering supplies the extracted building image by segmentation of the classification. Each connected region is given a unique digital number. The minimum number of pixels that must be contained in a region can be specified in order to remove the isolated noise.

In order to make a decision and reveal the building interest areas in the spatial analysis model described in section 4.3, the region attributes for the extracted building segmentation image from post classification filtering are needed. Based on each region, the following seven types of post-filtering attributes are extracted:

- mean gray level
- area
- perimeter
- area to perimeter ratio
- width of principal axis (PA)
- height perpendicular to the PA
- logarithm of the ratio of height to width = $\log_{10}(\frac{height}{width})$

These attributes are used to filter out extracted regions which are unlikely to be building areas, based on the seven feature values for every region.

4.2.3 Vegetation detection using NDVI

The application of remote sensing for recognition and measurement of the variations of types and densities of living forests, fields, and crops, is based on their spectral responses. Because the leaves absorb a great deal of the red and blue light and reflect much of the green, healthy vegetation appears green to our eyes. In addition to reflecting green light, healthy leaves also reflect radiation in the near infrared portion of the spectrum, which is particularly useful for detection.

Many mathematical formulas for combining bands of multispectral data have been found to be useful as indicators of vegetation. Normalized Difference Vegetation Index (NDVI) transforms multispectral data into a single image band representing vegetation distribution. The NDVI values indicate the amount of vegetation present in the image. NDVI is calculated from

$$NDVI = \frac{NIR - R}{NIR + R} \tag{4.3}$$

In the building extraction procedure, the DSM is an important cue since it supplies the height information. As building and tree areas have similar heights above the ground, it is necessary to find a method to differentiating them. The image comprising the NDVI values has been used for this purpose.

A multispectral image is an important dataset for NDVI calculation. Multispectral images of 1 and 2 metre resolution were supplied by SpecTerra Systems Pty Ltd. These images were evaluated for use in the proposed automatic building extraction system. After a thorough analysis these images proved to be inadequate as it was not possible to generate a high accuracy DSM from them. As SpecTerra Systems could not supply any high resolution data, it was more practical and appropriate to use colour aerial images with 0.6 metre resolution. The visible vegetation index (VVI) was used to substitute the NDVI and calculated using the following formula:

$$visible\ vegetation\ index = \frac{G - R}{G + R} \tag{4.4}$$

The advantage of using colour, demonstrated by the experiments in section 4.3.2, is that colour information is an important evidence source and can aid in differentiating buildings from trees.

4.3 Spatial Analysis Model Combining Four Data Sources

The VVI and DSM are two key parameters which assist to determining the differences between vegetated and non-vegetated objects on flat terrain. Simplistically, the area which has height above some limit, must either be a tree or building. Areas with low VVI which are above the surface area are likely to be buildings, whereas areas with high VVI and which are above the surface are likely to be trees. Areas with high VVI and located on the height of the earth surfaces are likely to be grassland or cultivated areas. The spatial analysis model combining four data sources, comprising the 2D and 3D information, has been used to determine which areas are possible building areas.

4.3.1 Spatial analysis model

The spatial analysis model used to determine the correct building regions is a multistage decision method implemented using the ArcView GIS software from ESRI. A series of consecutive decisions is made. Figure 4.2 shows the decision tree in the spatial analysis model, consisting of a number of connected decision nodes, none of which is expected to perform the complete segmentation of the image data set. Each decision node (DN) only performs part of the task as noted in the Figure. The advantages of using a multistage decision tree in spatial analysis model are that different data sources, different sets of features and even different algorithms can be used at each decision stage.

The fundamental purpose of Geographic Information Systems (GIS) is the analysis of geographic data. This data is organized into layers or thematic maps that can be graphically overlaid by the software system. However, the power of a GIS is not limited to creating composite imagery by performing data layering operations. Its principal usefulness is that the GIS supports a query system on the database which allows the user to perform both visual and analytical inquiries with a high degree of sophistication and complexity. ArcView is a powerful and easy-to-use tool which visualizes, explores, queries and analyzes data spatially. Map Query in ArcView is an analysis function based on multiple thematic layers. It allows the user to select areas spatially by defining a query

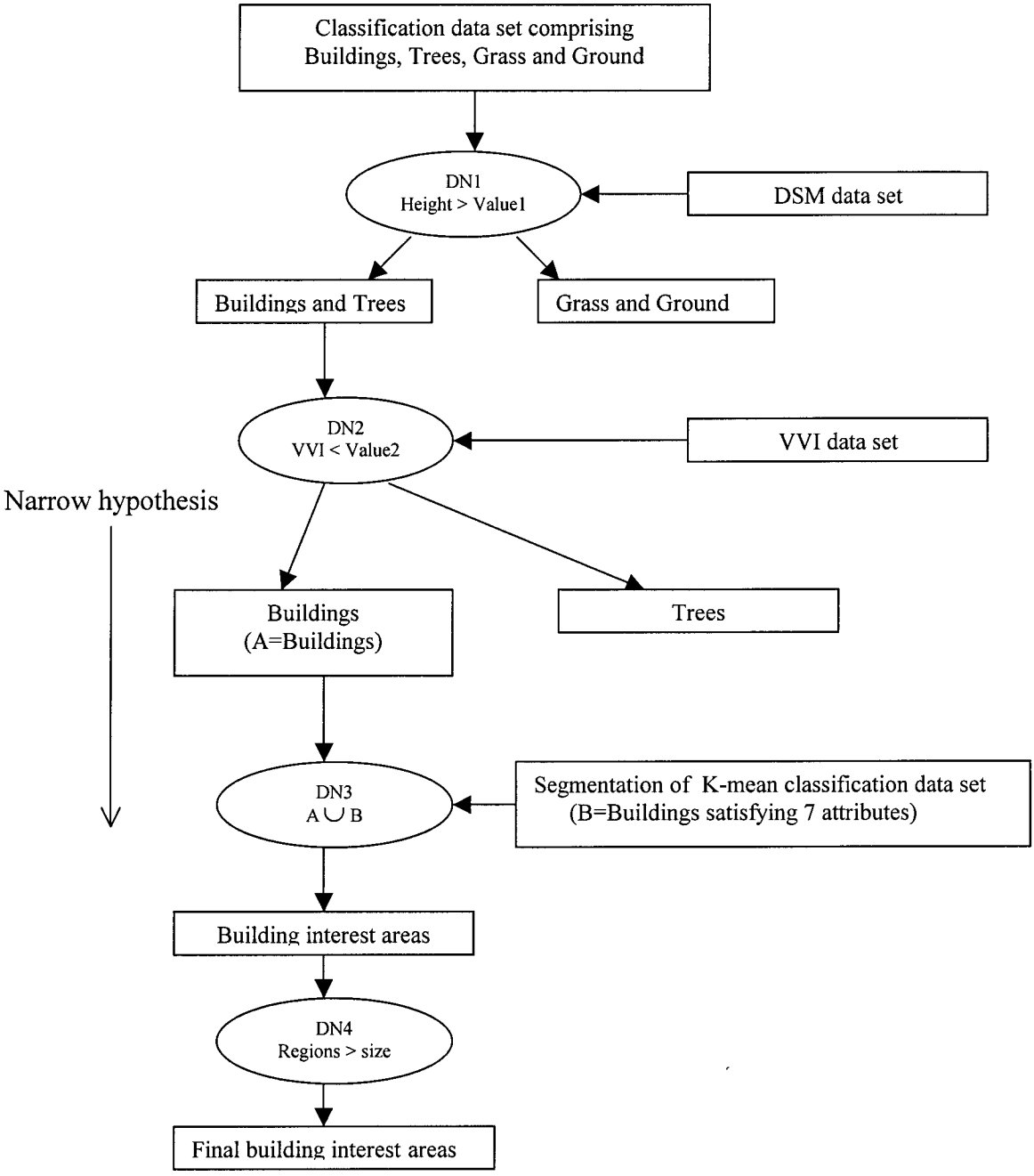


Figure 4.2 Decision trees in spatial analysis model combining four data sources

equation based on the values of the thematical layers. Four information layers consisting of the land cover classification, segmentation of classification, DSM and VVI are input into ArcView. Map Queries are used to implement the decision by the user defined query operations based on decision trees in Figure 4.2. When an area is concluded as a building interest area, it should have a height above the terrain surface, low VVI value and

satisfies the minimum number of pixels for building region. The majority of building interest areas can be extracted correctly using this process.

4.3.2 Results of spatial analysis

Figure 4.3 illustrates a pair of colour aerial images with 580 × 560 pixels in the row and

(left)

(right)

Figure 4.3 Stereo aerial images

column directions respectively. The focal length of the camera is 153mm and resolution is 0.6 meter. The flying height is 3070 meter.

Figure 4.4 illustrates the DSM map obtained from stereo image matching method. Figure 4.5 is the result of unsupervised classification using the K-Means method. Yellow is assigned to building areas, red and green are the vegetation areas and blue is assigned to ground areas. Classification for some areas is not correct, because some ground areas are classified as buildings. Based on the classification of Figure 4.5, using a post classification procedure in ENVI software package without filtering attributes criteria, a segmentation image was created to show the building areas in the classification map as illustrated in Figure 4.6. After this stage, most of the building areas have been detected, but several dark roof buildings have been missed and some road areas have been wrongly assigned as buildings.

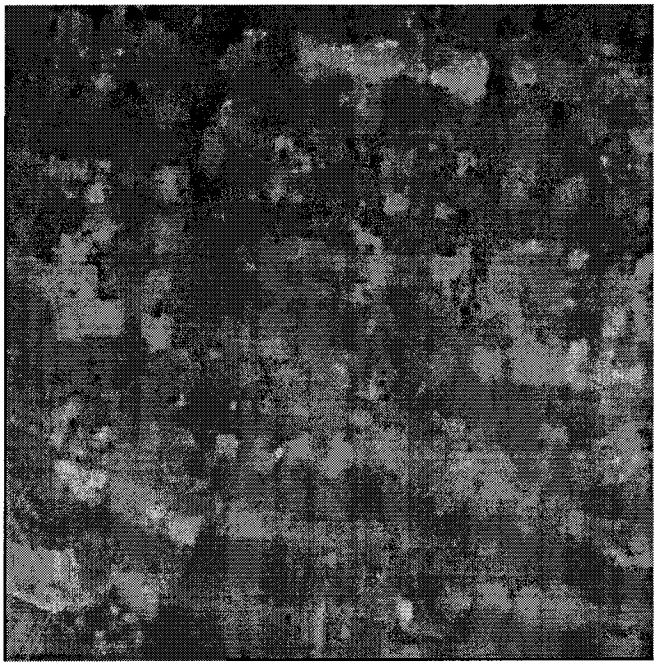


Figure 4.4 DSM from stereo image matching

Figure 4.5 Unsupervised classification by clustering

The results of processing by VVI are shown in Figure 4.7, where high VVI represent the vegetation and the areas with low VVI represent the ground and building areas. There is also a difference revealed between grass and trees.

Figure 4.6 Building areas extracted by segmentation in Figure 4.5

Figure 4.7 VVI image revealing vegetation

The four information data sources of Figures 4.4~4.7 are input into the spatial analysis model. Using the Map Queries operation in ArcView, the possible building areas are extracted in Figure 4.8. By implementing a region growing algorithm, the small spots which do not belong to buildings can be deleted. The final building interest areas are derived by the module of low-level image analysis and interpretation and shown in Figure 4.9.

Figure 4.10 is the ortho image. The building interest areas are overlaid on the ortho image as shown in Figure 4.11. Comparing Figure 4.10 with Figure 4.11, the road areas wrongly assigned as buildings in Figure 4.5 have been deleted and all the correct building areas from classification have been successfully maintained. Several red and dark roof buildings have been missed. Because the further processing introduced in next chapters is based on the building interested areas shown in Figure 4.11, the missed the buildings will never be recovered.

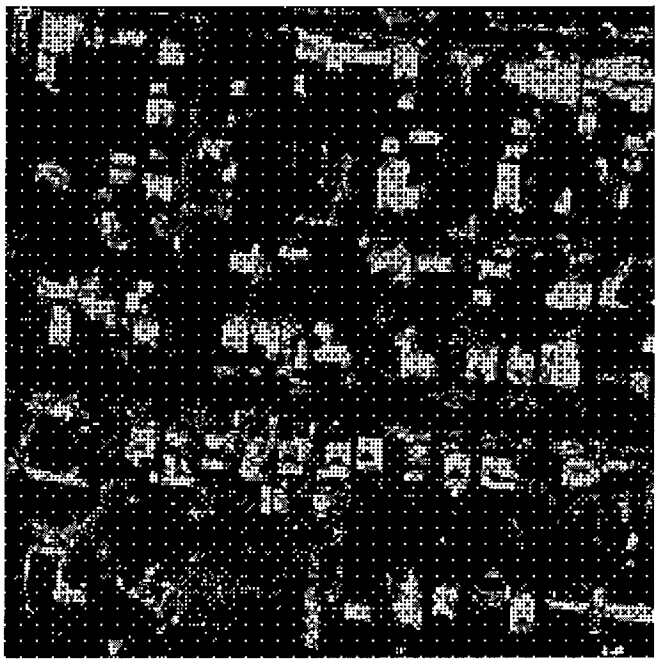


Figure 4.8 Building areas from spatial analysis model combining four data sources

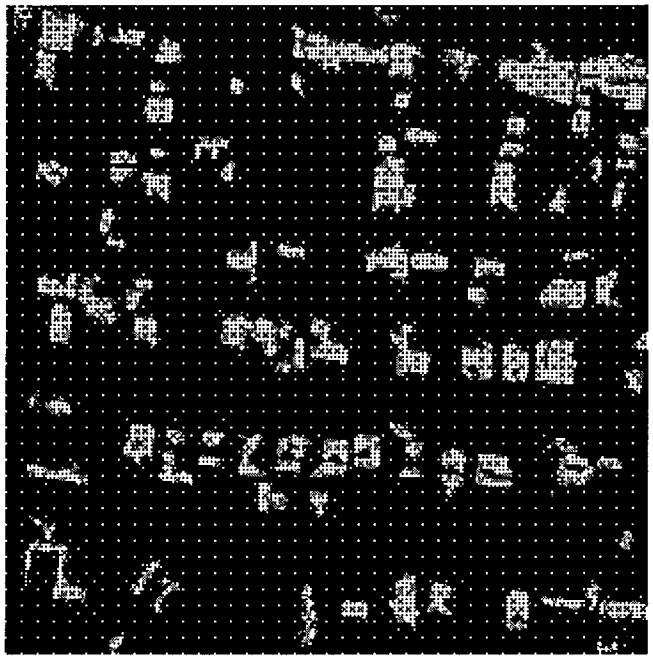


Figure 4.9 Final building interest areas

Figure 4.10 The ortho image

Figure 4.11 The building interest areas overlapped on ortho image

This is an example of how the application of low-level image analysis and interpretation (LLIAI) module operates for the extraction of building interest areas. Based on the above

test data, although several buildings are missed, the results are encouraging. The example shows that the low-level image analysis and interpretation is an efficient module to supply building interested areas for further processing. The extracted building interest areas will be input into level set modelling based image segmentation (LSMBIS) module, which delineates the precision building boundaries. The details of LSMBIS module will be described in next chapter.

CHAPTER 5

Level Set Modelling Based Image Segmentation

5.1 Introduction

From the process describe that in the previous chapter, the building interest areas have been detected. The interest areas only indicated buildings may exist and cannot supply the boundaries of buildings. In order to precisely delineate the building boundaries, the level set modelling based image segmentation (LSMBIS) module will be introduced in this chapter. The traditional techniques for tracking curved boundaries will be presented in Section 5.2. The details of LSMBIS will be described in Section 5.3. The LSMBIS module is based on mathematical and numerical theory. Shape modelling is carried out on the extracted building interest areas derived with previous chapter, thus the complicated and difficult 2D segmentation inference based on single information dataset

can be overcome. Because the LSMBIS module is integrated intellectually with other modules in the system, it can work effectively without user input. In section 6.4, the examples based on LSMBIS module to delineate the boundaries of buildings are presented.

5.2 Traditional Techniques for Tracking Curve Boundary

There are many methods can be used to tracking curved boundaries and more details can be found in Sethian (1999). In this section, the traditional Lagrangian method will be described. After studying the algorithm, the drawbacks can be investigated.

Consider a closed curve moving in the plane. Let $\gamma(0)$ be a smooth, closed initial curve in Euclidean plane R^2 as shown in Figure 5.1, and let $\gamma(t)$ be the one-parameter family of curves generated by moving $\gamma(0)$ along its normal vector field with speed $F(K)$, F is a given scalar function of the curvature K. If $\boldsymbol{p}$ is the position vector of the curve, t is time, and $\vec{n}$ is the unit normal to the curve, then $\vec{n} \cdot \boldsymbol{p}_t = F$.

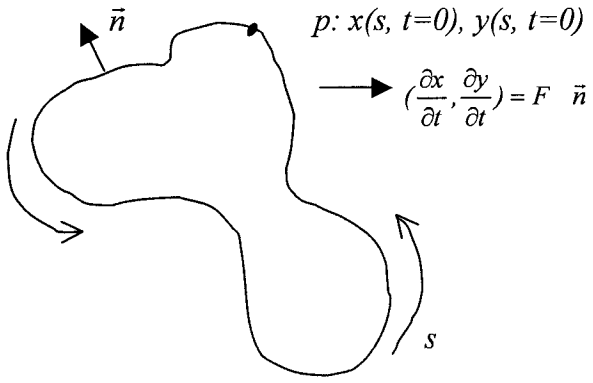


Figure 5.1 Parameterised propagating curve

Let $\boldsymbol{p}(s,t)$ be the position vector which parameterises $\gamma(t)$ by s, $0 \le s \le S$. Assume periodic boundary conditions, $\boldsymbol{p}(0, t) = \boldsymbol{p}(S, t)$. Since the curve is parameterised, the interior is on the left in the direction of increasing s as shown in Figure 5.1.

For a parametric curve $\begin{cases} x = x(t) \\ y = y(t) \end{cases}$, based on the definition of the curvature, the curvature at point s inside the speed function *F(K)* is shown as follows:

$$K = \frac{y_{ss}x_s - x_{ss}y_s}{(x_s^2 + y_s^2)^{3/2}}$$

Where $x_s = \frac{dx(t)}{dt}|_{t=s}$, $y_s = \frac{dy(t)}{dt}|_{t=s}$, $x_{ss} = \frac{d^2x(t)}{dt^2}|_{t=s}$, $y_{ss} = \frac{d^2y(t)}{dt^2}|_{t=s}$

The normal equation at point s is:

$$x_s[x(t)-x(s)]+y_s[y(t)-y(s)]=0$$

For a straight line *l*, if its slope is *k*, then the vector of the straight line *l* is:

$$\vec{l} = \vec{i}_x + k \cdot \vec{i}_y$$

Where $\vec{i}_x$ and $\vec{i}_y$ are unit vectors in *x* and *y* directions respectively.

For normal equation, its slope is $(-\frac{x_s}{y_s})$, the vector of normal is shown as follows:

$$\vec{n}' = \vec{i}_x + (-\frac{x_s}{y_s}) \cdot \vec{i}_y$$

In order to normalize the $\vec{n}'$, $\vec{n}'$ is divided by $\sqrt{1 + (-\frac{x_s}{y_s})^2}$

$$\vec{n} = \frac{\vec{n}'}{\sqrt{1+(-\frac{x_s}{y_s})^2}} = \frac{\vec{i}_x + (-\frac{x_s}{y_s}) \cdot \vec{i}_y}{\sqrt{1+(-\frac{x_s}{y_s})^2}} = \frac{y_s\vec{i}_x - x_s\vec{i}_y}{\sqrt{x_s^2 + y_s^2}}$$

The normal can be described as

$$\vec{n} = (y_s, -x_s)/(x_s^2 + y_s^2)^{1/2} \tag{5.1}$$

The equations of curve motion can be presented for individual components of *p(x, y)* as:

$$x_t = F[\frac{y_{ss}x_s - x_{ss}y_s}{(x_s^2 + y_s^2)^{3/2}}](\frac{y_s}{(x_s^2 + y_s^2)^{1/2}}),$$
$$y_t = -F[\frac{y_{ss}x_s - x_{ss}y_s}{(x_s^2 + y_s^2)^{3/2}}](\frac{x_s}{(x_s^2 + y_s^2)^{1/2}}) \quad (5.2)$$

Where x_t and y_t are the x and y changes at time t respectively. y_{ss} and x_{ss} are the second order derivatives for curve s respectively. x_s and y_s are the first order derivatives.

Since *(x(s, t), y(s, t))* describes the curve moving, this is a Lagrangian representation for tracking curved boundaries. A standard approach to modelling moving curves comes from discretizing the Lagrangian form of the equations (5.2). In this method, the parameterisation is discretized into a set of points whose positions at any time are used to reconstruct the curve.

Assume the parameterisation interval [*0, S*] is divided into M equal intervals of size Δs, then $M+1$ mesh points $s_i = i\Delta s, i = 0,....M$ are created. Time is divided into equal intervals of length Δt. Each mesh point $i\Delta s$ at each time step $n\Delta t$ is a marker point (x_i^n, y_i^n) on the moving curve. The aim of the method is to use a numerical algorithm that will supply new values (x_i^{n+1}, y_i^{n+1}) from the previous position (x_i^n, y_i^n).

At each marker point, parameter derivatives can be approximated using neighbouring mesh points. Based on Taylor's series, central difference approximations can be shown as follows:

$$\frac{dx_i^n}{ds} \approx \frac{x_{i+1}^n - x_{i-1}^n}{2\Delta s}$$
$$\frac{dy_i^n}{ds} \approx \frac{y_{i+1}^n - y_{i-1}^n}{2\Delta s} \quad (5.3)$$
$$\frac{d^2x_i^n}{ds^2} \approx \frac{x_{i+1}^n - 2x_i^n + x_{i-1}^n}{\Delta s^2}$$

$$\frac{d^2 y_i^n}{ds^2} \approx \frac{y_{i+1}^n - 2y_i^n + y_{i-1}^n}{\Delta s^2} \tag{5.4}$$

Similarly, using the forward difference approximations, time derivatives can be shown as:

$$\frac{dx_i^n}{dt} \approx \frac{x_i^{n+1} - x_i^n}{\Delta t}$$

$$\frac{dy_i^n}{dt} \approx \frac{y_i^{n+1} - y_i^n}{\Delta t} \tag{5.5}$$

Substituting Equations (5.3)~(5.5) into Equation (5.2), Equation (5.6) can be obtained.

$$(x_i^{n+1}, y_i^{n+1}) = (x_i^n, y_i^n) + \Delta t F(K_i^n) \frac{((y_{i+1}^n - y_{i-1}^n), -(x_{i+1}^n - x_{i-1}^n))}{((x_{i+1}^n - x_{i-1}^n)^2 + (y_{i+1}^n - y_{i-1}^n)^2)^{1/2}} \tag{5.6}$$

Where

$$K_i^n = 4 \frac{(y_{i+1}^n - 2y_i^n + y_{i-1}^n)(x_{i+1}^n - x_{i-1}^n) - (x_{i+1}^n - 2x_i^n + x_{i-1}^n)(y_{i+1}^n - y_{i-1}^n)}{((x_{i+1}^n - x_{i-1}^n)^2 + (y_{i+1}^n - y_{i-1}^n)^2)^{3/2}}$$

This is the equation for updating the positions of marker points in the curve. These discrete markers points are updated in time by approximating the spatial derivatives in the equations of motion, and advancing their positions. However, there are several problems with this approach.

- A very sensitive calculation can be obtained. This is because the fixed discretization interval Δs is no longer in the Equation (5.6). When the marker points come together, the numerator and denominator on right hand side of Equation (5.6) approach zero.
- Small errors in the computed particle positions are significantly amplified by the curvature term, and calculations become unstable unless an extremely small time step is used.
- In the absence of a smoothing curvature term, singularities develop in the propagating curve, and an entropy condition must be invoked to produce a reasonably weak solution which can overcome the formation of the singularity.

- The topological changes are difficult to manage as the evolving interface breaks and merges.
- Significant problems occur in the extension of this technique to three dimensions.

Lagrangian approximation provides a numerical method to describe the moving curve. It can be accurate for small-scale motions of curve. Since the method uses a local representation of the curve rather than a global one, it could not take into account the proper entropy conditions and weak solutions. Thus, the method can suffer instability and some limitations.

5.3 Image Segmentation using Level Set Modelling

In order to overcome the drawbacks of Lagrangian method, the level set method for propagating interfaces was proposed by Sethian(1985,1999) and Osher (1988). It is based on mathematical and numerical work of curve and surface motion. It offers a highly robust and accurate method for tracking interfaces moving under complex motions.

5.3.1 Fundamentals of level set shape recovery using curve propagation

Level set method represents the curve $\gamma(t)$ as the level set $\{\Phi = 0\}$ of a function Φ. Thus, given a moving, closed hypersurface $\gamma(0)$, a formulation for the motion of the hypersurface propagating along its normal direction with speed F is needed. F can be a function of various arguments, including the curvature, normal direction, etc. The main idea is to embed this propagating interface as the zero level set of a higher dimensional function Φ. Let $\Phi(\boldsymbol{p}, t = 0)$, where $\boldsymbol{p} \in R^N$ is defined by

$$\Phi(\boldsymbol{p}, t = t_1) = \pm d \tag{5.7}$$

where d is the distance from $\boldsymbol{p}$ to $\gamma(t = 0)$, and the plus sign is chosen if the point $\boldsymbol{p}$ is outside the initial hyersurface $\gamma(t = 0)$, minus sign is chosen if the point $\boldsymbol{p}$ is inside the initial hypersurface. Thus, the initial function $\Phi(\boldsymbol{p}, t = 0) : \boldsymbol{p} \in R^N$ can be defined as follows:

$$\gamma(t=0) = (\boldsymbol{p} \mid \Phi(\boldsymbol{p}, t=0) = 0) \tag{5.8}$$

Figure 5.2 illustrates the example of an expanding circle. Suppose the initial curve γ at $t=0$ is a circle in the xy plane and shown in Figure 5.2(a). Assume the solid blue circle is the level set $\{\Phi = 0\}$ of an initial surface $z = \Phi(x, y, t=0)$ in R^3 shown in Figure 5.2(b). At time $t=0$, the solid blue circle can be evolved into broken green curve shown in Figure 5.2(b). This broken green curve can be assigned as level set at time $t=1$ and shown as solid green circile in Figure 5.2(d). Based on the solid green circle in Figure 5.2(d), the final evolved curve becomes the broken purple circle. As illustration in Figure 5.2(c) and Figure 5.2(d), the one parameter family of moving curves $\gamma(t)$ can be matched with a one parameter family of moving surfaces in such a way that the level set $\{\Phi = 0\}$ always yields the moving curve.

Now, it is necessary to produce an equation for the evolving function $\Phi(\boldsymbol{p}, t)$ which contains the embedded motion of $\gamma(t)$ as the level set $\{\Phi = 0\}$. Let $\boldsymbol{p}(t), t \in [0, \infty]$ be the path of a point on the propagating curve. That is, $\boldsymbol{p}(t=0)$ is a point on the initial curve $\gamma(t=0)$, and $\boldsymbol{p}_t = F(\boldsymbol{p}(t))$ with the vector $\boldsymbol{p}_t$ normal to the curve at $\boldsymbol{p}(t)$. Since the evolving function Φ is always zero on the propagating hypersurface, the following level set equation can be obtained

$$\Phi(\boldsymbol{p}(t), t) = 0 \tag{5.9}$$

Using the chain rule in Sethian(1999, 1995), Equation (5.9) become as:

$$\Phi_t + \nabla\Phi(\boldsymbol{p}(t), t) \cdot \boldsymbol{p}_t(t) = 0 \tag{5.10}$$

Because speed F is in the outward normal direction as shown in Figure 5.1, then

$$\boldsymbol{p}_t(t) \cdot n = F \tag{5.11}$$

where $n = \nabla\Phi / |\nabla\Phi|$

Substituting the Equation (5.11) into Equation (5.10), the evolution equation for Φ can be obtained as following:

$$\Phi_t + F \mid \nabla\Phi \mid = 0 \tag{5.12}$$

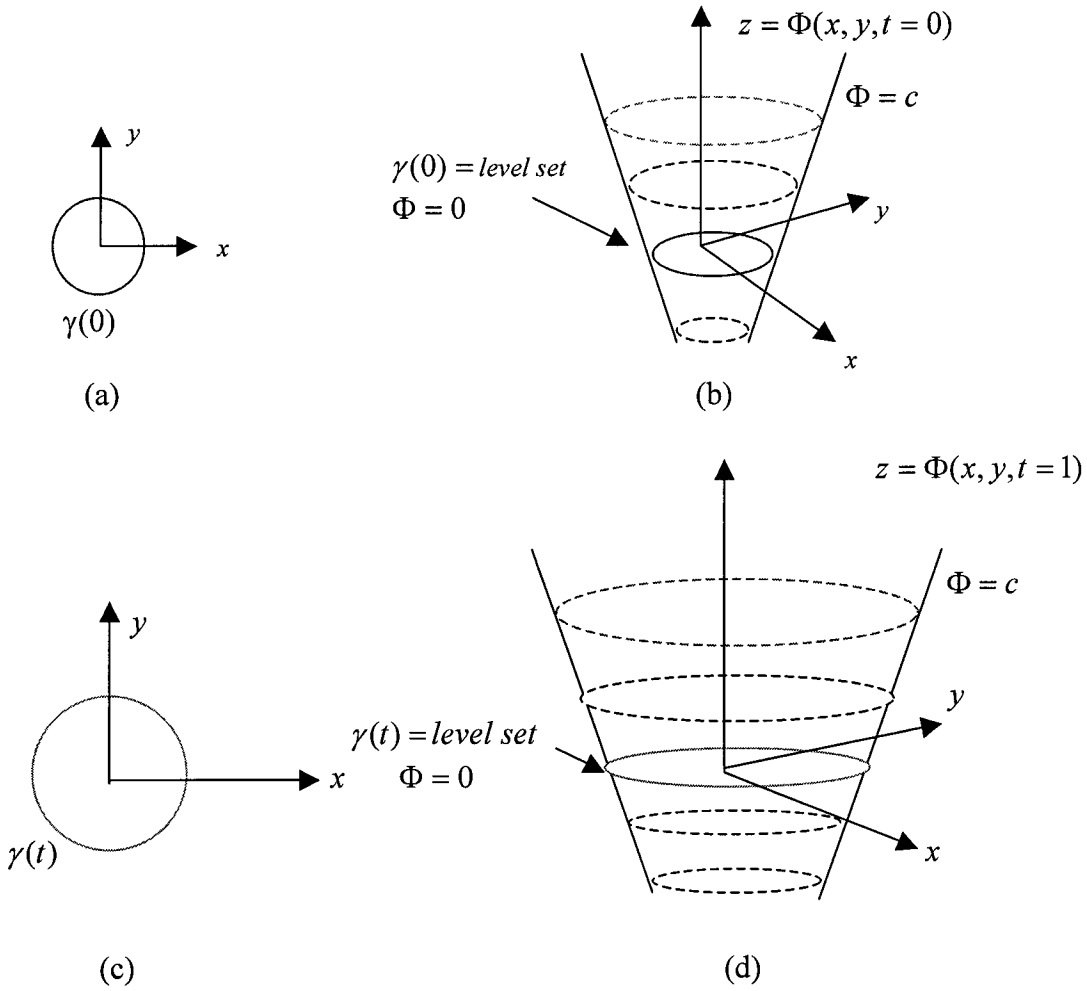


Figure 5.2 Level set formulation of equations of motion (Sethian, 1995)
(a) and (b) show the curve γ and the surface $z = \Phi(x, y)$ at *t=0.*
(c) and (d) show the curve γ and the surface $z = \Phi(x, y)$ at *t=1.*

In order to develop a numerical method to solve the Equation (5.12), a method based on the link between Hamilton-Jacobi equations and hyperbolic conservation laws is presented, with a given value of $\Phi(\boldsymbol{p}, t = 0)$. Formula (5.12) is a type of Hamilton-Jacobi equation. For certain of forms of the speed function *F,* the standard Hamilton-

Jacobi equation can be obtained, which can converge to correct solutions (Sethian 1999, Sethian and Strain 1992).

Based on the advantages of the Hamilton-Jacobi equation, the evolving function $\Phi(p,t)$ always remains a function as long as F is smooth. The propagating hypersurface $\gamma(t)$ may change topology, break, merge, and form sharp corners as the function Φ evolves.

The Eulerain Hamilton-Jacobi formulation concerns numerical approximation. Using a uniform mesh of spacing h, with grid nodes ij, and employing the standard notation that Φ_{ij}^n is the approximation to the solution $\Phi(ih, jh, n\Delta t)$, where Δt is the time step, Equation (5.12) can be written as:

$$\frac{\Phi_{ij}^{n+1} - \Phi_{ij}^n}{\Delta t} + (F)(\nabla_{ij}\Phi_{ij}^n) = 0 \tag{5.13}$$

Let $\nabla_{ij}\Phi_{ij}^n$ be the appropriate finite difference operator for the spatial derivative. The correct technique for approximating the spatial derivative in the above comes from using the appropriate entropy condition for propagating curves. In order to build a correct entropy satisfying approximation to the difference operator, the technology of hyperbolic conservation laws is needed. More details about entropy condition can be found in Sethian (1999) and Osher and Sethian (1988).

The speed function $F(K)$ is as follows:

$$F(K) = F_0 + F_1(K) \tag{5.14}$$

where K is the curvature of level set and shown as :

$$K = \nabla \frac{\nabla\Phi}{|\nabla\Phi|} = -\frac{\Phi_{xx}\Phi_y^2 - 2\Phi_x\Phi_y\Phi_{xy} + \Phi_{yy}\Phi_x^2}{(\Phi_x^2 + \Phi_y^2)^{3/2}}$$

In two space dimensions, based on the forward and backward difference approximations in Φ, Equation (5.13) can become:

$$\Phi_{ij}^{n+1} = \Phi_{ij}^{n} - F_0 \Delta t((\max(D_x^- \Phi_{ij}, 0))^2 + (\min(D_x^+ \Phi_{ij}, 0))^2$$

$$+ (\max(D_y^- \Phi_{ij}, 0))^2 + (\min(D_y^+ \Phi_{ij}, 0))^2)^{1/2} - \Delta t F_1 \mid \nabla \Phi \mid \quad (5.15)$$

where D_x^+ computes the new values at j using information at j and j+1,

$D_x^+ \Phi_{ij} = \Phi_{i(j+1)} - \Phi_{ij}$

D_x^- computes the new values at j using information at j and j-1;

$D_x^- \Phi_{ij} = \Phi_{ij} - \Phi_{i(j-1)}$

D_y^+ computes the new values at i using information at i and i+1;

$D_y^+ \Phi_{ij} = \Phi_{(i+1)j} - \Phi_{ij}$

D_y^- computes the new values at i using information at i and i-1;

$D_y^- \Phi_{ij} = \Phi_{ij} - \Phi_{(i-1)j}$

The advantages of the level set shape recovery using curve propagation are:

- The evolving function $\Phi(\boldsymbol{p},t)$ always remains a function as long as F is smooth.
- Because the entropy condition has been used in level set shape recovery based on Hamilton-Jacobi equations and hyperbolic conservation laws, the formation of the singularity can be overcome.
- The intrinsic geometric properties of the curve may be easily determined from the level function Φ because F_1 is related to curvature K. Topological changes in the evolving curve can be handled.
- The above level set approach can be used in high spatial dimensions.

5.3.2 Boundary values for level set shape recovery

Level set shape recovery method is used to find the boundary of an evolving curve. In order to extract object shapes from an image, the curve should be forced to surround the

object boundaries. When the evolving curve reaches the boundaries of the objects, all the points on the curve stop evolving anymore and the computation is ended.

In Equation (5.15), if the speed function can be related to the image, it can be a criterion to stop the curve propagation. In Equation (5.14), the speed F has two components. F_0 is the advection term and is independent of the moving curve's geometry. Since the second term is related to the curvature of curve, it is dependent to the curve geometry. In order to make the speed approach to zero in the areas of shape boundaries, a quantity k_I is multiplied with Equation (5.14). The speed function (5.14) can be as follows:

$$F = k_I(F_0 + F_1(K)) = k_I(-1.0 - 0.025K) \tag{5.16}$$

If there are boundaries in the image, the intensity will change sharp. As shown in Figure 5.3 (Gonzalez and Woods 1992), the first derivative corresponds to the local maximum and second derivative corresponds to the zero crossings. The Figure illustrates how intensity changes result in zero crossings of the second derivative which correspond to the boundaries in the image.

The second derivative alone is a high pass operator. Therefore the zero crossings would be correlated with the fine details of the image. A two dimensional Gaussian filter is used to smooth the high frequency detail before application of the second derivative. The two operations can be combined into one filter, known as the Laplacian of Gaussian (LOG), which can be calculated by Equation (5.17).

$$\nabla G^2(x,y) = [\frac{x^2 + y^2}{\sigma^2} - 2]e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{5.17}$$

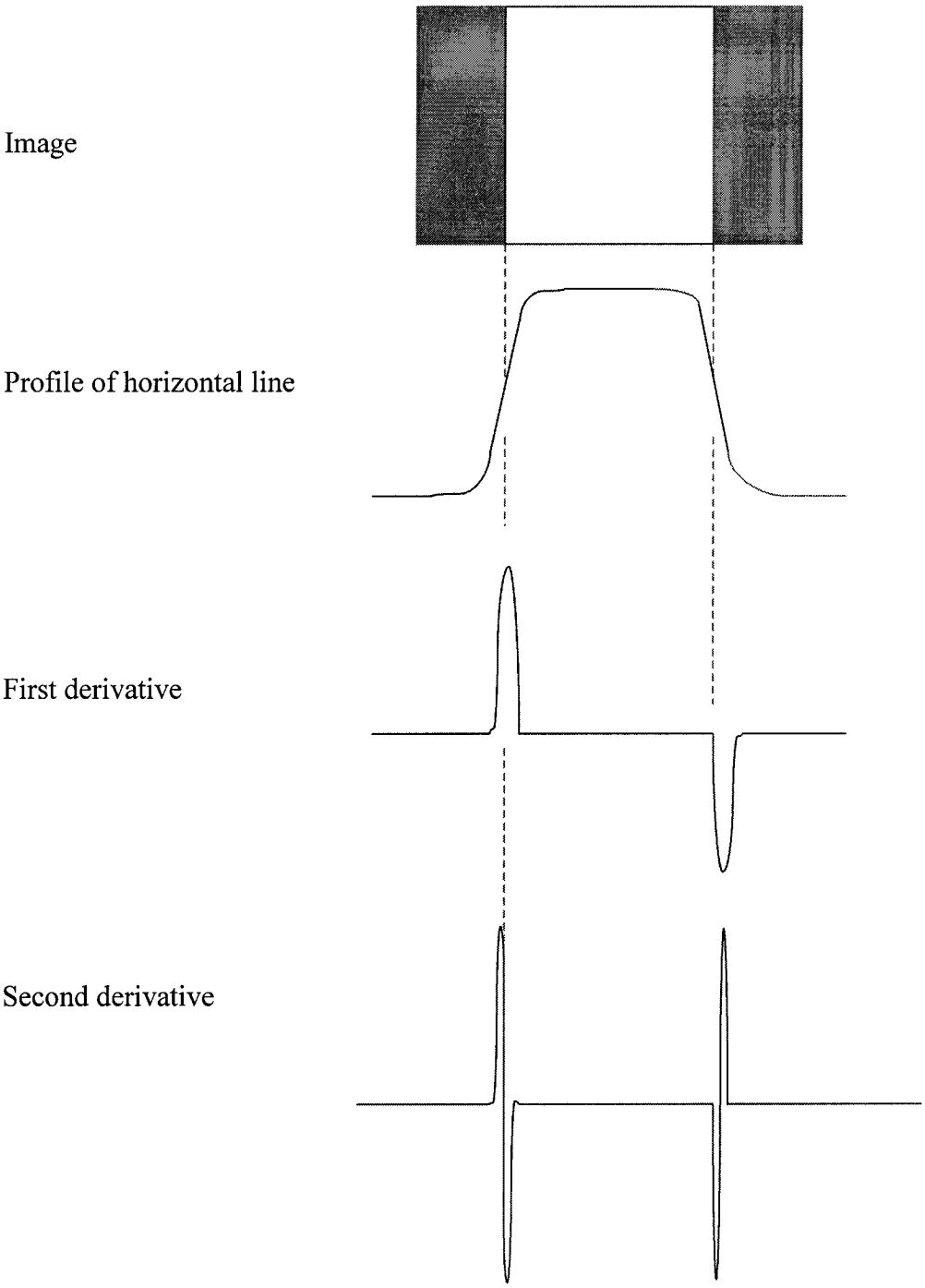


Figure 5.3 The relationship between the local maximum, zero crossings and grey level change (Gonzalez and Woods, 1992)

Figure 5.4 shows a cross-section of the LOG operator. From Equation (5.17), the width of the central lobe w is related to σ, which determines the degree of smoothing.

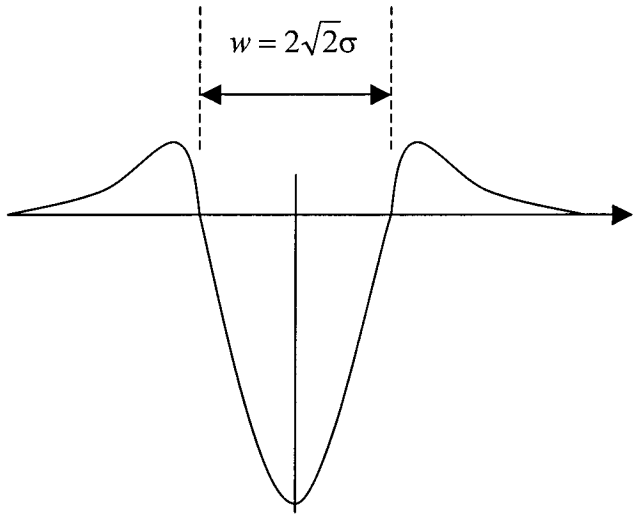


Figure 5.4 Cross section through LOG operator

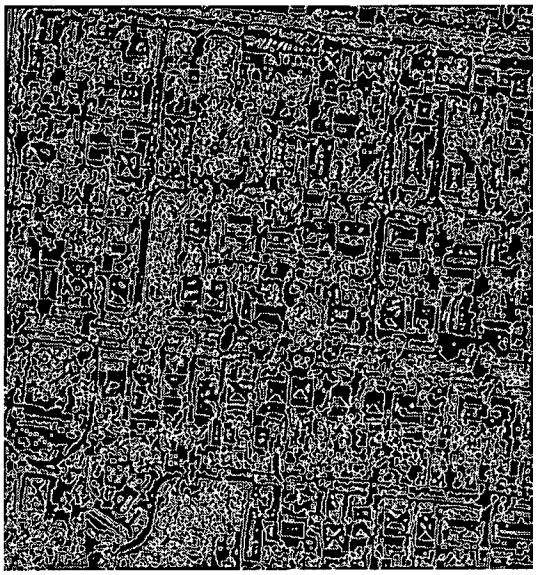
If k_I is defined as Equation (5.18), around the boundaries, $|LOG * I(x,y)|$ supplies the maximum value, then k_I approaches to zero. Based on Equation (5.16), the speed is closed to zero too, and the process of curve propagation will stop.

$$k_I(x,y) = \frac{1}{1+|LOG * I(x,y)|} \tag{5.18}$$

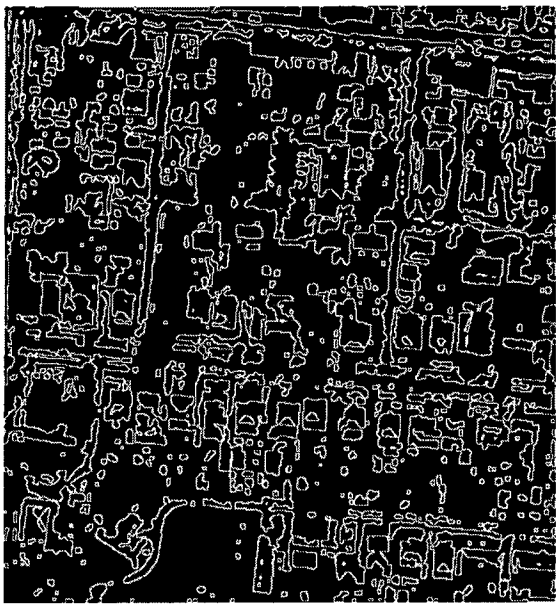
Figure 5.5 shows the convolution of the LOG with the image to find zero crossings which corresponding to the boundaries in the image. Figure 5.5(a) is the processed image. Figures 5.5(b)~ 5.5(d) show the convolution with the different sizes of LOG to produce zero crossings at different scales of resolution. In Figure 5.5, the processing window size is 9x9. When σ increases, the density of zero crossings becomes coarse.

More detailed examples of the level set shape recovery using the speed function based on the image edges can be found in Section 5.4

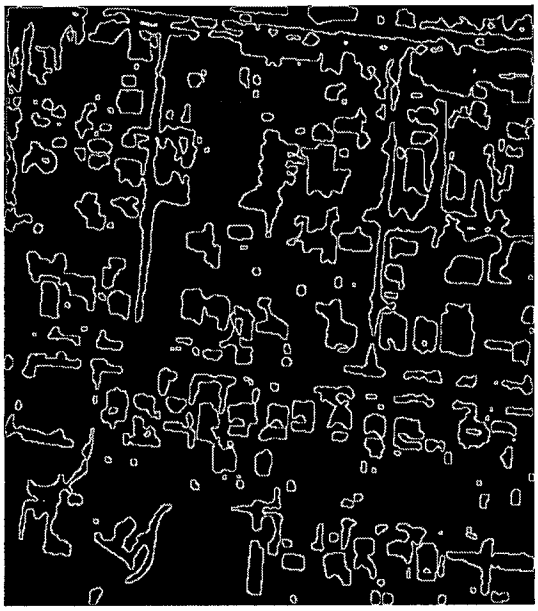
(a) The processed image



(b) Zero crossings with σ =1.25



(c) Zero crossings with σ =2



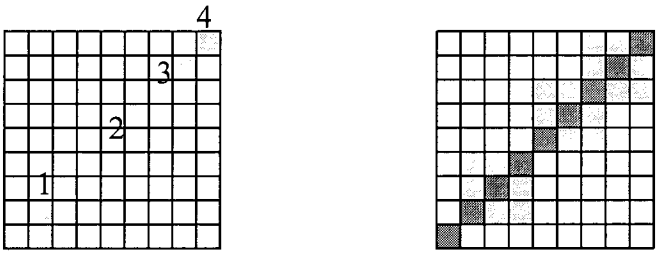
(d) Zero crossings with σ =3.25

Figure 5.5 LOG convolving with the image

5.3.3 Region growing algorithm for evolving curve recovery

As shown in Figure 5.2, after curve $\gamma(0)$ evolves iteratively, the new level set $\Phi = 0$ can be obtained as curve $\gamma(t)$. Based on curve $\gamma(t)$, new iterative evolving can start. The level set $\Phi = 0$ can be constructed from Equation (5.15) Φ_{ij}^{n+1}. Since the approximated construction of the zero level set can create a discontinuous curve away from the zero level set, reinitialization of the level set function during fixed time steps is needed to make sure the level set evenly evolves around the curve. For every reinitialization of the level set, it is necessary for the discontinuous curve to become a closed curve.

A seed based pixel-to-pixel region growing procedure has been used to following the curve. A seed can be any point on the tracking curve. In a similar method to that shown in Section 3.2.3, the method tries to link the points, which belong to the same curve but also separated from each other, into a closed curve. The example of this operation is given in Figure 5.6.



(a) Local curve points (b) After dilation

Figure 5.6 The break points linking sample using dilation

Figure 5.6(a) shows the points in the curve. When break point 1 appears, point 1 and the untracked points such as points 2,3,4 will extend to its 8 neighbours. The purpose of this operation is to link the break points. Figure 5.6(b) shows that point 1 can connect with point 2 and point 2 can connect with point 3 because of the dilation of points 1,2,3,4. The blue points construct the final curve. This region growing algorithm is efficient because only the points on the curve will be taken into account in the procedure, it is not necessary to process every pixel of the region.

5.4 Implementation of Module of Level Set Modelling

The theory of the level set modelling based image segmentation (LSMBIS) module has been introduced in the last section. In this section, the details and implementation of LSMBIS are described.

A brief flow diagram of LSMBIS is shown in Figure 5.7 and includes following main steps:

- Step 1, the processed image is convolved with the LOG to supply a zero crossings map using Equation (5.17).
- Step 2, based on the building interest areas from low-level image analysis and interpretation in Chapter 4, define an initial curve $\gamma(0)$. This initial curve is a square curve in the processed area. The minimum distance between a pixel to initial curve can be the initializing value for $\Phi(\boldsymbol{p}, t = 0)$.
- Step 3, based on Equation (5.14), the curvature from Φ_{ij}^{n} can be calculated. The speed F can be obtained using Equations (5.16) and (5.18). Finally, Φ_{ij}^{n+1} can be updated using Equation (5.15).
- Step 4, since Δt is small in Equation (5.15), after iterative evolving, a new level set $\Phi = 0$ can be constructed from Φ_{ij}^{n+1} using linear approximation. For a point *p(i, j)*, if

$$\max(\Phi_{i-1,j-1}, \Phi_{i-1,j}, \Phi_{i,j-1}, \Phi_{i,j}, \Phi_{i,j+1}, \Phi_{i+1,j}, \Phi_{i+1,j+1}) < 0$$

$$\min(\Phi_{i-1,j-1}, \Phi_{i-1,j}, \Phi_{i,j-1}, \Phi_{i,j}, \Phi_{i,j+1}, \Phi_{i+1,j}, \Phi_{i+1,j+1}) > 0$$

  then $p(i, j) \notin \gamma(t)$ is not included in new level set curve.
- Step 5, seed based region growing algorithm to construct the closed evolving curve.

- Step 6, based on reinitialized curve $\gamma(t_1)$, $\Phi(\boldsymbol{p}, t = t_1) = \pm dis\tan ce$ from $\boldsymbol{p}$ to $\gamma(t_1)$ can be calculated. The evolving processing will stop when the evolving curve is around the boundaries because the evolving speed is close to zero.

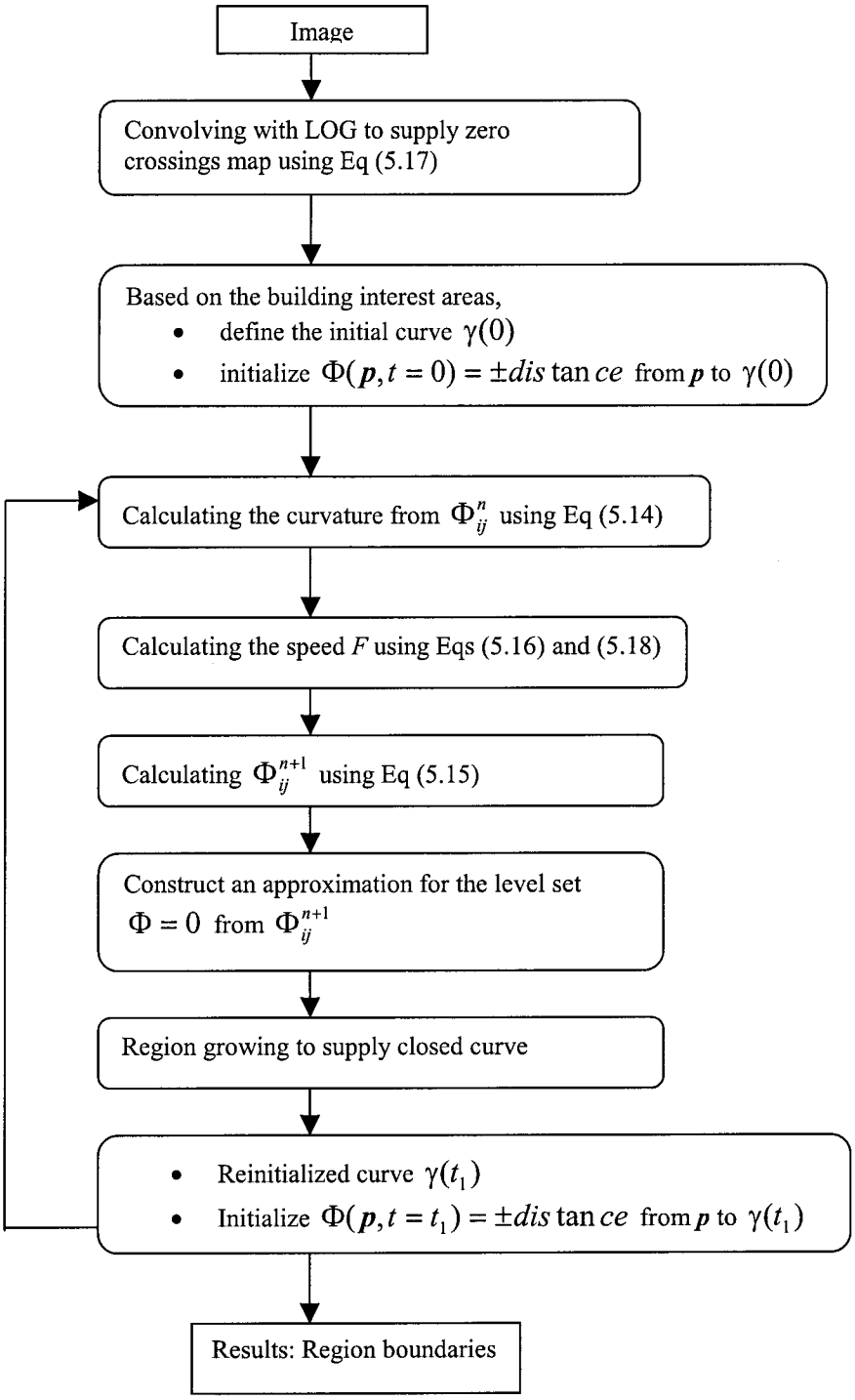


Figure 5.7 Brief diagram of LSMBIS

Figure 5.8 shows an example of building boundary extraction using LSMBIS module. The example demonstrates that when the evolving curve reaches the boundaries of the building, all the points on the curve stop evolving and the computation is ended.

Figure 5.8(a) shows the zero crossings in the building interest obtained by convolving the image with LOG operator. Zero crossings overlaid on images are illustrated in Figure 5.8(b). The initial curve is defined in Figure 5.8(c) as shown in blue. At Δt =0.00005, the curve will be reinitialized at 50 time steps. Figure 5.8(d) shows the result of curve reinitialization after 10 times. Figure 5.8(e) shows the final result of the evolving curve using level set modelling. The final extracted building boundary overlaid on the image is shown in Figure 5.8(f). The building boundary has been extracted correctly.

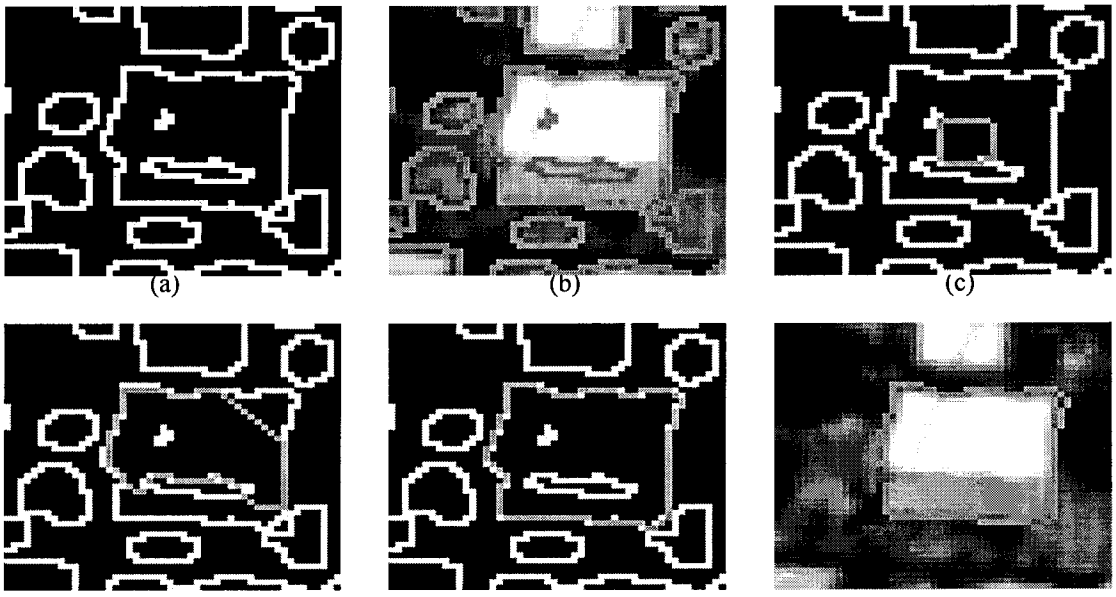


Figure 5.8 The example 1 of delineation the boundary of the building area

Figure 5.9 shows a second example of building boundary extraction using LSMBIS module. Figure 5.9(a) shows the zero crossings in the building interest area. Zero crossings overlaid on images are illustrated in Figure 5.9(b). The initial curve is defined as a square in building interest area. Figure 5.9(c)~5.9(e) show the results of curve evolving after 10,15,20 times reinitialization respectively. Figure 5.9(f) shows the final result of curve evolving using level set modelling. The final extracted building boundary

result of curve evolving using level set modelling. The final extracted building boundary overlaid on the image is shown in Figure 5.9(g). The building boundary has been correctly extracted.

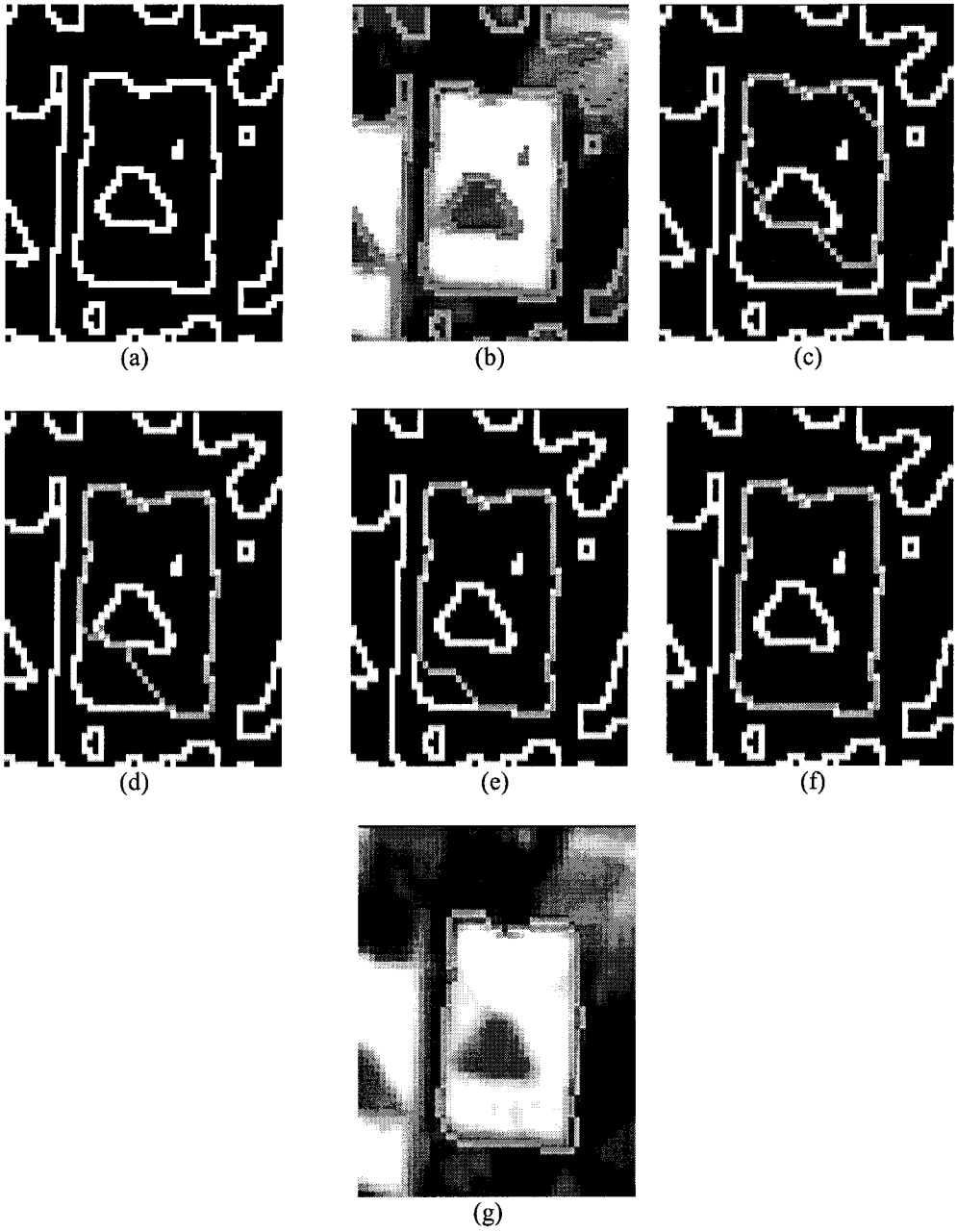


Figure 5.9 The example 2 of delineation the boundary of the building area

Figure 5.10 shows a third example of building boundary extraction using LSMBIS module. Figure 5.10(a) shows the zero crossings in the building interest area. Zero crossings overlaid on images are illustrated in Figure 5.10(b). Figure 5.10(c)~5.10(e)

show the results of curve evolving after 10,15,20 times reinitialization respectively. Figure 5.10(f) shows the final result of curve evolving after 25 iterations using level set modelling. The final extracted building boundary overlaid on the image is shown in Figure 5.10(g). The building boundary has been successfully extracted.

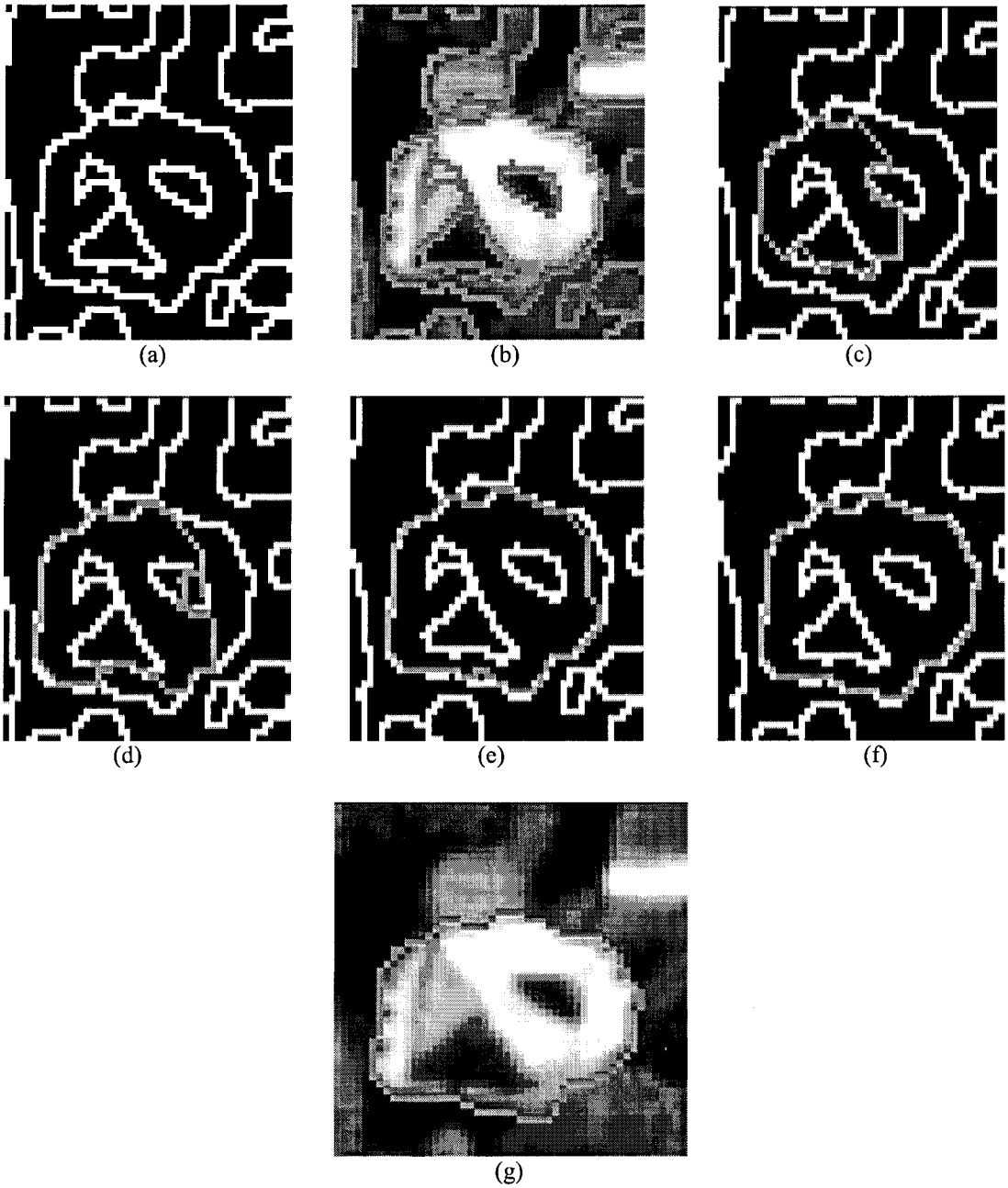


Figure 5.10 The example 3 of delineation the boundary of the building area

Figure 5.11 shows a fourth example of building boundary extraction using LSMBIS module. Figure 5.11(a) shows the zero crossings in the building interest area. Zero crossings overlaid on images are illustrated in Figure 5.11(b). Figure 5.11(c)~5.11(e) show the results of curve evolving after 2, 5, 10 times reinitialization respectively. Figure 5.11(f) shows the final result of curve evolving after 20 iterations using the level set modelling. The final extracted building boundary overlaid on the image is shown in Figure 5.11(g). The building boundary has been successfully extracted.

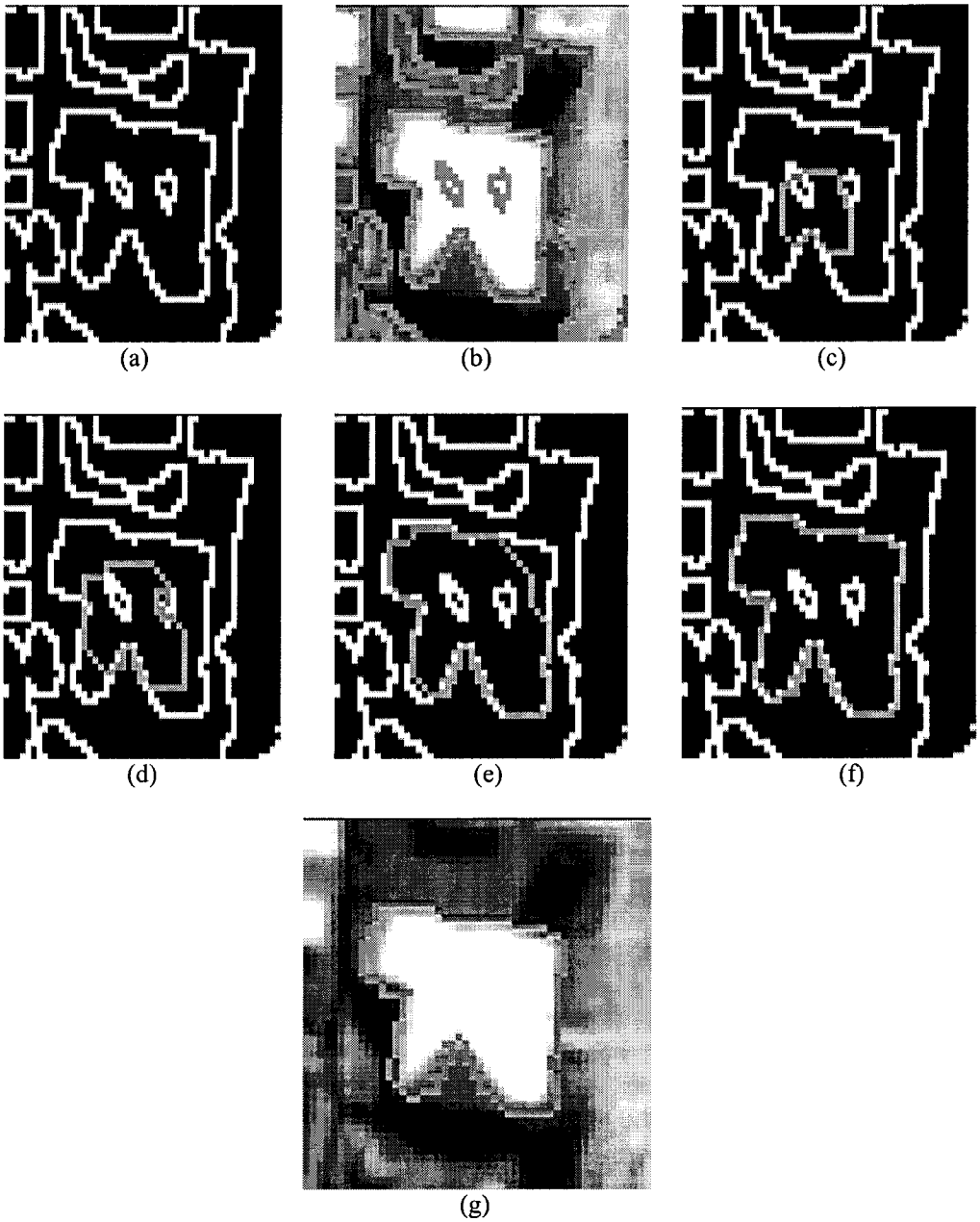
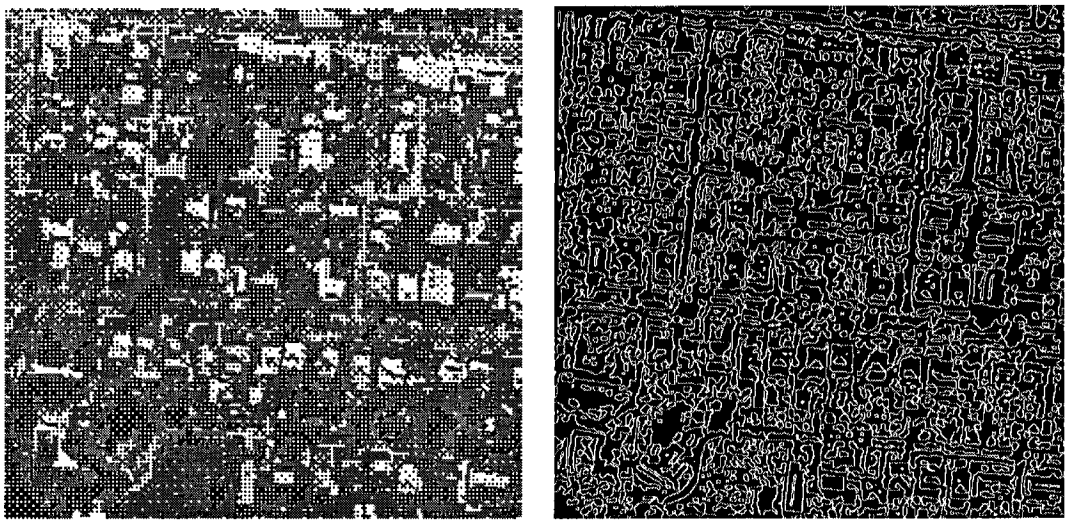


Figure 5.11 The example 4 of delineation the boundary of the building area

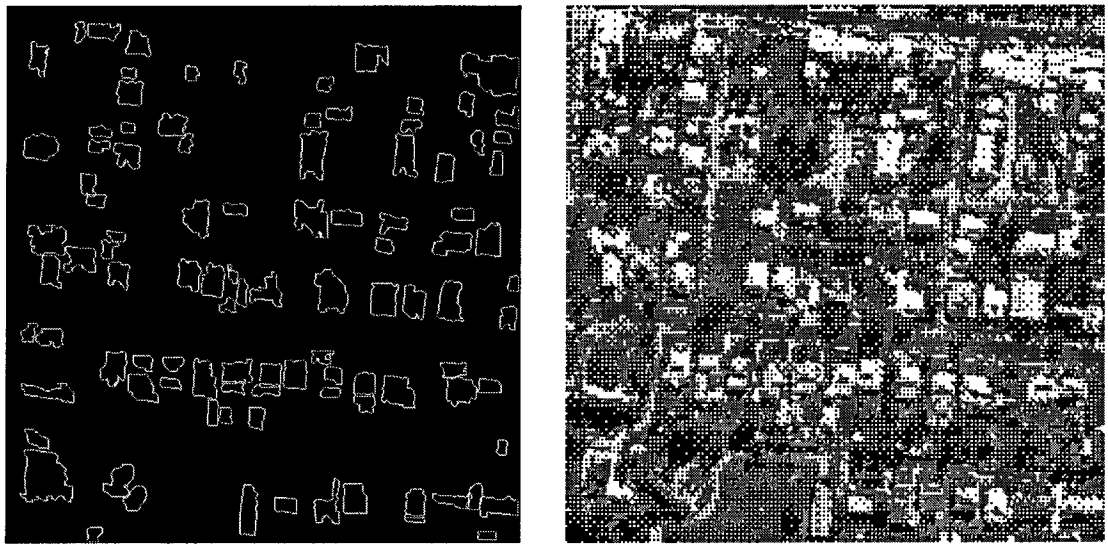
For the whole image which has been processed in Chapter 4 and shown in Figure 5.12(a), level set modelling based image segmentation (LSMBIS) module has been used to delineate the boundaries of the buildings. Based on the approximate building areas obtained from low level image analysis and interpretation, for every building interest area, shape modelling with level set method is implemented. Figure 5.12(b) shows the zero crossings map which has $\sigma=2$ and window size=$13x13$. Figure 5.13(a) is the final building boundaries extracted by LSMBIS. The level set results overlaid on the ortho image are shown in Figure 5.13(b).

The level set method has successfully delineated most of the building boundaries. The examples show LSMBIS is effective module for automatic building extraction system. However it could not supply the correct boundaries for areas where several building interest areas are mixed together, such as areas at the top of the image. Some areas are wrongly assigned as building areas in the image because the building interest areas from previous processing are not correct. The method which will evaluate the extracted building areas and delete the wrong areas will be introduced in next chapter.



(a) (b)

Figure 5.12 The processed image and corresponding zero crossings



(a) (b)

Figure 5.13 Results from LSMBIS module

CHAPTER 6

Multi-source Evidential Reasoning based Region Evaluation

6.1 Introduction

The preliminary data analysis in Chapter 4 illustrated that a single source of evidence does not provide a consistent means of interpretation of the image for extraction of buildings. Each data source used in this thesis has strengths and weaknesses. In order to evaluation the reliability of extracted building regions, an approach combining evidence from a number of data sources is need. Region evaluation with redundant data can help reduce imprecision while complementary data can provide a more complete description. A multi-source evidential reasoning based region evaluation (MEBRE) module has been developed for this task. The module uses three data sources: DSM derived in Chapter 3, the classified image in Chapter 4 and regions from level set modelling based image

segmentation (LSMBIS) as derived in Chapter 5. Evaluation at feature level has been implemented in the MEBRE module by the Dempster-Shafer method, which is a statistical-based data fusion classification algorithm and used when the data contributing to determination of the analysis of the images is subject to uncertainty. One of the advantages of this module is that it uses the spectral and spatial characteristics of the features. This approach effectively combines and counterbalances the multiple evidence from different data sources, in order to extract correct building regions instead of relying on scoring techniques. The architecture of MEBRE module for decision making is presented in Section 6.2. Data fusion using Dempster-Shafter theory is described in Section 6.3. Multi-source region evaluation is examined in Section 6.4.

6.2 The Architecture of MEBRE Module

Images are subject to poor resolution, noise, occlusions, shadows and other artifacts that cause errors in the processing. Therefore the processing of a single data source has uncertainties that affect the quantity and quality of information extraction. The use of multiple redundant data sources can overcome many of these shortcomings for decision making. As an example, as described in Chapter 5, region boundaries are extracted using level set modelling based image segmentation (LSMBIS) module. Initial values for LSMBIS are derived from low-level image analysis and interpretation. If these initial values are subject to uncertainties, the region boundary may not be correctly extracted. Therefore, the MEBRE module applies the Dempster-Shafer data fusion technique to combine the information from multiple sources of derived data, such as DSM, classified images and regions from level set method, to supply reliable building regions. Figure 6.1 illustrates the multi-source evidential reasoning based region evaluation (MEBRE) module. More detail explanation of Figure 6.1 will be described in Section 6.4, after the theory of Dempster-Shafer is introduced.

Data fusion techniques have the ability to interpret simultaneously information from multiple data sources of the same scene, in order to obtain correct conclusions. Since the data sources may highlight different characteristics, they are also partly complementary.

None of the data sources provides completely decisive and reliable information. In addition, the information is often imprecise and uncertain. The application of a data fusion technique aims to improve the quality of a decision by increasing the number of date sources while decreasing its imprecision and uncertainty, by making use of redundant and complementary data. A further advantage of Dempster-Shafer evidence theory is that it provides estimations of imprecision and uncertainty of the information derived from different sources (Klein 1999 and Hegarat-Mascle 1997). In order to easily refer the names of multiple data sets, the short terms of the names have been defined as:

Clustered image: classification image described in Chapter 4

Regions of LevelSet: building regions extracted by level set modelling based image segmentation in Chapter 5

DSM: digital surface model described in Chapter 3.

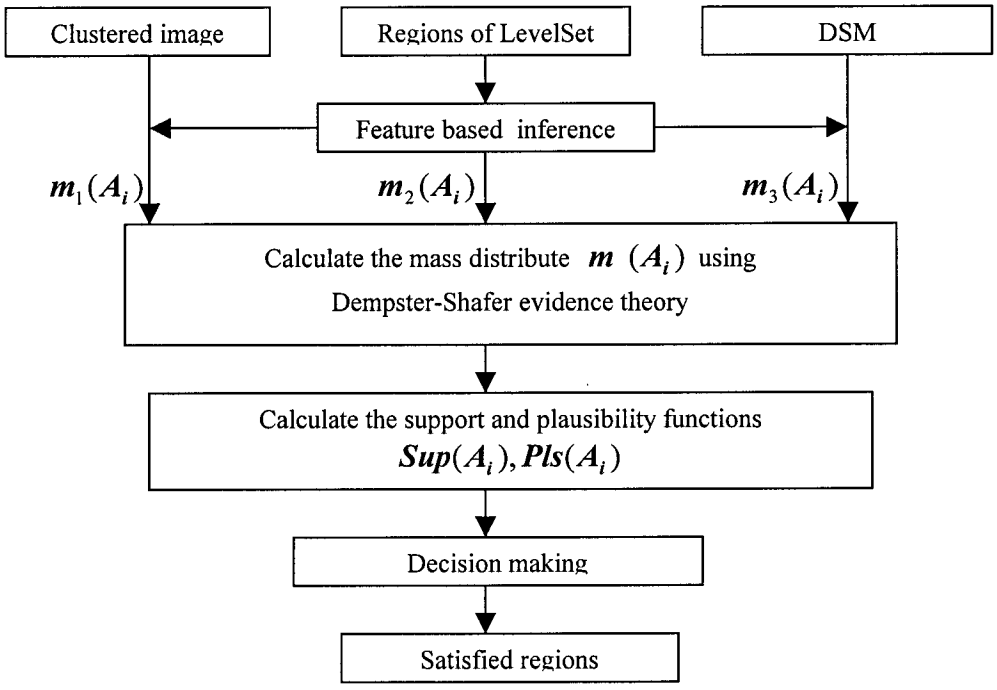


Figure 6.1 Multi-source evidential reasoning based region evaluation (MEBRE) module

6.3 Data Fusion Using Dempster-Shafer Theory

Assume a set of n propositions making up the hypothesis space and is denoted by Θ. 2^{Θ} is the subsets of Θ. Based on the information from the data source, a probability mass $\boldsymbol{m}$ can be assigned to any proposition or union of propositions. For $\forall A \in 2^{\Theta}$, $\boldsymbol{m}$ is defined for every element A and the mass value $\boldsymbol{m}(A)$ is in the interval [0,1]. The following mass equations can be obtained:

$$\boldsymbol{m}(\varnothing)=0$$
$$\boldsymbol{m}(\Theta)=\sum_{A\subset 2^{\Theta}}\boldsymbol{m}(A)=1 \tag{6.1}$$

where $\varnothing$ is the empty set.

In image classification, Θ is the set of hypotheses about pixel classes. The Dempster-Shafer theory allows the consideration of any subset of Θ. Applied to image classification problems, it means that not only single classes but also any union of classes can be represented. Including all possible unions and excluding the null set, the number of classes is equal to 2^n-1. For example, if n=3, $2^3-1=7$ classes given by C1, C2, C3, C1 $\cup$ C2, C1 $\cup$ C3, C2 $\cup$ C3, and C1 $\cup$ C2 $\cup$ C3 (Klein 1999, Hegarat-Mascle 1997 and Shafer 1976).

The Dempster-Shafer theory provides a representation of both imprecision and uncertainty through the definition of two parameters: support (*Sup*) and plausibility (*Pls*), which are obtained from the probability mass function $\boldsymbol{m}$. *Support* for a given proposition means that all masses assigned directly by the data sources are summed. *Plausibility* for a given proposition means all masses not assigned to its negation are summed. For $\forall A \in 2^{\Theta}$ and $\forall B \in 2^{\Theta}$, two parameters are defined respectively as follows:

$$Sup(A)=\sum_{B\subseteq A}\boldsymbol{m}(B)$$
$$Pls(A)=\sum_{B\cap A\neq\varnothing}\boldsymbol{m}(B) \tag{6.2}$$

An uncertainty interval is defined by [*Sup*(*A*), *Pls*(*A*)] where

$$Sup(A) \leq Pls(A)$$

$$Pls(A)=1-Sup(\overline{A}), \quad A \cup \overline{A} = \Theta, \; A \cap \overline{A} = \varnothing \tag{6.3}$$

$\overline{A}$ is the complementary hypothesis of A. $Sup(\overline{A})$ is called the dubiety and represents the degree to which the evidence impugns a proposition. It supports the negation of the proposition.

The *support* value of hypothesis A may be interpreted as the minimum uncertainty value about A. Its *plausibility* may be interpreted as the maximum uncertainty value of A. The uncertainty interval gives a measurement of the imprecision about the uncertainty values.

The Dempster-Shafer method allows compatible propositions to combine the probability masses from several data sources and obtain a single value for the probability of the intersection (union) of the propositions. Assume there are A_i object types in data source 1 and B_j object types in data source 2, the total probability mass committed to a subset A is

$$m(A)=\frac{\sum_{\substack{i,j \\ A_i \cap B_j = A}} m_1(A_i)\, m_2(B_j)}{1-K}, \quad K \neq 1$$

$$K = \sum_{\substack{i,j \\ A_i \cap B_j = \varnothing}} m_1(A_i)\, m_2(B_j) \tag{6.4}$$

The hypotheses about single classes and unions of classes are called simple hypotheses and compound hypotheses respectively. The final decision of class types can be made by a combination of the probability masses from multiple data sets based on the decision rule. There are different decision rules. When the probability mass of simple hypotheses is not null, a decision rule such as the maximum of support over simple hypotheses can be used. The formula is as following:

$$\max(Sup(A))$$
$$Sup(A) \geq Sup(\overline{A}) \tag{6.5}$$

Table 6.1 shows the computed probability mass, *plausibility* and *support* values for each simple and compound hypothesis of two data sources (Hegarat-Mascle et al 1997). Assume three classification types C1, C2, C3 in the images. Since there is no general answer to the problem of the probability mass definition, the probability masses should be defined based on the application. If two classes Ci and Cj cannot be distinguished by a data source, it makes sense to give a non null probability mass to their union $Ci \cup Cj$. The probability masses for classes and their union for two data sources have been showed in Table 6.1.

In this table, for data set 1, the probability masses for class C1, C2, C3 and the union of their classes are defined as:

$$\boldsymbol{m}_1(C1 \cup C2) = \boldsymbol{m}_1(C1 \cup C3) = \boldsymbol{m}_1(\Theta) = 0$$
$$\boldsymbol{m}_1(C2) = \boldsymbol{m}_1(C3) = \boldsymbol{m}_1(C2 \cup C3) = t$$

t is a constant probability and normalizing the probability masses for all classes in data set 1, the probability mass for C1 class can be:

$$\boldsymbol{m}_1(C1) = 1 - \boldsymbol{m}_1(C1 \cup C2) - \boldsymbol{m}_1(C1 \cup C3) - \boldsymbol{m}_1(C2 \cup C3) - \boldsymbol{m}_1(C2) - \boldsymbol{m}_1(C3) - \boldsymbol{m}_1(\Theta)$$
$$= 1\text{-}3t$$

Similarly, the probability masses can be assigned to all the classes in data set 2 as shown in Table 6.1. The combination of probability masses of two data sets is obtained from Equation (6.4). The above definition of probability masses has been used by Hegarat-Mascle (1997) to implement an unsupervised classification.

For simple class C1, according to Equation (6.2) and the values in the table, the *support* and *plausibility* functions can be calculated as follows:

$$Sup(C1) = \boldsymbol{m}(C1)$$
$$Pls(C1) = \boldsymbol{m}(C1) + \boldsymbol{m}(C1 \cup C2) + \boldsymbol{m}(C1 \cup C3) + \boldsymbol{m}(C1 \cup C2 \cup C3) = \boldsymbol{m}(C1)$$

Table 6.1 : Dempster-Shafer calculation from two data sources
$1-k=2(t+u-4ut)$

A	m_1(A)	m_2(A)	m(A)	Sup(A)	Sup($\overline{A}$)
C1	$1-3t$	u	$\frac{2u(1-3t)}{1-k}$	$\frac{2u(1-3t)}{1-k}$	$\frac{2t(1-u)}{1-k}$
C2	t	$1-3u$	$\frac{2t(1-3u)}{1-k}$	$\frac{2t(1-3u)}{1-k}$	$\frac{2u(1-t)}{1-k}$
C3	t	u	$\frac{4tu}{1-k}$	$\frac{4tu}{1-k}$	$\frac{2(t+u-6tu)}{1-k}$
C1 $\cup$ C2	0	0	0	$\frac{2(t+u-6tu)}{1-k}$	$\frac{4tu}{1-k}$
C1 $\cup$ C3	0	u	0	$\frac{2u(1-t)}{1-k}$	$\frac{2t(1-3u)}{1-k}$
C2 $\cup$ C3	t	0	0	$\frac{2t(1-u)}{1-k}$	$\frac{2u(1-3t)}{1-k}$
Θ	0	0	0	1	0

For compound hypothesis class $C1 \cup C2$, according to Equation (6.2) and the values in the table, the *support* and *plausibility* functions can be calculated as follows:

$$Sup(C1 \cup C2) = \boldsymbol{m}(C1) + \boldsymbol{m}(C2) + \boldsymbol{m}(C1 \cup C2) = \boldsymbol{m}(C1) + \boldsymbol{m}(C2)$$

$$Pls(C1 \cup C2) = \boldsymbol{m}(C1) + \boldsymbol{m}(C2) + \boldsymbol{m}(C1 \cup C2) + \boldsymbol{m}(C1 \cup C3)$$
$$+ \boldsymbol{m}(C2 \cup C3) + \boldsymbol{m}(C1 \cup C2 \cup C3)$$
$$= \boldsymbol{m}(\mathrm{C1}) + \boldsymbol{m}(\mathrm{C2})$$

By the above inference, *Sup*(C1) and *Pls*(C1), *Sup*($C1 \cup C2$) and *Pls*($C1 \cup C2$) have the same values respectively. The *support* and *plausibility* functions for other classes can be obtained using the same method. *Supports* for the negation of the proposition can be calculated using Equation (6.3).

In table 6.1, since all the compound hypothese $\boldsymbol{m}(C1 \cup C2), \boldsymbol{m}(C1 \cup C3)$ and $\boldsymbol{m}(C2 \cup C3)$ have null probability masses after combination, *Sup* and *Pls* are equal according to Equation (6.2) and the uncertainty interval as described in Equation (6.3) reduces to zero. This means that there is no longer imprecision after the two data sets are combined.

6.4 Multi-sources Region Evaluation

In this section, the theory fundamentals of multi-sources region evaluation in MEBRE module are introduced, including the method of choosing the initial probabilities, and their combination for evaluating building areas.

6.4.1 Fundamentals of multi-sources region evaluation

If there are more than 2 data sources, m_n is the basic probability mass provided by source n ($1 \le n \le p, p \ge 3$). Using Dempster-Shafer evidential theory, the combination of all the data sources is defined as follows:

$$m(A) = \frac{\sum_{A_1 \cap A_2 \ldots\ldots \cap A_p = A} \prod_{1 \le n \le p} m_n(A_i)}{1 - K}$$

$$K = \sum_{A_1 \cap A_2 \ldots\ldots \cap A_p = \varnothing} \prod_{1 \le n \le p} m_n(A_i) \qquad (6.6)$$

In Figure 6.1, $m_n(A_i)$ is the probability mass for class A in data set n. Regions of LevelSet are spatial features from previous processing. Corresponding feature areas in DSM and clustered image can be extracted respectively. For each feature area in the different data sources, probability masses $m_n(A_i)$ can be defined. Based on Dempster-Shafer theory, combined probability masses $m(A_i)$ from multiple data sources are calculated using Equation (6.6). Using the *support* and *plausibility* functions in Equation (6.5), the MEBRE module can then be used to predict the building regions.

As mentioned earlier, Dempster-Shafer theory has been used for unsupervised pixel level classification (Hegarat-Mascle et al 1997). Because only non spatial features were considered, classification errors occur at the boundaries of features. The evaluation procedure by Dempster-Shafer evidential reasoning described in this study is based on spatial features. This means that the determination of probability masses and their combination are based on features. For each region of LevelSet, there are corresponding

areas obtained from the DSM and clustered image. The total possible evidence combinations are shown in Figure 6.2. The clustered image, regions of LevelSet and

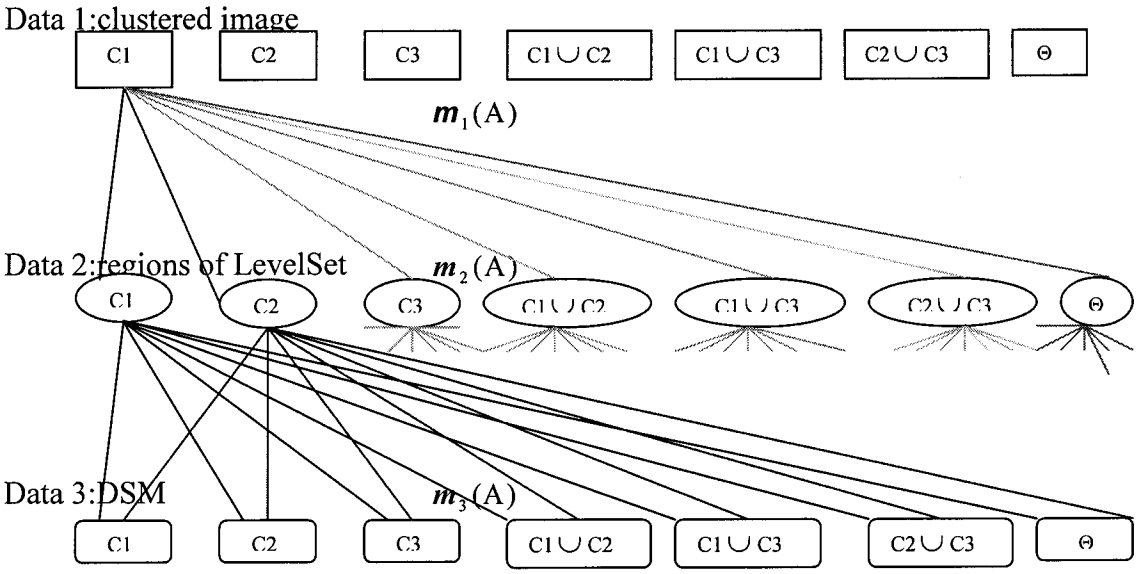


Figure 6.2 Calculation probability masses based on three data sets

DSM are assigned as data sets 1,2, and 3 respectively. Classes C1, C2 and C3 represent trees, buildings and ground respectively. For determining probability mass $\boldsymbol{m}$(C1), Class C1 in data set 1 is combined with all the classes in data sets 2 and 3 to find propositions which belong to class C1. Although there are 49 possible combinations for calculating probability mass $\boldsymbol{m}$(C1), only the combinations which belong to class C1 can be added to probability mass $\boldsymbol{m}$(C1). Intersections of other classes, such as C1 ∪ C2 in data set 1 with C1 in data set 2 and 3, belong to class C1, they are also added to probability mass $\boldsymbol{m}$(C1). Similarly, the probability masses for other classes can be calculated as shown in Figure 6.2.

Table 6.2 shows the computed probability mass, *plausibility* and *support* values for each simple and compound hypothesis based on three data sources. For data set 1, the clustered image, the probabilities for all the simple and compound classes are represented by probability t as an example. Similarly, for data sets regions derived from LevelSet and DSM, the probabilities for the classes are represented by u and s respectively.

Table 6.2 : Dempster-Shafer calculation from three data sources

$1-k = 4(ut + st + us - 9uts)$

A	m_1(A) clustered image	m_2(A) regions LevelSet	m_3(A) DSM	m(A)	Sup(A)	Sup($\overline{A}$)
C1 trees	$1-3t$	u	s	$\frac{4us(1-3t)}{1-k}$	$\frac{4us(1-3t)}{1-k}$	$\frac{4t(s+u-6su)}{1-k}$
C2 buildings	t	$1-3u$	s	$\frac{4st(1-3u)}{1-k}$	$\frac{4st(1-3u)}{1-k}$	$\frac{4u(s+t-6st)}{1-k}$
C3 ground	t	u	$1-3s$	$\frac{4ut(1-3s)}{1-k}$	$\frac{4ut(1-3s)}{1-k}$	$\frac{4s(u+t-6ut)}{1-k}$
C1 ∪ C2	0	0	s	0	$\frac{4s(u+t-6ut)}{1-k}$	$\frac{4ut(1-3s)}{1-k}$
C1 ∪ C3	0	u	0	0	$\frac{4u(s+t-6st)}{1-k}$	$\frac{4st(1-3u)}{1-k}$
C2 ∪ C3	t	0	0	0	$\frac{4t(s+u-6su)}{1-k}$	$\frac{4us(1-3t)}{1-k}$
Θ	0	0	0	0	1	0

6.4.2 Initial probability mass definition for region evaluation

The definition of probability functions for region evaluation remains a largely unsolved problem. In image processing, the definition of the initial probability masses can be obtained at three different levels. At the most abstract or highest level, information representation is derived in a way similar to method in artificial intelligence, where probability masses are assigned to propositions, often provided by experts (Gordon and Shortliffe, 1985). At the middle level, the definition of probability masses is derived from attributes, and may involve simple geometrical models (Van Cleynenbreugel et al., 1991). This definition is suitable to model-based pattern recognition, but it is difficult to use in image fusion for the classification of complex structures such as buildings in urban areas for which no model exists. At the pixel level, probability masses are obtained from statistical pattern recognition. The most widely used approach assigns probability masses based on simple hypotheses (Rasoulian et al., 1990). Because there are no probability masses for compound hypothesis, this limits the power of Dempster-Shafer evidential method.

For the multi-source evidential reasoning based region evaluation (MEBRE) module, the probability masses are assigned based on a reasoning approach where knowledge about the information provided by each image is used. This method is more reliable and is able to take into account a larger variety of situations. For data set 1, the clustered image, $m_1(A)$ has no null values to C1, C2, C3 and C2 $\cup$ C3 as shown in Table 6.2. Since C2 for buildings and C3 for the ground may have the same texture in the image, there are ambiguities between these two classes in clustered image. In Table 6.2, C2, C3 and C2 $\cup$ C3 are assigned the same probability, t. Null probability masses are assigned to the other compound hypotheses. Since all the masses sum to 1, the probability mass for C1 is $1-3t$. For each extracted region of LevelSet, the numbers of pixels representing trees, ground, grasses and building can be calculated respectively. The probability t can be defined based on pixel assigned to each building region:

total=NumofTree+NumofGround+NumofGrass+NumofBuilding

probBuilding=NumofBuilding/total

Since $t \in [0, \frac{1}{3}]$ and varies from 0 to 1/3, the calculated probability mass of pixels being buildings from the above formula can be normalized in the t range.

$$t = \frac{probBuilding}{3} \tag{6.7}$$

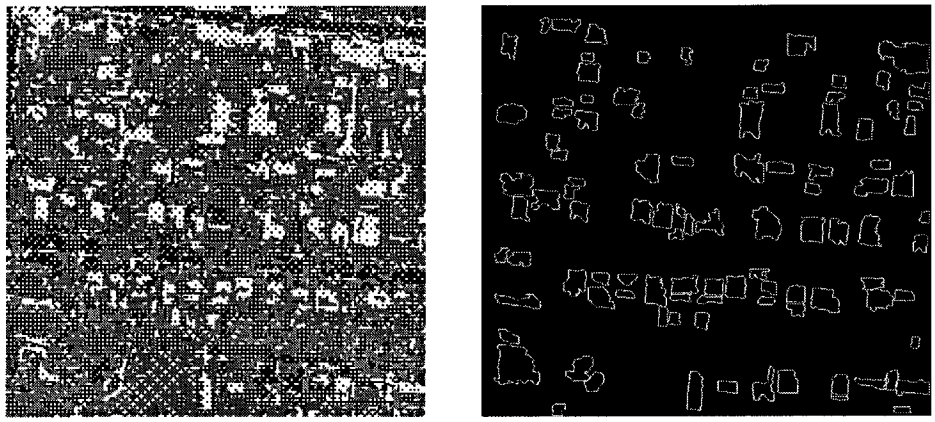
For data set 2, regions of derived by LevelSet, since the extracted building regions are from the processing of low-level image analysis and interpretation (LLIAI) and level set modelling based image segmentation (LSMBIS), they are more reliable and are assigned higher probabilities. As shown in Table 6.2, buildings in class C2 are assigned a probability of $1-3u$ and other non-null classes are assigned a probability of u. If the other classes are assigned lower probabilities, the class of building will have higher probability since the sum of probabilities is 1.

For data set 3, DSM image, since buildings and trees are above ground areas in the DSM, these two classes areas are easily confused. As shown in the Table 6.2, class C1 representing trees, C2 of buildings and C1 $\cup$ C2 are assigned the same probabilities, s.

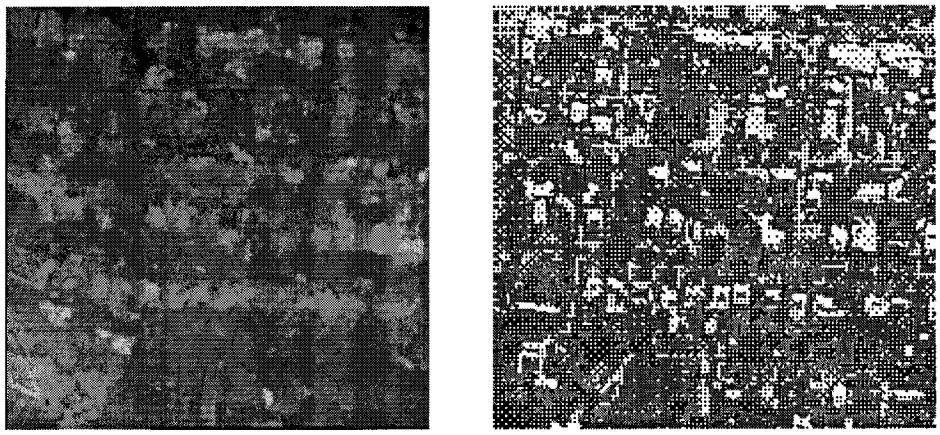
The probability mass of class C3, the ground is *1-3s*. The calculation of *s* is based on the mean of the DSM values. For example, if the mean of elevations for a building area is high, the probability *s* assigned to that building area in the DSM is assigned as higher than other areas.

6.4.3 Experiments with region evaluation module

The application of the multi-source evidential reasoning based region evaluation (MEBRE) module has been tested on one area. The data sets and original left image are shown in Figure 6.3.



a) Data1: Clustered image b) Data2: Regions of LevelSet



c) Data3: DSM d) Left image

Figure 6.3 Three data sets in MEBRE module and left image

The evaluation using two data sources

Firstly, two data sets, clustered image and regions of LevelSet are input to the MEBRE module. Based on the decision rule Equation(6.5), most building areas can be reliably

extracted, as shown in Figure 6.4. The building regions derived from level set modelling overlaid on ortho image are shown in Figure 6.5. Comparing Figure 6.4 with Figure 6.5, it is noted that some correct areas have also been deleted although some incorrect building areas have been detected and eliminated.

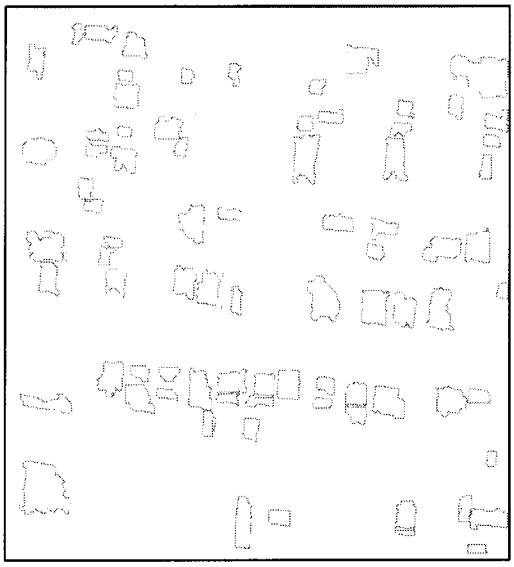


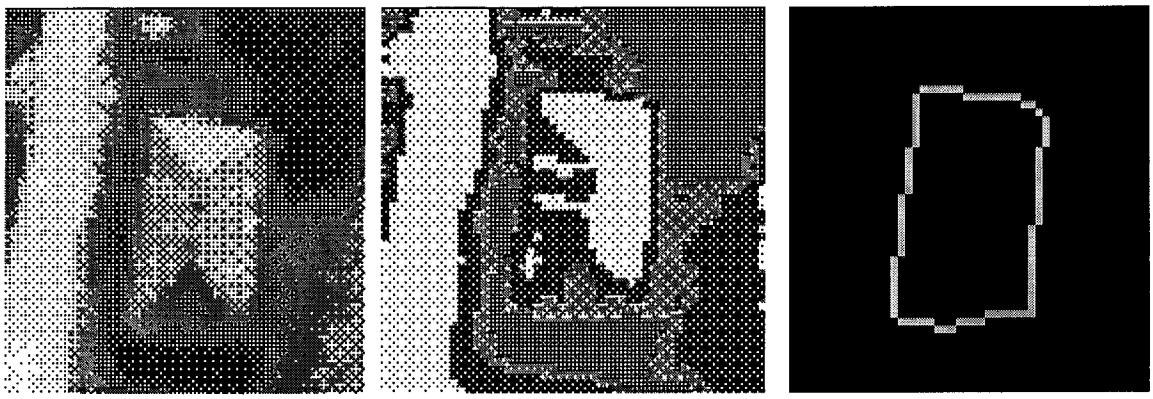
Figure 6.4 Result from Dempster-Shafer calculation based on two data sources

Figure 6.5 Regions of LevelSet overlaid on ortho image

Figure 6.6 Data fusion using two data sets clustered image and regions of LevelSet

In Figure 6.6, the green regions are false building areas and correctly deleted by fusion of the two data sets. Red regions are correct building regions that has been wrongly deleted. For example, the building No.35 indicated by arrow is a correct building area and deleted from building regions by data fusion.

Figure 6.7 shows the small areas of building No.35 in ortho image (a), clustered image (b) and region of LevelSet (c). In the clustered image, red, green, blue and yellow represent



(a) (b) (c)

Figure 6.7 Small region areas for building No.35

trees, grass, ground and building respectively. Table 6.3 shows the numbers of pixels for 4 classes in Figure 6.7 (b) corresponding to the building region in Figure 6.7 (c). Based on the Equation (6.7), the normalized probability of building t is 0.16.

Table 6.3 Pixel numbers in clustered image

Classes / Number	Tree	Grass	Ground	Building
	17	90	194	276

Table 6.4 shows computed probability masses, *plausibility* and *support* values for each simple and compound hypothesis for the model based on the two data sources, ie clustered image and building regions of LevelSet. In the case where the probability masses of simple hypotheses are not null, a decision rule based on the maximum value of the *support* over all hypotheses will always favour compound hypotheses. For example,

in Table 6.4, compound hypotheses C1 $\cup$ C2 has *Sup(A)=0.78* and *Sup($\overline{A}$)=0.22.* It satisfies the decision rule. In order to avoid this bias, a decision rule involving only the simple hypotheses is used. Based on Equation (6.5), there is no *Sup(A)* which is the max and $Sup(A) \geq Sup(\overline{A})$. In this analysis, building No. 35 is evaluated as non-building class, which is the reason for the incorrect deletion of building No. 35 from two data sets model.

Table 6.4 Building No.35 evaluation using two datasets: clustered image and regions of LevelSet

t=0.16, u=0.15, 1-k=0.428

A	m_1 (A) clustered img	m_2 (A) regions LS	m (A)	Sup(A)	Sup($\overline{A}$)
C1	0.52	0.15	0.36	0.36	0.64
C2	0.16	0.55	0.41	0.41	0.59
C3	0.16	0.15	0.22	0.22	0.78
C1 $\cup$ C2	0	0	0	0.78	0.22
C1 $\cup$ C3	0	0.15	0	0.59	0.41
C2 $\cup$ C3	0.16	0	0	0.64	0.36
Θ	0	0	0	1	0

The evaluation using three data sources

Based on the decision rule given in Equation (6.5), the three data sources, the clustered image, regions of LevelSet and DSM are input to MEBRE module and used by the Dempster Shafer algorithm to produce building areas shown in Figure 6.8. The regions of LevelSet overlaid on ortho image are shown in Figure 6.9. Comparing Figure 6.8 with Figure 6.9, it is clear that using three data sets will result in a much better extraction of buildings than the fusion of two data sets.

In this case, five incorrect building areas have been detected and deleted in the final result, which are the green areas shown in Figure 6.10. The correct building areas shown in Figure 6.3 (b) are unchanged after data fusion in Figure 6.10. Only one red area which is incorrectly shown as a building has not been detected, because of its higher elevation values in DSM information layer. This area contains ground and trees. The ground makes

the area being classed as a building and the height of tree indicated a building area, so the Dempster-Shafer processing could not detected the error.

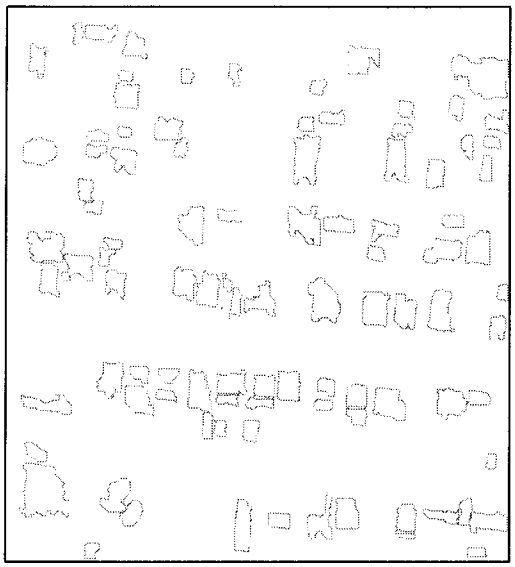


Figure 6.8 Result from Dempster-Shafer calculation based on three data sources

Figure 6.9 Regions of LevelSet overlaid on ortho image

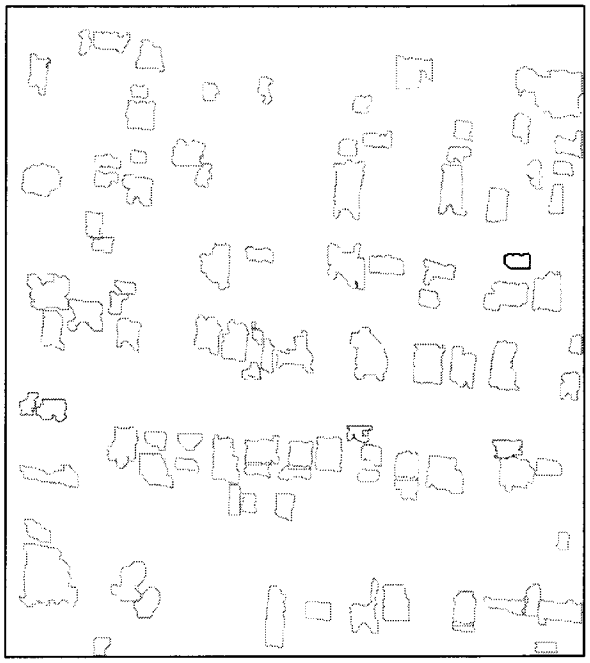


Figure 6.10 Data fusion using three data sets
Clustered image, regions of LevelSet and DSM

For building No. 35, Table 6.5 shows computed probability masses, *plausibility* and *support* values for each simple and compound hypothesis of three data sources. Based on the decision rule Equation (6.5), it is clear why building No. 35 has been evaluated as buildings from three data sets.

Table 6.5 Building No.35 evaluation using three datasets: clustered image, regions of LevelSet and DSM (definition 1 of probabilities)

t=0.16, u=0.15, s=0.33, 1-k=0.22

A	m_1(A) clustered img	m_2(A) regions LS	m_3(A) DSM	m(A)	Sup(A)	Sup($\overline{A}$)
C1 trees	0.52	0.15	0.33	0.47	0.47	0.53
C2 buildings	0.16	0.55	0.34	0.53	0.53	0.47
C3 ground	0.16	0.15	0	0.01	0	1
C1 ∪ C2	0	0	0.33	0	1	0
C1 ∪ C3	0	0.15	0	0	0.47	0.53
C2 ∪ C3	0.16	0	0	0	0.53	0.47
Θ	0	0	0	0	1	0

Since choosing the initial probability masses is important, several experiments have been made to investigate how they affect the results of the region evaluation. Assume Table 6.5 is definition 1 of the probabilities for building No. 35, while Table 6.6 shows definition 2. The initial probabilities for data sets 2 and 3 are kept the same, while the probabilities for data set 1 are changed. Since grass is included in class C3 as ground and are possibly confused with the class of tree, the probability of the compound class C1 ∪ C3 is increased. Normalizing the probabilities requires the probability for class C1 to be decreased. From Table 6.6, it is noted that the initial probability for class C2 is unchanged. After data fusion, building No. 35 still can be correctly assigned as building. Although the initial probabilities for clustered image have been changed, the building can be detected because it has enough supporting probabilities.

For definition 3 of probabilities, in Table 6.7, the probabilities for data sets 2 and 3 are same as in definition 1, the probability of t is decreased from its value in Table 6.2 while the compound probability C1 ∪ C3 is increased. Since the supporting probabilities

for the class of building are not high enough, it is not possible to detect building No. 35 using this scenario.

Table 6.6 Building No.35 evaluation using three datasets: clustered image, regions from LevelSet and DSM (definition 2 of probabilities)

A	m_1(A) clustered img	m_2(A) regions LS	m_3(A) DSM	m(A)	Sup(A)	Sup($\overline{A}$)
C1 trees	0.27 ⩛	0.15	0.33	0.47	0.47	0.53
C2 buildings	0.16	0.55	0.34	0.53	0.53	0.47
C3 ground	0.16	0.15	0	0.01	0.01	0.99
C1 ∪ C2	0	0	0.33	0	0.99	0.01
C1 ∪ C3	0.25 ⩚	0.15	0	0	0.47	0.53
C2 ∪ C3	0.16	0	0	0	0.53	0.47
Θ	0	0	0	0	1	0

For definition 4 of probabilities in Table 6.8, the probabilities for data sets 1 and 2 are same as definition 1, the initial probability for the class of building *s* is decreased from its value in Table 6.2 and the compound probability C2 ∪ C3 is increased. In this case building No. 35 can be detected.

Table 6.7 Building No.35 evaluation using three datasets: clustered image, regions from LevelSet and DSM (definition 3 of probabilities)

A	m_1(A) clustered img	m_2(A) regions LS	m_3(A) DSM	m(A)	Sup(A)	Sup($\overline{A}$)
C1 trees	0.52	0.15	0.33	0.65	0.65	0.35
C2 buildings	0.10 ⩛	0.55	0.34	0.34	0.34	0.66
C3 ground	0.10 ⩛	0.15	0	0.01	0.01	0.99
C1 ∪ C2	0	0	0.33	0	0.99	0.01
C1 ∪ C3	0.18 ⩚⩛	0.15	0	0	0.66	0.34
C2 ∪ C3	0.10 ⩛	0	0	0	0.35	0.65
Θ	0	0	0	0	1	0

Table 6.8 Building No.35 evaluation using three datasets: clustered image, regions from LevelSet and DSM (definition 4 of probabilities)

A	m_1(A) clustered img	m_2(A) regions LS	m_3(A) DSM	m(A)	Sup(A)	Sup($\overline{A}$)
C1 trees	0.52	0.15	0.23 ⩛	0.47	0.30	0.70
C2 buildings	0.16	0.55	0.24 ⩛	0.53	*0.57*	*0.43*
C3 ground	0.16	0.15	0	0.01	0.13	0.87
C1∪C2	0	0	0.23 ⩛	0	0.87	0.13
C1∪C3	0	0.15	0	0	0.43	0.57
C2∪C3	0.16	0	0.3 ⩚	0	0.70	0.30
Θ	0	0	0	0	1	0

For definition 5 of probabilities in Table 6.9, the probabilities for data sets 1 and 3 are same as in definition1, the probability u in Table 6.2 is decreased and the compound probability C1∪C2 and C2∪C3 are increased, resulting in correct detection of the class of building.

Table 6.9 Building No.35 evaluation using three datasets: clustered image, regions from LevelSet and DSM (definition 5 of probabilities)

A	m_1(A) clustered img	m_2(A) regions LS	m_3(A) DSM	m(A)	Sup(A)	Sup($\overline{A}$)
C1 trees	0.52	0.12⩛	0.33	0.45	0.45	0.55
C2 buildings	0.16	0.45⩛	0.34	0.55	0.55	0.45
C3 ground	0.16	0.12⩛	0	0	0	1
C1∪C2	0	0.10⩚	0.33	0	1	0
C1∪C3	0	0.11⩛	0	0	0.45	0.55
C2∪C3	0.16	0.10⩚	0	0	0.55	0.45
Θ	0	0	0	0	1	0

From above tests, it is noted that all buildings can be extracted if they are assigned sufficiently high probabilities from multiple data sets. If the probabilities of the union classes, which do not include the class of building, are too high, the class of building will not be correctly assigned as the building class.

For the whole test image, based on the definition of initial probability masses in Table 6.2, there are 96 buildings which are more than 100 pixels in size. The method has successfully detected 85 buildings, which is a detection rate of 88.54%. The result is encouraging.

In order to test different combinations of data sets by Dempster-Shafer data fusion, the two data sets of the clustered image and DSM have been input to MEBRE module. Fusing of these two data sets can only be implemented at pixel level because there is no spatial boundary information. The pixels assigned as the class of building based on the fusion of these two datasets are shown in Figure 6.11. A Sobel edge detector has been used to derive building boundaries as shown in Figure 6.12, which are overlaid on the ortho image displayed in Figure 6.13. It is clear that the clustered image and DSM modelling cannot supply unambiguous building regions. Some road areas are still wrongly assigned as building areas. This shows that the building regions derived from level set modelling is an important information layer in MEBRE module and that fusion based on three data sets implemented at feature level is important.

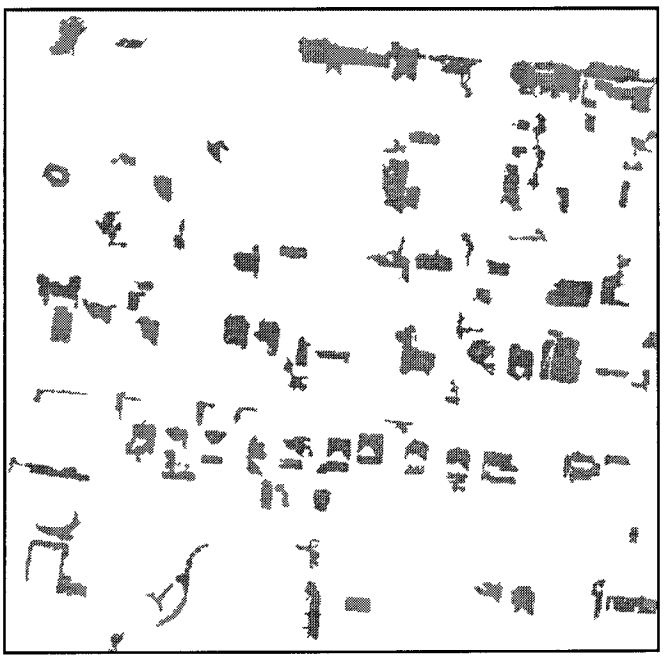


Figure 6.11 The pixels assigned as building class based on classification and DSM fusion

In this chapter, multi-source evidential reasoning based region evaluation (MEBRE) module has been described in detail. Based on Dempster-Shafer theory and tested on one image set, the MEBRE module has been proven an efficient way to evaluate the reliability of the extracted building regions. It is an important procedure in the system of automatic building extraction for 3D reconstruction. In addition, after tuning the initial probabilities of the module, it is noted that when the probabilities of union of classes are increased, the buildings can still be extracted if their initial probabilities from multiple data sets are high enough. The processing of the previous chapters results in some buildings missing, which are not possible to be detected by MEBRE module. Refining the low-level image analysis and interpretation module is an important step to help detect these missing buildings. In next chapter, more test examples will be presented.

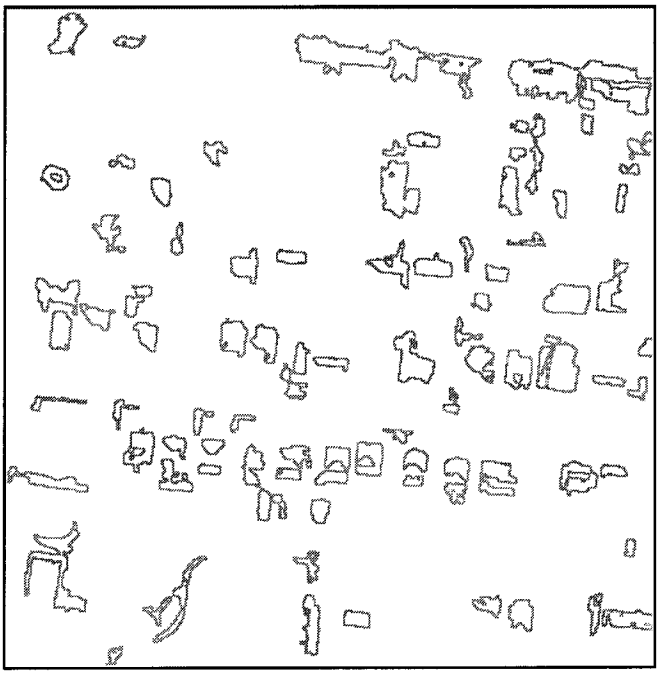


Figure 6.12 Building boundary extracted from Figure 6.11

Figure 6.13 Extracted building boundary overlaid on ortho image

CHAPTER 7

Test and Results

7.1 Introduction

The purpose of this chapter is to demonstrate the effectiveness of the system of automatic building extraction for 3D reconstruction with a number of tests of the techniques described in previous chapters. The test and results so far demonstrate that the four components of the system described in Figure 2.2 are important procedures to effective building extraction. No part can be ignored because each component contributes to the extraction process. The contributions of each process are important for the next level of processing. Four image pairs have been used to dependently test the system in this chapter. Test of the low-level image analysis and interpretation, level set modelling based segmentation and multi-source evidential reasoning based region evaluation will be

analyzed and discussed. The conditions under which these tests have been undertaken are same as the example in Chapter 6. For all the cases in this study, definition 1 described in Chapter 6 has been used as initial probability masses for data fusion.

7.2 Image Data

In order to effective extract the buildings, the high resolution aerial images have been used. The principle distance of the camera is 153mm and the images have been digitized with a pixel size of 0.3 metre. The flying height is 3070 metres.

The image of Test 1 is shown in Figure 7.1, which illustrates one of a pair of aerial images with 515 × 521 pixels in the row and column directions respectively. The image contains a number of white roof and two red buildings. The buildings have a distinguishable contrast against the background.

Figure 7.1 Test area 1

The image in Test 2 is shown in Figure 7.2, which is one of a pair of aerial images with 584 × 521 pixels in the row and column directions respectively, in which the buildings have different shapes and colour. The red and white roof buildings have significant contrast compared with road surface and trees, but the dark roof building such as one in the middle

of the left image is surrounded by the trees and has blurred contrast compared with the surrounding areas.

The image in Test 3 is shown in Figure 7.3 and has a size of 522 × 584 pixels. The majority of buildings have white or red roofs, but there are some dark roof buildings as well. There is good contrast between the buildings and the background.

Figure 7.2 Test area 2

Figure 7.3 Test area 3

The image for Test 4 is shown in Figure 7.4, which has a size of 579 × 515 pixels. The image contains extensive trees which surround the buildings. Two dark roof buildings in the bottom of image have a poor contrast with the background and one of them is hardly recognizable. Each data set will be processed in order.

Figure 7.4 Test area 4

7.3 Description of Experiments

7.3.1 Test 1

The image in Figure 7.1 has been processed as follows. Figure 7.5 illustrates the DSM map obtained from stereo image matching method. Figure 7.6 displays the result of unsupervised classification using the K-Means clustering, where yellow is assigned to building areas, red and green are the vegetation areas, and blue is ground areas. Classification of some areas is not correct, because some ground areas are classified as buildings. Also, some cars are assigned as buildings in the left of image.

Based on the classification of Figure 7.6, using a post classification procedure, a segmented image is created to show the building areas in the classification map in Figure 7.7. Most building areas have been detected, but one red roof building has been completely missed and some road areas have been wrongly assigned as buildings.

The results of processing by VVI are shown in Figure 7.8, where high VVI represents the vegetation and low VVI represents the ground and building areas.

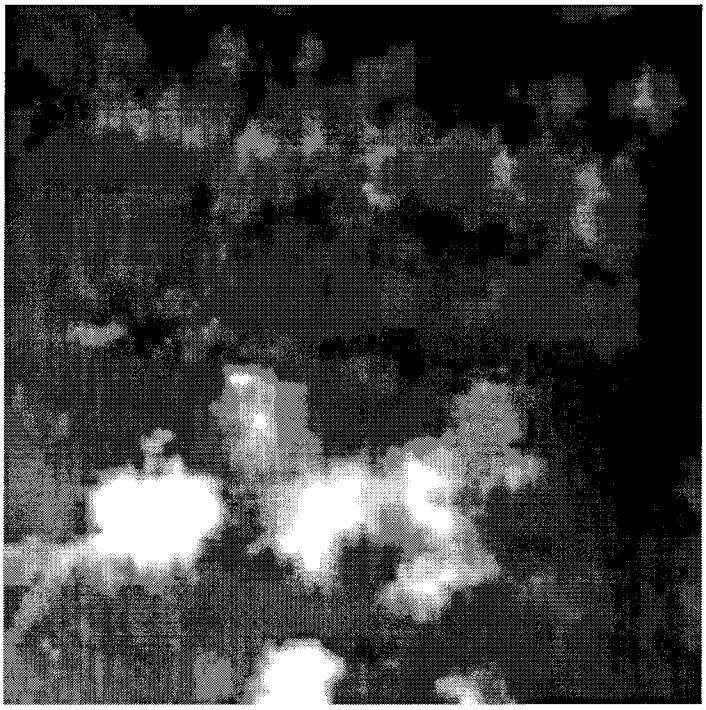


Figure 7.5 DSM from stereo image matching of Test 1

Figure 7.6 Unsupervised classification by clustering

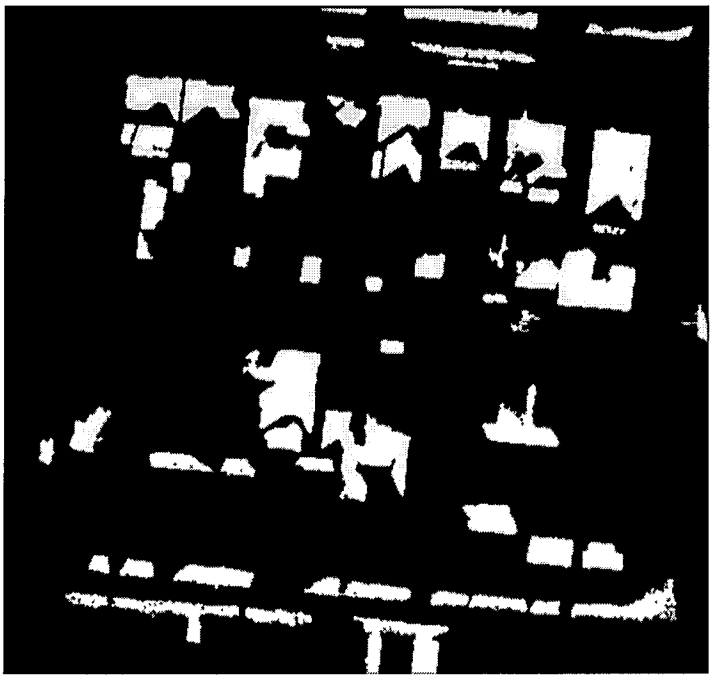


Figure 7.7 Building areas extracted by segmentation on Figure 7.6

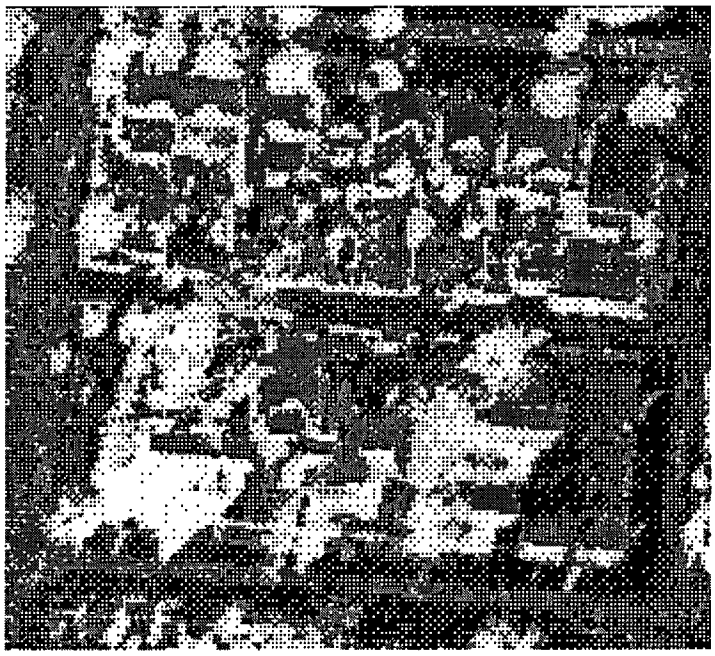


Figure 7.8 VVI image revealing vegetation

The four information data sources of Figures 7.5~7.8 are input into the spatial analysis model. Using the Map Queries operation in ArcView, the possible building areas are

extracted. By implementing the region growing algorithm, the small spots which do not belong to buildings can be deleted from the result. The final building interest areas are derived by analyzing the colour and stereo images as shown in Figure 7.9.

The building interest areas are overlaid on the ortho image in Figure 7.10 as shown in Figure 7.11. Comparing the Figure 7.11 with Figure 7.10, most the road areas wrongly assigned as buildings have been deleted and the correct building areas from classification have been successfully maintained.

Based on the building interest areas in Figure 7.9, level set modelling based image segmentation (LSMBIS) has been used to delineate the boundaries of the buildings, as shown in Figure 7.12. Building regions from the level set modelling overlaid on ortho image are shown in Figure 7.13.

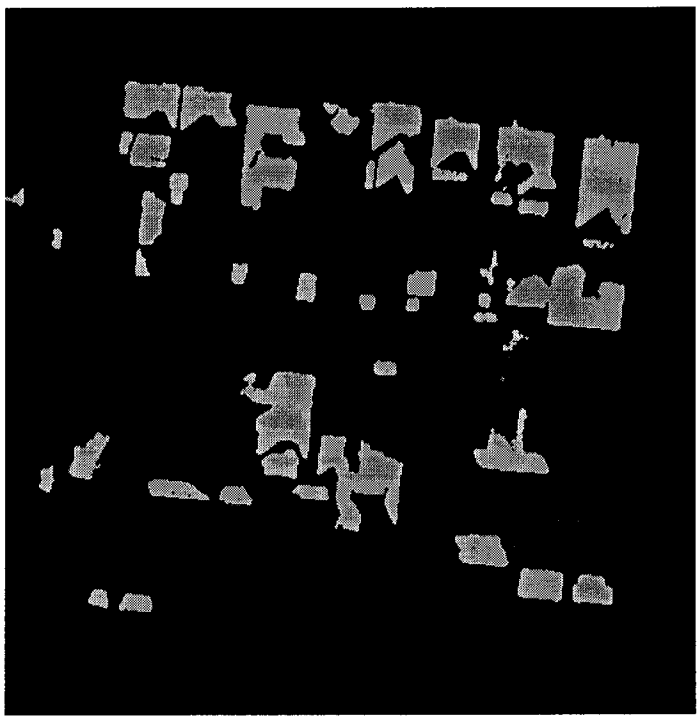


Figure 7.9 Final building interest areas derived after map query operation

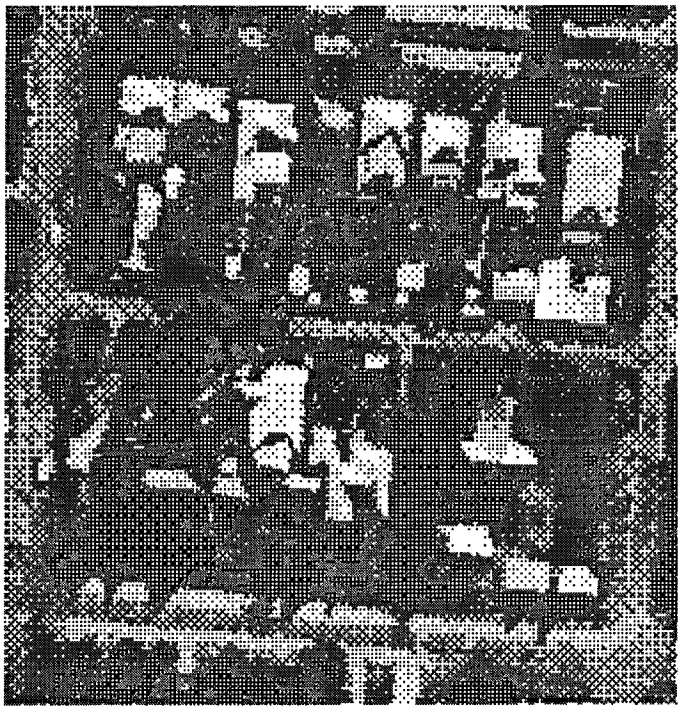


Figure 7.10 The ortho image of Test 1

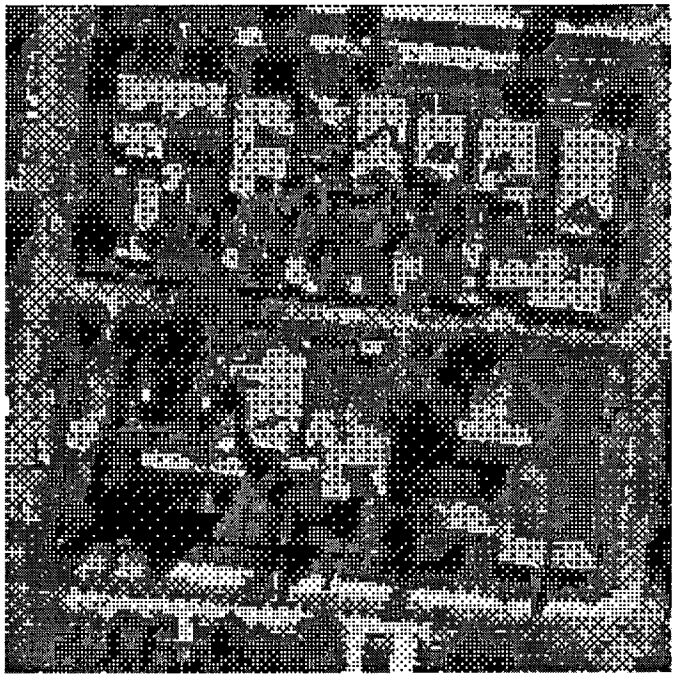


Figure 7.11 The building interest areas overlaid on ortho image in Test 1

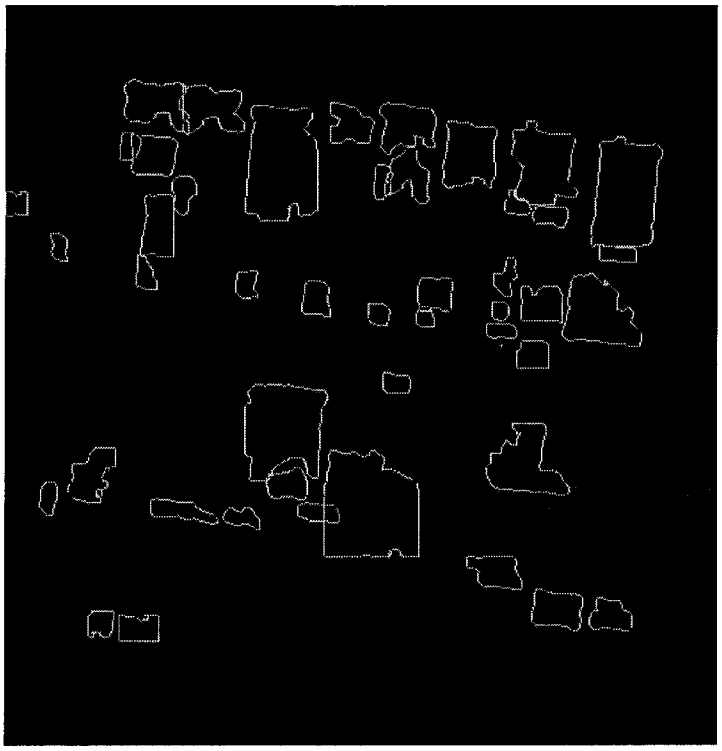


Figure 7.12 Building regions from level set shape modelling in Test 1

Figure 7.13 Building regions from level set modelling overlaid on ortho image in Test 1

In Figure 7.13, it is noticeable that some regions, which belong to road and car areas still are assigned as building areas after the level set shape modelling. This is because the building interest areas supply wrong information and caused the interpretation of some building regions to be unreliable. Thus, it is necessary to use multi-source evidential reasoning based region evaluation (MEBRE) module to evaluate these regions by data fusion.

The results of the application of the Dempster Shafer data fusion algorithm are shown in Figure 7.14. The areas, which were incorrectly assigned as buildings and, have been detected and deleted are shown in green. The blue regions are correctly extracted buildings.

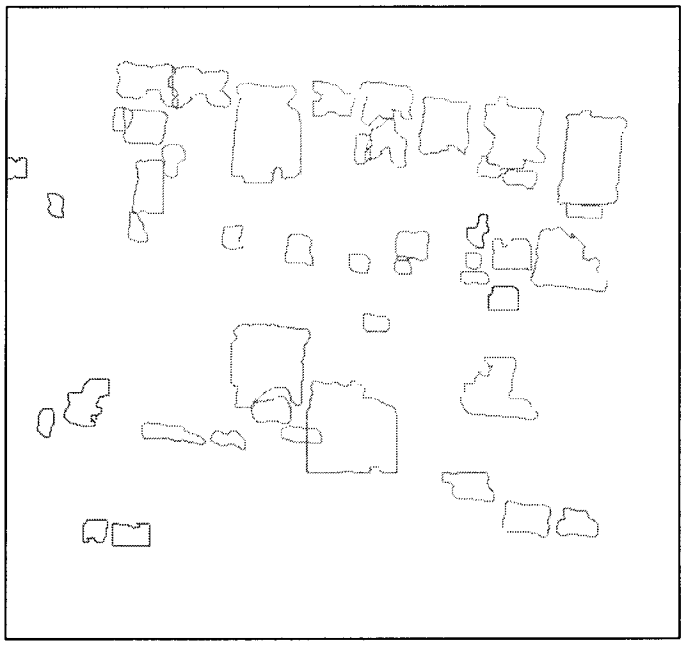


Figure 7.14 Data fusion using three data sets

The building regions derived from MEBRE module evaluation overlaid on ortho image are shown in Figure 7.15. This test demonstrates that the fusion of three data sets is an effective way to evaluate the reliability of the extraction of building regions, because it has deleted all incorrect regions and the correct building areas remain unchanged. In this case, similar roof types have been chosen. One red roofed building is completely missed, while all other buildings have been successfully extracted. One large building exists in

the left bottom of image, which has a partly bright roof and partly dark. As shown in Figure 7.15, only the bright parts of this building have been successfully detected. There are 32 buildings in the image with a size greater than 100 pixels. The detection rate is 96.8%.

Figure 7.15 The building regions from the proposed system overlaid on ortho image for Test 1

7.3.2 Test 2

Since majority of buildings in Test 2 in Figure 7.2 are red and dark roofed building, the classification is more difficult than in Test 1. The result of unsupervised classification is shown in Figure 7.16, where most of the building areas have been detected, but some dark roofed buildings and small buildings are completely missed. Classification of large road areas in the left part of the image is not correct. Also, some cars are assigned as buildings in the top of image.

The four information data sources are input into the low-level image analysis and interpretation module. The final building interest areas overlaid on ortho image are derived as shown in Figure 7.17. Although some large road areas have been wrongly classified as building in LLIAI module, information from DSM and VVI assists to

differentiate between the road areas and building areas. Finally, the road areas wrongly assigned as buildings have been deleted and the correct building areas from classification have been successfully maintained. All the white roofed buildings have been indicated in Figure 7.17, while all the dark roofed buildings are missed. Some red roofed buildings are partly presented.

Figure 7.16 Unsupervised classification by clustering in Test 2

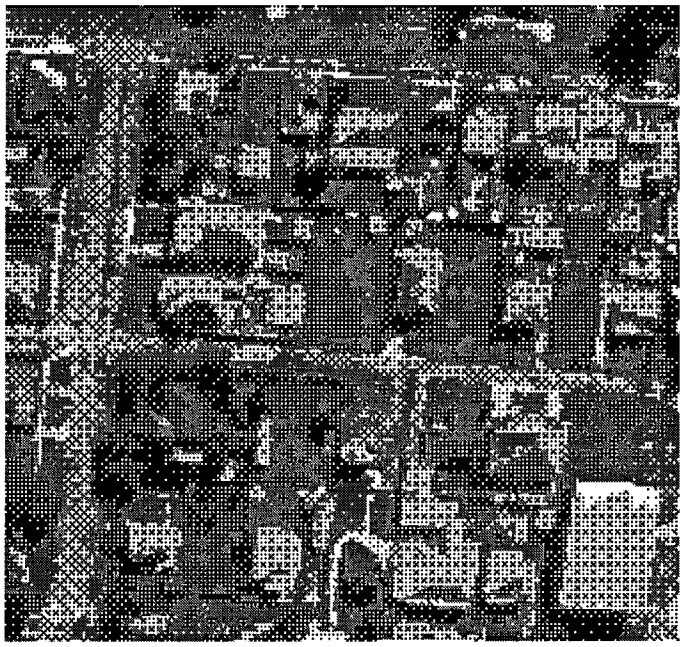


Figure 7.17 The building interest areas overlaid on ortho image in Test 2

Based on the building interest areas in Figure 7.17, the level set modelling based image segmentation (LSMBIS) module has been used to delineate the boundaries of the buildings as shown in Figure 7.18, which are overlaid on ortho image in Figure 7.19.

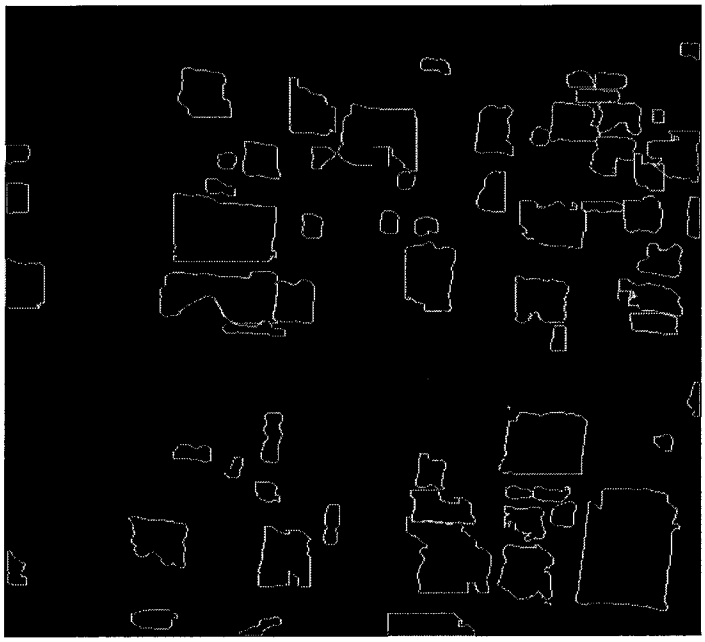


Figure 7.18 Building regions from level set shape modelling in Test 2

Figure 7.19 Building regions from level set modelling overlaid on ortho image in Test 2

In Figure 7.19, some regions in the top and bottom of image, which belong to ground and car areas still are assigned as building areas after level set shape modelling. Subsequently, MEBRE module has been used to evaluate the reliabilities of the building regions as shown in Figure 7.20. The green regions, which were previously incorrectly assigned as building areas, have been detected and deleted. But four red regions are still kept because of their higher elevation values and the texture in DSM and VVI information layers respectively. Four red regions do not have regular building shapes and could be deleted from the building interest areas in Figure 7.17. But this may also delete some correctly assigned building interest areas with red or dark roof, since for most of red and dark roofed buildings, only parts of building areas can be assigned as building interest areas. Hence a reasonable threshold should be chosen to delete small areas, which do not belong to buildings, from building interest areas. In this case, in order to detect more red roofed buildings, a lower threshold has been set, to maintain building interest areas larger than 100 pixels. As a result, four red regions in Figure 7.20, which could not be deleted from the earlier processing have been wrongly kept. The building regions derived by the MEBRE module are overlaid on ortho image as shown in Figure 7.21.

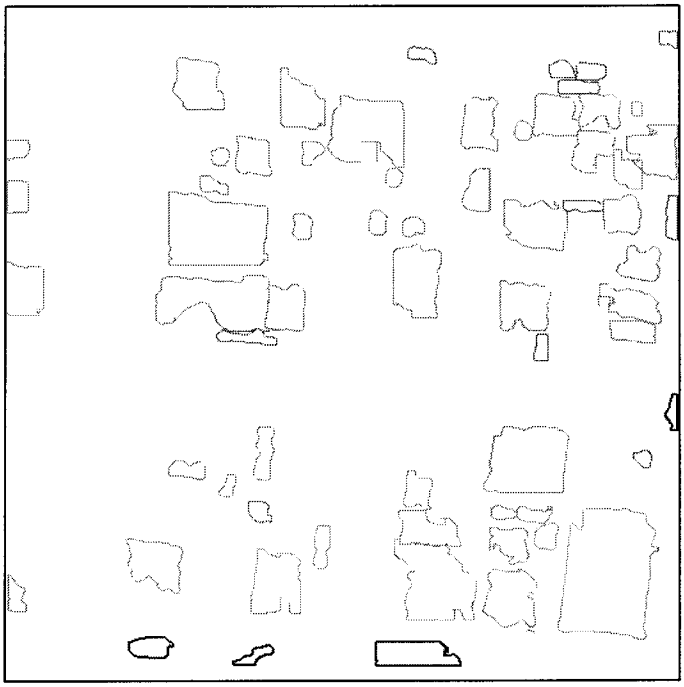


Figure 7.20 Data fusion using three data sets

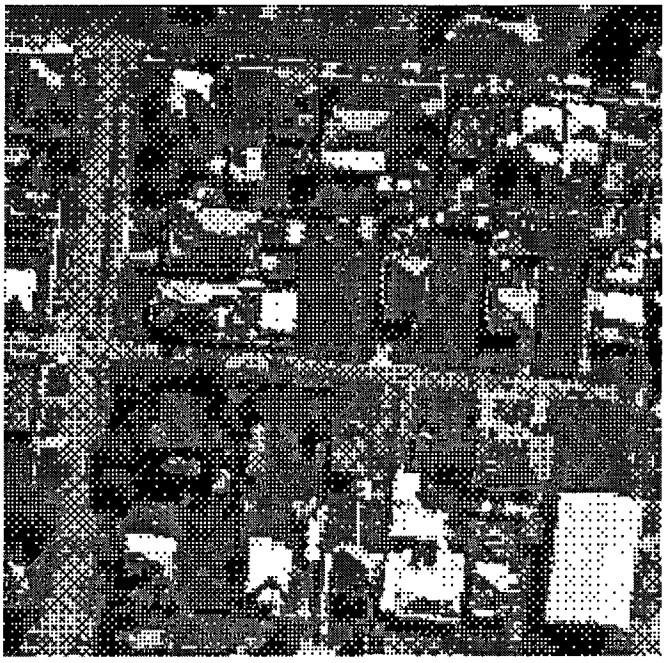


Figure 7.21 The building regions from the proposed system overlaid on ortho image in Test 2

There are 61 buildings in the image and 50 buildings have been correctly detected. Four non-building areas have been assigned as buildings. The detection rate is 81.9%.

7.3.3 Test 3

Figure 7.22 is the result of unsupervised classification on Test 3 as shown in Figure 3. Most of the building areas have been detected, but the dark roofed buildings are completely missed. Some red roofed buildings are partly detected. Also, one car is assigned as a building in the top of image.

After the low-level image analysis and interpretation processing. The final building interest areas overlaid on ortho image are derived as shown in Figure 7.23. In LLIAI module, the larger road area in left bottom of image and a car area in the top of image have been deleted based on the information from DSM and VVI, while the correct building areas from classification have been successfully kept. Based on Figure 7.23, the boundaries of the buildings are delineated by LSMBIS module as shown in Figure 7.24. Building regions from the level set modelling overlaid on ortho image are shown in

Figure 7.25. Although most of the building boundaries are correct, three regions in the top and bottom left are wrongly assigned as building areas. In this case, in order to delete the small spots in building interest areas processing, the threshold for building interest areas is set as 170 pixels. This value maintains some of the red roofed buildings, such as one on the right top of image in Figure 7.23, but it does not successfully delete three wrongly assigned building interest areas on the left top and bottom of image.

Figure 7.22 Unsupervised classification by clustering in Test 3

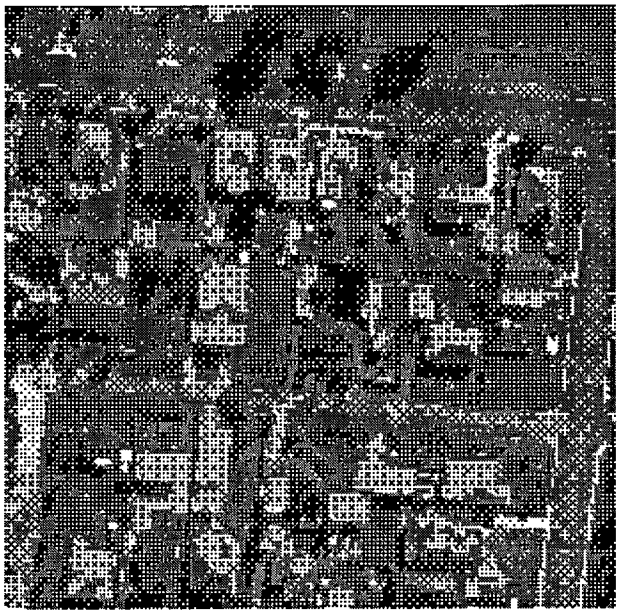


Figure 7.23 The building interest areas overlaid on ortho image in Test 3

MEBRE module has been used to evaluate the reliabilities of the building regions as shown in Figure 7.26. The three wrong building areas mentioned before are detected and deleted after data fusion. The building regions derived from MEBRE module evaluation overlaid on ortho image are shown in Figure 7.27. There are 50 buildings in the image and 40 building are detected. The detection rate is 80%.

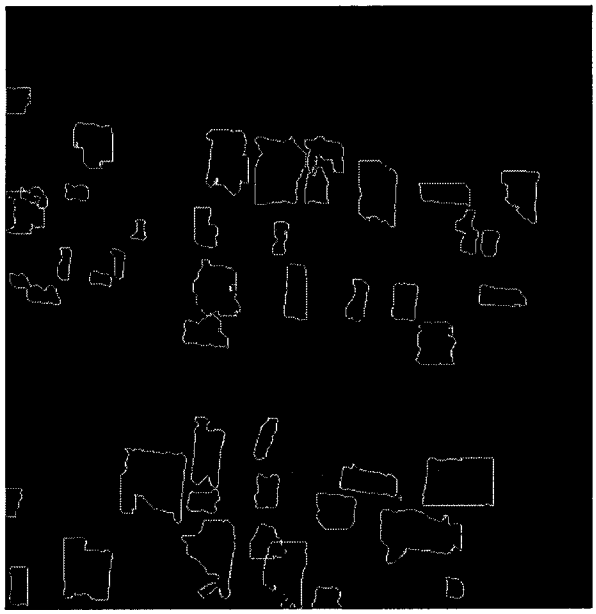


Figure 7.24 Building regions from level set shape modelling in Test 3

Figure 7.25 Building regions from level set modelling overlaid on ortho image in Test 3

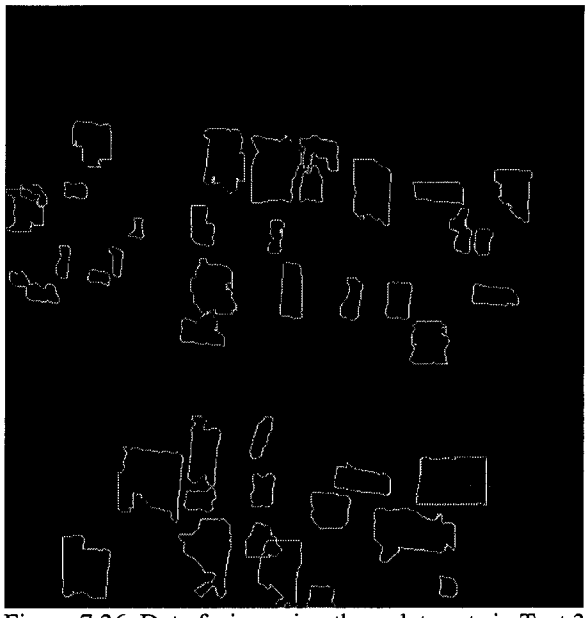


Figure 7.26 Data fusion using three data sets in Test 3

Figure 7.27 The building regions from the proposed system overlaid on ortho image in Test 3

As described in Chapter 6, the determination of initial probabilities are important for data fusion. In order to test the effectiveness of the initial probabilities for Test 3, the initial probabilities have been revised for this case. As in Chapter 6, variations of the initial probabilities were tested and the detection rates were as shown in Table 7.1.

Table 7.1 Definitions for the initial probabilities and the detection rates

Initial prob.	Total buildings	Regions of LevelSet	Detected buildings	Deleted wrong blds	False evaluation	Detection rates
Defintion1	50	43	40	3	0	80%
Definition2	50	43	40	3	0	80%
Definition3	50	43	31	3	9	62%
Definition4	50	43	40	3	0	80%
Definition5	50	43	40	3	0	80%

From Table 7.1, it is clear that the detection rate will be high if the class of buildings have sufficiently high initial probabilities in the multiple data sets. For definition 3, since the initial probabilities for the simple and compound classes of buildings are not high enough, many buildings were not correctly evaluated and the detection rate was lower. For definitions 2, 3 and 4, although the initial probabilities have been changed, the detection rates were not changed because the class of buildings still has enough supporting probabilities. After tuning the initial probabilities, it was proved that the assignment of the initial probabilities by definition1 is a reasonable choice.

7.3.4 Test 4

Figure 7.28 is the result of the classification on Test 4 shown in Figure 7.4. As occurred for the previous cases, all the white roofed buildings have been correctly classified, but many road areas have been wrongly classified as buildings. All the dark roofed buildings have been completely missed, such as two dark roofed buildings in bottom of the image are assigned as the class of trees. Figure 7.29 is the VVI colour image. The range of colour changes from green, yellow to brown, where green represents the lowest VVI value and the brown represents the highest. The two dark roofed buildings mentioned above, are shown as brown and have similar VVI values as trees. Although DSM supplies high elevation values for the two dark roofed buildings, the VVI and classification cannot give supporting information. This is why after LLIAI processing, the dark roofed buildings still cannot be detected as building interest areas. The final building interest areas overlaid on ortho image are derived as shown in Figure 7.30. The boundaries of the buildings are delineated by LSMBIS module as shown in Figure 7.31. Figure 7.32 shows the building regions overlaid on ortho image. Two regions which belong to roads are wrongly assigned as buildings in the top left of Figure 7.31. The building regions from

MEBRE module evaluation overlaid on ortho image are shown in Figure 7.33. Two false building regions have been successfully deleted by MEBRE module, but there is a swimming pool in the left bottom of Figure 7.33 which is wrongly assigned as buildings. A total of 26 buildings exist in Figure 7.4 and 21 buildings are correctly detected as shown in Figure 7.33. The detection rate is 80.8%. Although trees are mixed within the buildings, the system still can be effective to detect the buildings.

Figure 7.28 Unsupervised classification by clustering in Test 4

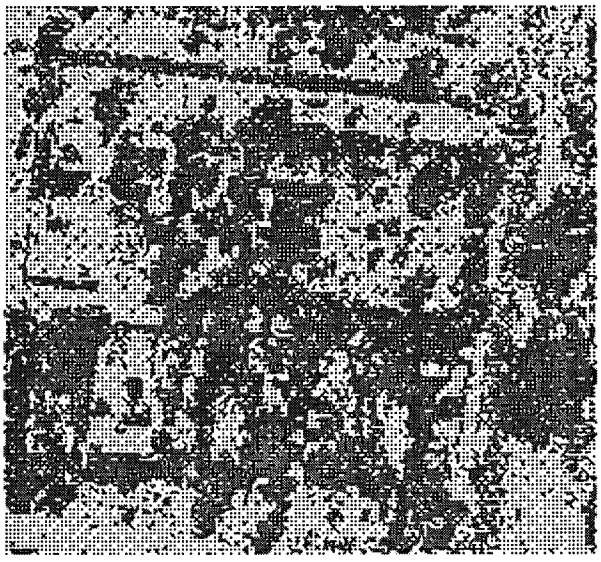


Figure 7.29 VVI colour image revealing the vegetation in Test 4

Figure 7.30 The building interest areas overlaid on ortho image in Test 4

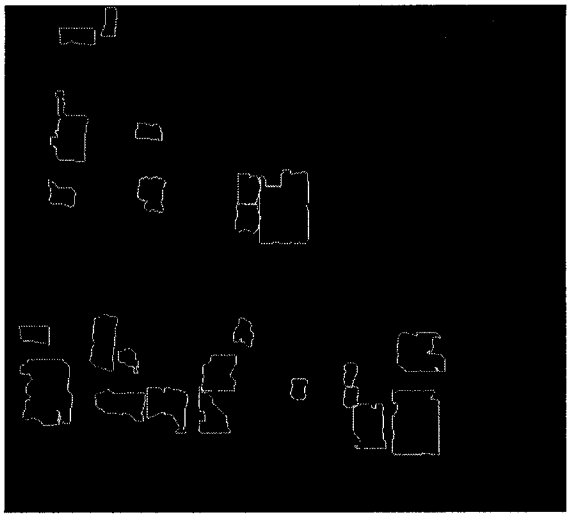


Figure 7.31 Building regions from level set shape modelling in Test 4

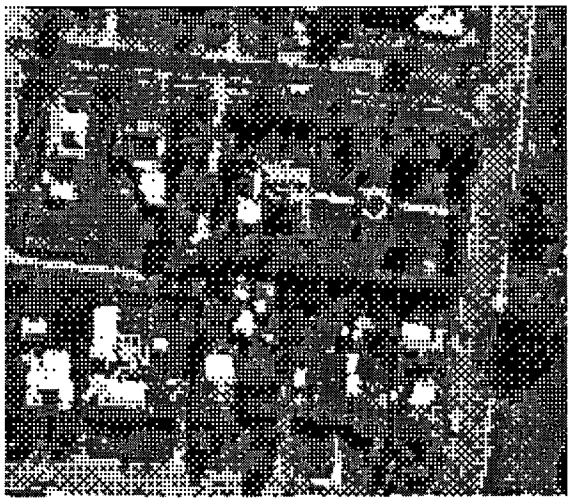


Figure 7.32 Building regions from level set modelling overlaid on ortho image in Test 4

Figure 7.33 The building regions from the proposed system overlaid on ortho image in Test 4

7.4 Summary of Test Results

Table 7.2 gives the over all results of all tests including the 4 tests in this Chapter and the one in Chapter 6.

Table 7.2 Results of all tests in this study

Tests In this study	Total buildings	Regions of LevelSet	Detected buildings	Deleted wrong blds	False evaluation	Detection rates
Test in Ch6	96	91	85	5	1	88.5%
Test1 in Ch7	32	39	31	8	0	96.8%
Test2 in Ch7	61	65	50	11	4	81.9%
Test3 in Ch7	50	43	40	3	0	80%
Test4 in Ch7	26	24	21	2	1	80.8%

The system described in this thesis combines stereo image matching, multispectral image analysis, shape modelling with level set method and Dempster-Shafer data fusion theory, to locate building areas in the test images. The test results demonstrate that each procedure is an important step in determining the accuracy of the system. The testing results also prove that the system of automatic building extraction using interpretation techniques is effective and robust and provides a success rate of greater than 80%.

The purpose of the automatic building extraction is to supply more accurate DTM. One example shown here demonstrates that the extracted building boundaries can be used to overcomes errors in the DSM caused by matching on tops of man-made objects. Figure

7.34 is the extracted building boundaries. Based on a combination of the 3D DSM derived from the image matching and 2D building boundaries, the elevations derived in building regions can be removed from the DSM, as they do not represent the terrain surface, thus leading to a more accurate DTM. In the case of the buildings, the elevations can then be interpolated from the surrounding terrain. Where trees exist, the DSM heights can be reduced by the tree heights. Figures 7.35 and 7.36 illustrate the DSM and 3D

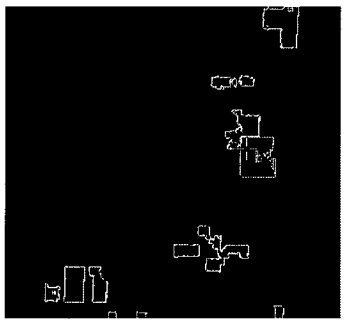


Figure 7.34 Extracted building boundaries

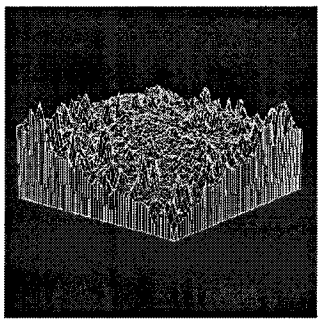


Figure 7.35 DSM from matching

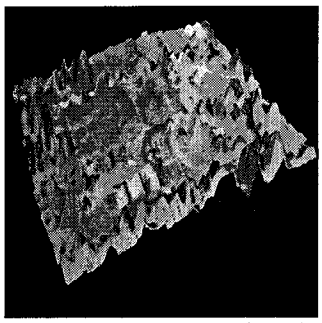


Figure 7.36 3D perspective view from matching

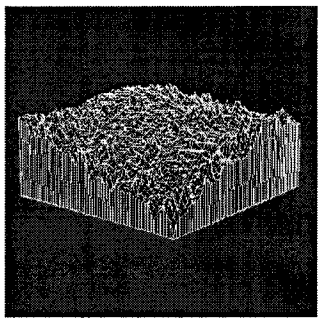


Figure 7.37 more accurate DTM

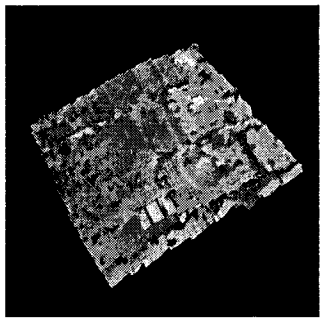


Figure 7.38 3D perspective view from more accurate DTM

perspective view derived directly from stereo image matching. A more accurate DTM and 3D perspective view produced by the method in this study are shown in Figures 7.37 and 7.38. It is noted that the points on the top of buildings in DSM have been put on the ground in Figures 7.37 and 7.38. The DTM represents the bare earth which excludes buildings. In this study, only building areas have been investigated and the corresponding building boundaries have been extracted to reconstruct DTM over these areas. The tree areas should also be extracted for DTM reconstruction over urban areas. The above example shows that the extracted building boundaries do improve DTM reconstruction

CHAPTER 8

Conclusions

8.1 Conclusions of This Study

In this thesis, an automatic building extraction system for 3D reconstruction using interpretation techniques has been presented. The system offers effective algorithms for building extraction for DTM reconstruction from complex photographs over urban areas. Since reliable results for building extraction cannot be achieved by using a single step, the system developed was based on the integration of multiple information from different image interpretation techniques. DSM providing 3D information, and classification, post classification processing and NDVI (VVI in the tests), which represent 2D information, have been used to differentiate buildings from other objects above the ground surface. Based on the building interest areas detected by the above process, 2D shape modelling using the level set method was then used to precisely extract the buildings to guide DTM

reconstruction. The extracted buildings were then evaluated based on the evident from the multiple data sets using data fusion method.

As described in Chapter 1, the main objectives of this study were:

- To improve the process of DTM determination by recognizing types of terrain cover and correction of their effects on the computed DSM.

- To develop robust building extraction procedures that include multiple data sources.

- To establish practical and effective modules in the building extraction process

From the previous chapters, it can be seen that these objective have been successfully achieved. The main contributions of this thesis include:

- **Implementation of robust image matching algorithms for DSM extraction**
  Because there are numerous conventional image matching algorithms and approaches, this thesis concentrates on the development of a matching method, which detects the breaklines in the computation of disparities in the images for areas where discontinuities caused by buildings and trees exist. A robust hybrid-based least squares imaged matching algorithm has been developed. The contribution of this approach is that the disparity map should preserve the discontinuities in the images and supply more a robust DSM for further processing. It overcomes the disadvantages of the least squares method and detects the breaklines and disparity changes during the matching.

- **Establishment of low-level image analysis and interpretation (LLIAI) module based on multiple data sources for extracting building interest areas**
  In this study, the low-level image analysis and interpretation module based on multiple data sources was developed to detect the building interest areas in high resolution images. DSM, land cover classification, the results of the post

classification procedure and NDVI (VVI for the aerial photography case), were integrated into the spatial analysis model to extract building interest areas for further shape modelling. The experiments show this module is an important step in the system. It differentiates buildings from other objects in the images and enables shape modelling to be implemented over the small building interest areas. It overcomes the problems created in 2D segmentation inferencing, when it is based only on single information dataset.

- **Development of level set modelling based image segmentation (LSMBIS) module for delineating the building boundaries**

  To extract building boundaries to guide DTM reconstruction, the level set modelling based image segmentation module was developed. The module offers a highly robust and accurate method for tracking curve interfaces moving under complex motions. The level set method has been applied for the first time in this thesis for complicated urban image analysis. The multiple data sources used for interpretation and integration in the previous procedures, result in the initial modelling of shapes for extracted building interest areas, enabling the level set shape modelling to be achieved without user input. The module also overcomes the complicated and difficult 2D inference in a single image. Since the evolving function in the level set method is associated with edge information in the image through a speed function, when the evolving curve reaches the boundaries of the buildings, all the points on the curve stop evolving further and the computation is ended. Therefore, precise building boundaries can be obtained.

  Since the approximated construction of the zero level set can create a discontinuous curve away from the zero level set, re-initialization of the level set function during fixed time steps is needed to make sure the level set evenly evolves around the curve. For every re-initialization, the discontinuous curve is needed to become a closed curve. A modified seed based pixel-to-pixel region growing procedure has been presented to follow the curve. A seed can be any point on the tracking curve. The method successfully links the points, which

belong to the same curve but are also separated from each other, into a closed curve.

- **Development of multi-source evidential reasoning based region evaluation (MEBRE) module for evaluating extracted building regions.**
  In order to evaluation the reliability of extracted buildings, multi-source evidential reasoning based region evaluation module was developed. Dempster-Shafer theory has been applied in this context for the first time for evaluating multiple data sources for extracting buildings and their reliability. The theoretical basis for the module is also given. The module is unique because it integrates the spectral characteristics and spatial features and evaluates the reliability of the extracted buildings at feature level.

  Similar textures for the road, building roofs and vehicles on road surface, are a common phenomenon in the complicated urban images. They usually lead to misclassification, and thus cause some generated building interest areas to be assigned to an incorrect class. When the level set shape modelling is based on the wrongly assigned building interest areas derived from the earlier steps in the procedure, an incorrect region will be supplied. Region evaluation with redundant data can help eliminate incorrectly extracted regions, while complementary data can provide a more complete description. To remove the wrongly assigned building regions, the three data sets were combined by data fusion. The use of the three data sets rather than two, ie clustered image, regions of LevelSet and DSM, can improve the reliability of system. The comparison of using two and three data sets has been demonstrated in Chapter 6.

  The experiments also demonstrate that the initial probabilities in MEBRE module are important to correctly evaluating the extracted building boundaries. The buildings can be extracted if their initial probabilities derived from the multiple data sets are appropriately assigned. A number of test areas, which include buildings with different sizes, shape and roof colour have been investigated. The

tests are encouraging and demonstrate that the MEBRE module is important procedure for effective building extraction, and the determination of more accurate elevations of the terrain surface. The detection rate largely depends on the building interest areas. As shown in the case studies, some dark roof buildings have been missed in building interest areas and never recovered. MEBRE module cannot recover the missed buildings, but it does supply an effective way to evaluate the reliability of the extracted buildings.

8.2 Further Studies

This thesis has investigated and established an automatic building extraction system for extracting buildings for 3D terrain reconstruction over urban areas. While significant achievements have been made with the system, further studies are also recommended.

- **Modelling more reliable low-level image analysis and interpretation**

In this thesis, building boundaries are extracted based on the previous processes. Thus, the results of building extraction rely on the quality of DSM and low-level image analysis and interpretation. When the building interest areas are not reliable, the level set modelling based image segmentation (LSMBIS) module may produce regions which are wrongly assigned as building areas. Also, the buildings which are missed in the low-level image analysis and interpretation module, will never be recovered from the later processing. Therefore, LLIAI module needs to be further modify to supply more reliable building interest areas. Firstly, multi-spectral images should be more efficient to differentiate between buildings and trees. As mentioned in Chapter 4, because the multi-spectral images were not available for this study, the VVI method was proposed to substitute the NDVI. Because dark roofed buildings have similar VVI values as trees, all dark roofed buildings have be missed in LLIAI module. This limitation may be overcome if multi-spectral images are available. Secondly, the spatial analysis model in the module could comprise more information than the four sets used in this study. This may result in more reliable building interest areas. In addition, the spatial analysis model should also have the ability to combine

information from high resolution aerial images and coarse resolution multi-spectral image. These techniques should be tested on a range of data sets and different scenes. The suitable data sets and modelling method can be obtained by investigation.

- **Integration of matching algorithms**

As mentioned earlier, the proposed matching algorithms can be packaged in a complete software package. If the methods of computing a DSM based on the robust image matching proposed in this thesis prove effective, a more accurate DSM should be available. This module can be input into the automatic building extraction system instead of using DSM from LH Systems' Socet Set v4.2.

- **Modelling more data sources for MEBRE**

The experiments in this study have demonstrate that three data sources can supply more reliable result than two. In some cases, even when three data sources had been used, some extracted building, which are wrongly assigned, cannot be detected as such, and deleted in MEBRE module. Modelling of more data sources from different sensors in the MEBRE module may lead to improved decisions. In addition, GIS data, such as land parcels, can be important information for evaluating the extracted buildings. The correctly extracted buildings should be within known land parcels. Which data sources can be effectively used and how they can be combined with the other data sources in MEBRE module require further investigation.

REFERENCES

Ackermann, F. (1999). Airborne Laser Scanning --- Present Status and Future Expectations. *ISPRS Journal of Photogrammetry and Remote Sensing* 54(2-3): 64-67.

Ackermann F. (1991). MATCH-T: Automatic Mensuration of Digital Elevation Models. *Proc. Sistemas Fotogramétricos Analíticos Digitales: Una nueva generación emergente. Sociedad Española de Cartografía, Fotogrametría y Teledetección,* Barcelona: 67-74.

Ackermann, F. (1984). Digital Image Correlation: Performance and Potential Application in Photogrammetry. *Photogrammetric Record*, 11(64): 429-439.

Allision, D., Zemerly. M. J. A. and Muller, J-P. (1991). Automatic Seed Point Generation for Stereo Matching and Multi-Image Registration. *Proceedings of International Geoscience and Remote Sensing Symposium,* Espoo, Helsinki, Finland, 2417-2421.

Alwan, R. H. and Naji, M. A. (1996). Automatic Stereo Image Matching Using Edge Detection Technique. *International Archives of Photogrammetry and Remote Sensing,* 31(3): 29-35.

Aschwanden P. and Guggenbuhl W. (1993). Experimental Results from a Comparative Study on Correlation-type Registration Algorithms. *In Foustner and Ruwiedel, editors, Robust Computer Vision,* 268-289, Winckmann.

Atkinson, K. B. (1996). Close Range Photogrammetry and Machine Vision. *Whittles publishing.*

Axelsson, P. (1999). Processing of Laser Scanner Data----Algorithms and Applications. *ISPRS Journal of Photogrammetry and Remote Sensing* 54(2-3):138-147.

Ballard, D. H. and Brown, C. M. (1982). *Computer Vision*, Prantice-Hall, Englewood Cliffs, N. J.

Baillard, C and Dissard, O. (2000). A Stereo Matching Algorithm for Urban digital Elevation Models. *Photogrammetric Engineering and Remote Sensing*, 66(9):1119-1128.

Baillard C., Dissard O. Jamet O., and Maitre H. (1998) Extraction and Textural Characterisation of Above Ground Areas from Aerial Stereo Pairs: a Quality Assessment, *ISPRS Journal of Photogrammetry and Remote Sensing*, 53: 130-141.

Baltsavias, E. (1999). A Comparison between Photogrammetry and Laser Scanning, *ISPRS Journal of Photogrammetry and Remote Sensing,* 54(1):83-94.

Baltsavias E., Mason S. and Stallmann D. (1995). Use of DTMs/DSMs and Orthoimages to Support Building Extraction. *Automatic Extraction of Man-made Objects from Aerial and Space Images.* Birkhauser Verlag, Basel, 199-210.

Baltsavias, E. P. (1991). Multiphoto Geometrically Constrained Matching. *Swiss Federal Institute Technology, Zurich, Institute of Geodesy and Photogrammetry,* Report No. 49.

Barnard, S. T. and Fischler, M. A. (1982). Computational Stereo. *Computing Surveys*, 14(4):553-572.

Barnard, S. T. and Thompson, W. B. (1980). Disparity Analysis of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* PAMI-2(4):333-340.

Brockelband, D. C. and Tam, A. P. (1991). Stereo Elevation Determination Techniques for SPOT Imagery. *Photogrammetric Engineering and Remote Sensing,* 57(8):1065-1073.

Brown, L. G. (1992). A Survey of Image Registration Techniques. *ACM Computing Surveys,* 24(4):325-376.

Brunn, A. (2001). Statistical Interpretation of DEM and Image Data for Building
Extraction. *Third International Workshop on Automation Extraction of Man-Made*
Objects from Aerial and Space Images. 10-15 June, Centro Stefano Franscini, Monte
Verita, Ascona, Switzerland.

Brunn, A. and Weidner, U. (1998). Hierarchical Bayesian Nets for Building Extraction using Dense Digital Surface Models. *ISPRS Journal of Photogrammetry and Remote Sensing* 53: 296-307.

Chen A., Donovan G., Sowmya A. and Trinder J. C. (2002). Inductive Clustering: Automating Low-level Segmentation in High Resolution Images. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 34(*3A): 73-78.

Chen L., and Hsu W., (2000). Extraction of Man-Made Buildings in Multispectral Stereoscopic Images, *International Archives of Photogrammetry and Remote Sensing,* 33(3):169-176.

Chui C. K.. (1992). An Introduction to Wavelets, *Academic Press,* New York.

Cochran, S. D. and Medioni, G. (1992). 3-D Surface Description from Binocular Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 14(10):981-994.

Collin, R., Jaynes, C., Cheng, Y., Wang, X. and Stoller, F. (1998). The Ascender System: Automated Site Modeling from Multiple Aerial Images. *Computer Vision and Image Understanding* , 72(2):143-162.

Collins, R., Hanson, A., Riseman, E. and Schultz, H. (1995). Automatic Extraction of Buildings and Terrain from Aerial Images. *Automatic Extraction of Man-Made Objects from Aerial and Spaces Images,* Grun, A., Kubler, O. and Agouris, P. eds. 169-178. Birkhauser Verlag, Basel.

Collin, R., Jaynes, C., Cheng, Y. Wang, X., Stolle, F., Hanson, A. and Riseman, E. (1992). The Ascender System: Automated Site Modelling from Multiple Aerial Images. *Computer Vision and Image Understand-Special Issue on Building Detection and Reconstruction from Aerial Images* 72(2):143-162.

Daubechies I., (1988). Orthonormal Bases of Compactly Supported Wavelets, *Communications in Pure and Applied Mathematics,* 41:909-996.

Daubechies I., (1992). Ten Lectures on Wavelets, *Society of Industrial and Applied Mathematics publications*, Philadhelphia USA.

Day, T. and Muller, J-P. (1989). Digital Elevation Model Production by Stereo-Matching SPOT Image-Pairs: A Comparison of Algorithms. *Image and Vision Computing,* 7(2):95-101.

Dhond, U. and Aggarwal, J. K. (1989). Structure from Stereo - A Review. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6):1489-1510.

Doorn, B., Agouris, P., Al-Tahir, R., Stefanidis, A. and Zilberstein, O. (1990). Digital Stereo Matching: In Perspective. *Technical Notes in Photogrammetry*, No 10. Dept. of Geodetic Science and Surveying. The Onio State University.

Drumheller, M. Poggio, T. (1986). On Parallel Stereo. *IEEE Proceedings of International Conference on Robotics and Automation*, 1439-1448.

Eastman, R. D. and Waxman, A. M. (1987). Using Disparity Functionals for Stereo Correspondence and Surface Reconstruction. *Computer Vision, Graphics and Image Processing*, 39: 73-101.

Fischer, A., Koble, T., Lang, F., Cremers, A., Forstner, W., Plumer, L. and Steinhage, V. (1998). Extracting Buildings from Aerial Images Using Hierarchical Aggregation in 2D and 3D. *Computer Vision and Image Understand-Special Issue on Building Detection and Reconstruction from Aerial Images* 72(2): 185-203.

Foerstner, W. (1982). On The Geometric Precision of Digital Correlation. *International Archives of Photogrammetry and Remote Sensing,* 24(3):176-189.

Foerstner, W. (1986). A Feature Based Correspondence Algorithm for Image Matching. *International Archives of Photogrammetry and Remote Sensing,* 26(3/3): 150-166.

Foerstner, W. and Guelch, E. (1987). A Fast Operator for Detection and Precise Location of Distinct Points, Corner and Centres of Circular Features. *Proceedings of ISPRS Intercommission Workshop on Fast processing of Photogrammetric Data,* Interlaken, 281-305.

Fua, P. (1996). Model-based Optimization: Accurate and Consistent Site Modeling. *International Archives of Photogrammetry and Remote Sensing,* 26(3): 222-233.

Gabet L., Giraudon G. and Remouard L. (1997). Automatic Generation of High Resolution Urban Zone Digital Elevation Model, *ISPRS Journal of Photogrammetry and Remote Sensing* , 52(4): 33-47.

Gordon, T. and Shortliffe, E. (1985). A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space. *Artificial Intelligence* 26: 323-357.

Greenfeld, J. S. and Schenk, A. K. (1989). Experiments with Edge Based Stereo Matching. *Photogrammetry Engineering and Remote Sensing*, 55(12):1771-1777.

Greenfeld, J. S. (1991). An Operator-Based Matching System. *Photogrammetric Engineering and Remote Sensing* 57(8):1049-1055.

Grennfeld. J. S. (1987). A Stereo Vision Approach to Automatic Stereo Matching in Photogrammetry. *Report No. 381, Department of Geodetic Science and Surveying,* Ohio State University, Columbus, Ohio.

Grimson, W. E. L. (1981). A Computer Implementation of A Theory of Human Stereo Vision. *Philosophical Transactions of the Royal Society of London*, B292: 217-253.

Grimson, W. E. L. (1983). Surface Consistency Constraints in Vision. *Computer Vision, Graphics and Image Processing,* 24:28-51.

Grimson, W. E. L. (1985). Computational Experiments with a Feature Based Stereo Algorithm. *IEEE Transactions on Pattern Analysis and Matching Intelligence*, PAMI 7(1): 17-34.

Gruen, A. W. (1985). Adaptive Least Squares Correlation: A Powerful Image Matching Technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography*, 14(3):175-187.

Gruen, A. W. (1997). Automation in Building Reconstruction. *Photogrammetric Week'97*, 175-186. Wichmann Verlag, Karlsruhe.

Gruen, A. W. and Baltsavias, E. P. (1987). Geometrically Constrained multiphoto Matching. *Proceedings of ISPRS Intercommission Workshop on Fast processing of Photogrammetric Data,* Interlaken, 204-230.

Gruen, A. W. and Stallmann, D. (1991). High Accuracy Edge Matching with an Extension of the MPGC-matching Algorithm. *Proceedings of the Society of Photo-Optical Instrumentation Engineers,* 1526: 42-55.

Gonzalez, R. C. and Wintz, P. (1987). *Digital Image Processing.* Second Edition, Addison-Wesley, U. S. A.

Gonzalez, R. C. and Woods, R. E. (1992). *Digital Image Processing.* Addison-Wesley, U. S. A.

Guelch, E. (1988). Results of Test on Image Matching of ISPRS WG III/4. *International Archives of Photogrammetry and Remote Sensing,* 27(3): 254-271.

Haala N. et al (1998). 3D urban GIS from laser altimeter and 2D map data, *International Archives of Photogrammetry and Remote Sensing*, 32(3/1): 339-346.

Haala N., and Hahn M., (1995). Data Fusion for the Detection and Reconstruciton of Building, *Automatic Extraction of Man-Made Objects from Aerial and Space Image,* Birkhauser Press, Edited by A. Gruen.

Habib, A. and Schenk, T., (1999). A New Approach from Matching Surfaces from Laser Scanners and Optical Sensors. *IAPRS, La Jolla, California*, Vol.32, Part 3W14.

Hahn. M. (1993). Measurement by Image Matching - State - of - the - Art in Digtial Photogrammetry. *Photogrammetric Week'93*, Fritsch, D., Hobbie, D. (eds.), 33-42. Wichmann Verlag, Karlsruhe.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). Robust Statistics: The Approach Based on Influence Functions. *John Wiley ad Sons, New York, NY.*

Hanson, A., Marengoni, M., Schultz, H., Stolle, F., Riseman, E. and Jaynes, C. (2001). Ascender II: A Framework for Reconstruction of Scenes from Aerial Images. *Third International Workshop on Automation Extraction of Man-Made Objects from Aerial and Space Images.* 10-15 June, Centro Stefano Franscini, Monte Verita, Ascona, Switzerland.

Hannah, M. J. (1974). Computer Matching of Areas in Stereo Imagery. PhD thesis, Stanford University, AI-Memo 219.

Hannah, M. J. (1988). Digital Stereo Image Matching Techniques. *International Archives of Photogrammetry and Remote Sensing,* 27(B3):280-293.

Hannah, M. J. (1989). A System for Digital Stereo Image Matching. *Photogrammetric Engineering and Remote sensing,* 55(12): 1765-1770.

Haralick, R. M. (1984). Digital Step Edges from Zero Crossing of Second Directional Derivatives. *IEEE Transaction on Pattern Analysis and Machine Intelligence,* PAMI-6(1):58-68.

Harvey, B. R. (1993). Survey Network Adjustments by the L_1 Method. *Australia Journal of Geod. Photogram. Surv*, No.59:39-52.

Hegarat-Mascle S., Bloch I. and Vidal-Madjar D. (1997). Application of Dempster-Shafer Evidence Theory to Unsupervised Classification in Multisource Remote Sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 35(4): 1018-1031.

Helava, U. V. (1988). Object Space Least Squares Correlation. *International Archives of Photogrammetry and Remote Sensing,* 27(3):321-331.

Heipke, C. (1992). A Global Approach for Least-Squares Image Matching and Surface Reconstruction in Objet Space. *Photogrammetric Engineering and Remote Sensing,* 58(3):317-323.

Henricsson O., Bignone F., Willuhn W., Ade F., and Kuebler O. (1996). Project AMOBE: Strategies, Current Status and Future Work. *International Archives of Photogrammetry and Remote Sensing,* 31(3):321-330.

Henricsson O., and Baltsavias E., (1997). 3-D Building Reconstruction with ARUBA: A Quantitative Evaluation, *Automatic Extraction of Man-Made Objects from Aerial and Space Image (II),* Monte Verita, 65-76.

Henricsson O., (1998) The Role Color Attributes and Similarity Grouping in 3-D Building Reconstruction, *Computer Vision and Image Understanding*, 72(2): 163-184.

Hoff, W. and Ahuja, N. (1989). Surfaces from Stereo: Integrating Feature Matching, Disparity Estimation, and Contour Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* PAMI-11(2):121-136.

Horaud, R. and Skordas, T.(1989). Stereo Correspondence through Feature Grouping and Maximal Cliques. *IEEE Transaction Pattern Analysis and Machine Intelligence*, PAMI-11(11):1168-1180.

Horiguchi S.,Ozawa S., Nagai S., and Sugiyama K., (2000). Reconstruction Road and Block from DEM in Urban Area, *International Archives of Photogrammetry and Remote Sensing,* 33(3):413-420.

Huber , P. J. 1981. Robust Statistics. John Wiley and Sons, New York, NY.

Hug, C. (1997). Extracting Artificial Surface objects from Airborne Laser Scanner Data. In Guren, A. Baltsavias, E. and Henricsson, O. (Eds.), *Automatic Extraction of Man-Made Objects from Aerial and Space images (II), Birkhauser, Verlag, Base*l, 203-212.

Huertas A., and Nevatia R., (1988). Detecting Buildings in Aerial Images, *Computer Graphics and Image Processing*, Vol. 41, No.2.

Jain, A. K. (1989). *Fundamentals of Digital Image Processing*. Prentice Hall.

Jaynes, C., Hanson, A., Riseman, A., and Schultz, H. (1997a). Building Reconstruction from Optical and Range Images. Proc. *International Conference on Computer Vision and Pattern Recognition (CVPR'97)*, 380-386.

Jaynes, C., Collins, R., Cheng, Y., Wang, X., Stolle, F., Schultz, H., Hanson, A., and Riseman, E. (1997b). Automatic Construction of Three-Dimensional Models of Buildings. In O. Firschein and T. Strat (eds), *RADIUS: Image Understanding for Imagery Intelligence, 223-236. San Francisco: Morgan Kaufmann.*

Kim, Y. C. and Aggarwal, J. K. (1987). Positioning 3-D objects Using Stereo Images. *IEEE Journal of Robotics and Automation*, RA-3(4):361-373.

Kim T., and Muller J. P., (1998). A Technique for 3-D Building reconstruction, *ISPRS Journal of Photogrammetry and Remote Sensing* , 64(9):923-930.

Klein L. (1999). Sensor and Data Fusion Concepts and Applications. *SPIE Optical Engineering Press.*

Lammi, J. (1996). 3-D Modelling of Buildings From Digital Aerial Imagery. *International Archives of Photogrammetry and Remote Sensing*, 31(2):213-217.

Lang, F. and Foerstner, W. (1995). Matching Techniques. *Proceedings 2nd Course In Digital Photogrammetry.*

Levy, M. (1988). A New Theoretical Approach to Relaxation Application of Edge Detection. *Proceedings of the 9th International Conference on Pattern Recognition,* 208-212.

Li, M. (1991). Hierarchical Multipoint Matching. *Photogrammetric Engineering and Remote Sensing,* 57(8): 1039-1047.

Lin, C. and Nevatia, R. (1998). Building Detection and Description from a Single Intensity Image. *Computer Vision, Graphics and Image Processing*, 72(2):101-121.

Liow Y. T., and Pavlidis T., (1990). Use of Shadows for Extracting Buildings in Aerial Images, *Computer Vision, Graphics and Image Processing*, 49: 242-277.

Lo, K. C. and Mulder, N. J. (1992). Automatic DEM Data Generation and Change Detection by Region Matching. *International Archives of Photogrammetry and Remote Sensing,* 29(3):476-480.

Lu, Y., Trinder J. and Kubik K. (2003). Automatic Building Extraction based on Multi-source Data Fusion. *Joint workshop of ISPRS WG I/2, I/5 and IC WG II/IV and EARSeL special interest group 3D remote sensing*, at the University of Hannover October 6-8, Germany.

Lu, Y. and Trinder J. (2003). Data Fusion Applied to Automatic Building Extraction. *International Archives of Photogrammetry and Remote Sensing,* 5-9, May, Anchorage, Alaska, USA.

Lu, Y. and Trinder J. (2002). Dempster-Shafer Data Fusion Theory for Automatic Building Extraction in 3D Reconstruction. *The 11th Australasian Remote Sensing and Photogrammetry Association Conference, September2-6, 2002,* Brisbane, Australia.

Marr, D. and Poggio, T. (1979). A Computational Theory of Human Stereo Vision. *Proceedings of Royal Society of London,* B204, 301-328.

Mallat S., (1989a) A Theory for Multi-resolution Signal Decompostion: the Wavelet Representation, *IEEE Trans. PAMI,* 11:674-693.

Mallat S. (1989b). Multifrequency Channel Decompositions of Images and Wavelet Models, *IEEE Trans. On Acoustics, Speech, and Signal processing,* 37(12): 2091-2110.

Mallat S., and Zhong S. (1992). Characterization of Signals from Multiscale Edges, *IEEE Trans. PAMI,* 14(7):710-732.

Masaharu H., and Hasegawa H., (2000). Three-Dimensional City Modeling from Laser Scanner Data by Extracting Building Polygons Using Region Segmentation Method, *International Archives of Photogrammetry and Remote Sensing,* 33(3):556-562.

Mcintosh K., Krupnik A., and Schenk T. (2000). Improvement of Automatic DSM Generation Over Urban Areas Using Airborne Laser Scanner Data, *International Archives of Photogrammetry and Remote Sensing,* 33(3):563-570.

Mcintosh, J. H. and Match, K. M. (1988). Matching straight lines. *Computer Vision, Graphics, and Image Processing,* 40: 386-408.

Mckeown D. M., (1991) Information fusion in Catographic Feature Extraction from Aerial Imagery, *Digital Photogrammetric Systems,* Wichmann, 103-110.

Medioni, G. and Nevatia, R. (1984). Matching Images Using Linear Features. *IEEE Transacitons on Pattern Analysis and Machine Intelligence,* PAMI-6(6):675-685.

Medioni, G. and Nevatia, R. (1985). Segment-Based Stereo Matching. *Computer Vision, Graphics, and Image Processing*, 31: 2-18.

Mei, X, Gong. P and Biging G. (2001). Image Matching Based on Tracking Matching Paths in the Similarity Space. *Photogrammetric Engineering and Remote Sensing*, 67(4): 453-460.

Moravec, H. P. (1977). Towards Automatic Visual Obstacle Avoidance. *Proceedings 5th International Joint Conference on Artificial Intelligence,* Cambridge, MA, pp. 584.

Nevatia, R., Huertas, A and Kim, Z. (1999). The MURI Project for Rapid Feature Extraction in Rrban Areas. *Automatic Extraction of GIS objects from Digital Imagery*, IAPRS, 32(3-2W5): 3-14.

Niederost, M. (2001) Automated Update of Building Information in Maps Using Medium-scale imagery. *Third International Workshop on Automation Extraction of Man-Made Objects from Aerial and Space Images.* 10-15 June, Centro Stefano Franscini, Monte Verita, Ascona, Switzerland.

Noronha, S. and Nevatia, R. (2001). Detection and Modeling of Buidings from Multiple Aerial Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-23(5):501-518.

Paparoditis, N., Maillet, G., Taillandier, F., Jung, F., Guigues, L. and Boldo, D. (2001). Multi-image 3D Feature and DSM Extraction for Change Detection and Building Reconstruction. *Third International Workshop on Automation Extraction of Man-Made Objects from Aerial and Space Images.* 10-15 June, Centro Stefano Franscini, Monte Verita, Ascona, Switzerland.

Paparoditis, N., Cord, M., Jordan, M. and Cocquerez, J. (1998). Building Detection and Reconstruction from Mid- and High-Resolution Aerial Imagery. *Computer Vision and Image Understand-Special Issue on Building Detection and Reconstruction from Aerial Images* 72(2), pp.122-142.

Price, K. (1985). Relaxation Matching Techniques—A Comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7(5):617-623.

Ohta, Y. and Kanade, T. (1985). Stereo by Intra and Inter-Scanline Search Using Dynamic Programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7(2):139-154.

Osher S. & Sethian J.A. (1988). "Fronts Propagating with Curvature Dependent Speed: Algorithms based on Hamilton-Jacobi formulation", *J. of Computational Phhysics*, (79):12-49

Quiguer, T. (1996). Rectangular Building 3D reconstruction in Urban Zones. *International Archives of Photogrammetry and Remote Sensing,* 26(3): 657-662.

Rasoulian, H., Thompson, W., Kazda, L. and Parra-loera, R. (1990). Application of the Mathematical Theory of Evidence to the Image Cueing and Image Segmentation Problem. *SPIE Signal and Image Processing Systems Performance Evaluation* 1310: 199-206.

Reeves, R., and Kubik K. (1998). Least Squares Matching in the Transform Domain. *ISPRS Commission 3 Symposium on Object Recognition and Scene Classification from Multispectral and Multisensor Pixels*, July 6-10, 1998, Columbus, Ohio, USA.

Richards J. and Jia X. (1999). *Remote Sensing Digital Image Analysis*.

Riply B. D. (1996). Pattern Recognition and Neural Networks. *Cambridge University Press.*

Rosenfeld, A., Hummel, R. A. and Zucker, S.W. (1976). Scene Labeling by Relaxation Operations. *IEEE Transactions on Systems, Man, and Cybernetics,* SMC-6(6):420-433.

Rosenfeld, A. (1982). Digital Picture Processing. *Academic Press, Inc. Ltd., London.*

Rosenfeld, A. ed. (1984). Multiresolution Image Processing and Analysis. *Springer-Verlag, Berlin.*

Rousseeuw, P. J. and Leroy, A. M. (1987). Robust Regression and Outlier Detection. *John Wiley and Sons, New York.*

Russ, J. C. (1995). The Image Processing Handbook. *CRC Press*, Inc, U.S.A.

Sahar L., and Krupnik A., (1999). Semiautomatic Extraction of Building Outlines from Large-Scale Aerial Image, *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(4): 459-465.

Saint-Marc, P., Chen J. and Medioni, G. (1991). Adaptive Smoothing: A General Tool for Early Vision. *IEEE Transactions on pattern Analysis and Machine Intelligence,* 13(6):514-529

Sarkar, S. and Boyer, K., (1991). Optimal Infinite Impulse Response Zero Crossing Based Edge Detectors. *Computer Vision, Graphics and Image Processing*, 54(2): 224-243.

Schenk, T., Postolov, Y. and Krupnik, A. (2000). A Comparison of Surface Matching Algorithms Tested Over an Urban Area Using Airborne Laser Data and Photogrammetric Data. *International Archives of Photogrammetry and Remote Sensing,* 33(3): 563-570.

Schenk, T. Li, J. C. and Toth C. (1991). Towards an Autonomous System for Orienting Digital Stereopairs. *Photogrammetric Engineering and Remote Sensing,* 57(8):1057-1064.

Sethian J. A. (1985). Curvature and the evolution of fronts, *Comm. In Math. Phys. 54*: 487-499.

Sethian J. A. (1995). Shape Modelling with Front Propagation: A Level Set approach, *IEEE Transactions on Pattern analysis and Machine Intelligence,* 17(2): 158-175.

Sethian J. A. and Strain J. (1992). Crystal Growth and Dendritic Solidification, *Journal of Computational Physics*, 99:231-253.

Sethian J. A. (1999). Level set methods and fast marching methods, *Cambridge University Press.*

Shafer G. (1976). A mathematical Theory of Evidence. *Princeton, NJ: Princeton University Press.*

Shapiro, L. (1980). A Structural Model of Shape. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* PAMI 2(2):111-126.

Shapiro, L. and Haralick, R. (1981). Structural Descriptions and Inexact Matching. *IEEE Transaction on Pattern Analysis and Machine Intelligence,* PAMI 3(5):504-519.

Shi, Z. and Shibasaki, R. (1995). Multiresolution Analysis Based Region Segmentation and Stereo Matching. *Proc. of 6th Symposium on Functional Graphic Information System*, Tokyo, 63-68.

Shi, Z. and Shibasaki, R. (1995). Automated Extraction of Man-Made Structures Using Region and Line-based Stereo Matching in Digital Aerial Images. *Proceeding of GIS AM/AF ASIA '95*, Tailand, G-1-1~ G-1-10.

Shufel, J. and McKeown, D (1993). Fusion of Monocular Cues to Detect Man-Made Structures in Aerial Imagery. *Computer Vision, Graphics, and Image Processing* , 57(3): 307-330.

Sohn, G. and Dowman, I. (2001). Extraction of Buildings from High resolution Satellite

Data. *Third International Workshop on Automation Extraction of Man-Made Objects from Aerial and Space Images.* 10-15 June, Centro Stefano Franscini, Monte Verita, Ascona, Switzerland.

Spreeuwears L., Schutte K., and Houkes Z., (1997). A Model Driven Approach to Extract Buildings from Multi-View Aerial Imagery, *Automatic Extraction of Man-Made Objects from Aerial and Space Image (II)*, Birkhauser Verlag, Berlin, pp109-118.

Stein A., Meer F. and Gorte B., (1999). Spatial statistics for Remote Sensing. *Kluwer Academic Publishers.*

Steinle, E. and Vogtle, T. (2001). Automated Extraction and Reconstruction of Buildings in Laser Scanning Data for Disaster Management. *Third International Workshop on Automation Extraction of Man-Made Objects from Aerial and Space Images.* 10-15 June, Centro Stefano Franscini, Monte Verita, Ascona, Switzerland.

Straub B. and Heipke C. (2001). Automatic Extraction of Trees for 3D City Models from Images and Height Data. *Third International Workshop on Automation Extraction of Man-Made Objects from Aerial and Space Images.* 10-15 June 2000, Centro Stefano Franscini, Monte Verita, Ascona, Switzerland.

Sowmya A. and Trinder. J. (2000). Modelling and Representation Issues In Automated Feature Extraction From Aerial and Satellite Images. *Journal of Photogrammetry and Remote Sensing*, 55:34-37.

Tu, X. W. and Dubuisson, B. (1990). 3D Information Derivation from a Pair of Binocular Images. *Pattern Recognition,* 23(3/4):223-235.

Van Cleynenbreugel, J., Osinga, S., Fierene, F. Suetens, P. and Oosterlinck A. (1991). Road Extraction from Multi-temporal Satellite Images by an Evidential Reasoning Approach. *Pattern Recognition Letter*, 12:371-380.

Vosselman, G. (1999). Building Reconstruction Using Planar Faces in Very High Density Height Data. *IAPRS, Munich,* 32(3-2W5): 87-92.

Vosselman, G. (1992). Relational Matching. *Lecture Notes in Computer Science*, Vol. 628, Springer Verlag, Berlin.

Wang, Z. (1990). Principles of Photogrammetry. *Publishing House of Surveying and Mapping*, Beijing.

Wolf, P. R. (1983) Elements of Photogrammetry. *McGraw-Hill Inc.,* New York.

Wrobel, B. (1987). Facet Stereo Vision (Fast Vision) - A New Approach to Computer Stereo Vision and to Digital Photogrammetry. *Proceedings of ISPRS Intercommission Workshop on Fast processing of Photogrammetric Data,* Interlaken Switzerland, 231-258.

Zhang, Z. (1995). Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting. *French National Institute for Research in Computer Science and Control,* Report No.2676 .

Zhang, Z., Deriche, R., Faugeras, O. and Luong, Q. (1995). A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry. *Artificial Intelligence, 78(199) 87-119.*

Zhang, Z. and Zhou, Y. (1989). A New Approach to Arrange the Approximate Epipolar Lines for SPOT Images. *Special issue of Seminar on the Wang Zhizhuo's Academic Thinking,* Chinese Society of Geodesy, Beijing, PRC, 46-50.

Zilberstein, O. (1992). Relational Matching for Stereopsis. *International Archives of Photogrammetry and Remote Sensing*, 29(3):711-719.

Publications from the

SCHOOL OF SURVEYING AND SPATIAL INFORMATION SYSTEMS
(formerly: SCHOOL OF GEOMATIC ENGINEERING)
THE UNIVERSITY OF NEW SOUTH WALES
ABN 57 195 873 179

To order, write or fax to:
Publications Officer, School of Surveying and Spatial Information Systems,
The University of New South Wales, UNSW SYDNEY NSW 2052, AUSTRALIA
Fax: +61 - 2 - 9313 - 7493

NOTE ON MAIL ORDERS: ALL ORDERS MUST BE PREPAID.
NO REQUESTS FOR RETURNS ARE ACCEPTED.
CREDIT CARDS ARE ACCEPTED. CONDITIONS APPLY. SEE ORDER FORM FOR DETAILS.
WHEN PAYING BY *CHEQUE*, THE CHEQUES MUST BE PAYABLE TO 'SCHOOL OF SURVEYING & SIS, UNSW',
IN AUSTRALIAN DOLLARS AND DRAWN ON AN AUSTRALIAN BANK.

MONOGRAPHS

Australian prices include GST #. Overseas prices include delivery by UNSW's air-lifted mail service (~2-4 weeks to Europe and North America) #. Rates for air mail through Australia Post on application.

For all (local and overseas) **orders, add $11.00 for processing and handling** (once per order).

(Prices effective July 2004)

		Price # Australia (incl. GST)	**Price** # Overseas
M1.	R. S. Mather, "The Theory and Geodetic Use of some Common Projections", (3rd edition), 125 pp, reprint 1998, ISBN 0-85839-007-8	$ 16.50	$ 15.00
M2.	R. S. Mather, "The Analysis of the Earth's Gravity Field", 172 pp, 1971 (no ISBN)	$ 8.80	$ 8.00
M8.	A. H. W. Kearsley, "Geodetic Surveying", 96 pp, revised reprint 1988, ISBN 0-85839-036-1	$ 13.20	$ 12.00
M11.	W. F. Caspary, "Concepts of Network and Deformation Analysis", 183 pp, 3rd corrected impression, 2000, ISBN 0-85839-044-2	$ 27.50	$ 25.00
M12.	F. K. Brunner, "Atmospheric Effects on Geodetic Space Measurements", 110 pp, 1988, ISBN 0-85839-048-5	$ 17.60	$ 16.00
M13.	B. R. Harvey, "Practical Least Squares and Statistics for Surveyors", (2nd edition, reprinted with corrections), 319 pp, 1998, ISBN 0-85839-059-0	$ 33.00	$ 30.00
M14.	E. G. Masters and J. R. Pollard (Eds.), "Land Information Management", 269 pp, 1991, ISBN 0-85839-061-2	$ 22.00	$ 20.00
M15/1	E. G. Masters and J. R. Pollard (Eds.), "Land Information Management - Geographic Information Systems - Advance Remote Sensing Vol. 1", 295 pp, 1993, ISBN 0-85939-064-7	$ 33.00	$ 30.00
M15/2	E. G. Masters and J. R. Pollard (Eds.), "Land Information Management - Geographic Information Systems - Advance Remote Sensing Vol. 2", 376 pp, 1993, ISBN 0-85839-065-5	$ 33.00	$ 30.00
M16.	A. Stolz, "An Introduction to Geodesy", 2nd extended edition, 148 pp, 2001, ISBN 0-7334-1736-1	$ 24.20	$ 22.00
M18.	J. M. Rüeger, "Electronic Surveying Instruments – A Review of Principles, Problems and Procedures", 1st ed., 166 pp, 2003, ISBN 0-7334-2083-4	$ 25.30	$ 23.00

UNISURV REPORTS - S SERIES

(Prices effective July 2004)

Australian Prices *:	S8 - S20	$11.00 #	S29 onwards	$33.00 #
Overseas Prices **:	S8 - S20	$10.00 #	S29 onwards	$30.00 #

* Australian prices include GST. #

** Overseas prices include delivery by UNSW's air-lifted mail service (~2-4 weeks to Europe and North America). # Rates for air mail through Australia Post on application.

For all (local and overseas) **orders, add $11.00 for processing and handling.**

S17. C. Rizos, "The Role of the Gravity Field in Sea Surface Topography Studies", Unisurv S17, 299 pp, 1980, ISBN 0-85839-029-9

S19. R. Coleman, "A Geodetic Basis for Recovering Ocean Dynamic Information from Satellite Altimetry", Unisurv S19, 332 pp, 1981, ISBN 0-85839-029-9

S36. A. R. Marshall, "Network Design and Optimisation in Close Range Photogrammetry", Unisurv S36, 249 pp, 1989, ISBN 0-85839-054-X

S37. W. Jaroondhampinij, "A Model of Computerised Parcel-Based Land Information System for the Department of Lands, Thailand", Unisurv S37, 281 pp, 1989, ISBN 0-85839-055-8

S39. C. Bosloper, "Multipath and GPS Short Periodic Components of the Time Variation of the Differential Dispersive Delay", Unisurv S39, 214 pp, 1990, ISBN 0-85839-057-4

S40. J. M. Nolan, "Development of a Navigational System Utilising the Global Positioning System in a Real Time, Differential Mode", Unisurv S40, 163 pp, 1990, ISBN 0-85839-058-2

S41. R. T. Macleod, "The Resolution of Mean Sea Level Anomalies along the NSW Coastline Using the Global Positioning System", Unisurv S41, 278 pp, 1990, ISBN 0-85839-060-4

S42. D. A. Kinlyside, "Densification Surveys in New South Wales - Coping with Distortions", Unisurv S42, 209 pp, 1992, ISBN 0-85839-062-0

S43. A. H. W. Kearsley (ed.), Z. Ahmad, B. R. Harvey and A. Kasenda, "Contributions to Geoid Evaluations and GPS Heighting", Unisurv S43, 209 pp, 1993, ISBN 0-85839-063-9

S44. P. Tregoning, "GPS Measurements in the Australian and Indonesian Regions (1989-1993)", Unisurv S44, 134 + xiii pp, 1996, ISBN 0-85839-068-X

S45. W.-X. Fu, "A Study of GPS and Other Navigation Systems for High Precision Navigation and Attitude Determinations", Unisurv S45, 332 pp, 1996, ISBN 085839-069-8

S46. P. Morgan et al, "A Zero Order GPS Network for the Australia Region", Unisurv S46, 187 + xii pp, 1996, ISBN 0-85839-070-1

S47. Y. Huang, "A Digital Photogrammetry System for Industrial Monitoring", Unisurv S47, 145 + xiv pp, 1997, ISBN 0-85839-072-8

S48. K. Mobbs, "Tectonic Interpretation of the Papua New Guinea Region from Repeat Satellite Measurements", Unisurv S48, 256 + xv pp, 1997, ISBN 0-85839-073-6

S50. M. D. Subari, "Low-cost GPS Systems for Intermediate Surveying and Mapping Accuracy Applications", Unisurv S50, 179 + xiii pp, 1997, ISBN 0-85839-075-2

S51. L.-S. Lin, "Real-Time Estimation of Ionospheric Delay Using GPS Measurements", Unisurv S51, 199 + xix pp, 1997, ISBN 0-7334-1664-0

S53. D. B. Lemon, "The Nature and Management of Positional Relationships within a Local Government Geographic Information System", Unisurv S53, 273 + xvi pp, 1997, ISBN 0-7334-1678-0

S54. C. Ticehurst, "Development of Models for Monitoring the Urban Environment Using Radar Remote Sensing", Unisurv S54, 282 + xix pp, 1998, ISBN 0-7334-1679-9

S55. S. S. Boey, "A Model for Establishing the Legal Traceability of GPS Measurements for Cadastral Surveying in Australia", Unisurv S55, 186 + xi pp, 1999, ISBN 0-7334-0685-X

S56. P. Morgan and M. Pearse, "A First-Order Network for New Zealand", Unisurv S56, 134 + x pp, 1999, ISBN 0-7334-0685-8

S57. P. N. Tiangco, "A Multi-Parameter Radar Approach to Stand Structure and Forest Biomass Estimation", Unisurv S57, 319 + xxii pp, 2000, ISBN 0-7334-0786-2

S58. M. A. Syafi'i, "Object–Relational Database Management Systems (ORDBMS) for Managing Marine Spatial Data: ADCP Data Case Study", Unisurv S58, 123 + ix pp, 2000, ISBN 0-7334-1704-3

S59. X.-Q. Lu, "Strategies for Improving the Determination of Displacements of Sea Surface Temperature Patterns Using Consecutive AVHRR Thermal Images", Unisurv S59, 209 + xiii pp, 2000, ISBN 0-7334-1721-3

S60. G. Dickson, "GPS-Controlled Photography: The Design, Development and Evaluation of an Operational System Utilising Long-Range Kinematic GPS", Unisurv S60, 417 + x pp, 2000, ISBN 0-7334-1725-6

S61. J. Wang, "Modelling and Quality Control for Precise GPS and GLONASS Satellite Positioning", Unisurv S61, 171 + x pp, 2001, ISBN 0-7334-1766-3

S62. Y. Wang, "Knowledge-Based Road Extraction from Aerial Images", Unisurv S62, 178 + xi pp, 2001, ISBN 0-7334-1767-1

S63. L. Ge, "Development and Testing of Augmentations of Continuously-Operating GPS Networks to Improve their Spatial and Temporal Resolution", Unisurv S63, 230 + xvi pp, 2001, ISBN 0-7334-1841-4

S64. H.-Y. Chen, "A Study on Real-Time Medium-Range Carrier-Phase-Based GPS Multiple Reference Stations", Unisurv S64, 182 + xxiv pp, 2001, ISBN 0-7334-1842-2

S65. G. Y. K. Shea, "A Web-Based Approach to the Integration of Diverse Data Sources for GIS", Unisurv S65, 233 + xv pp, 2001, ISBN 0-7334-1845-7

S66. M. Mirbagheri, "Analysis of Interferometric SAR for Topographic Mapping", Unisurv S66, 135 + xvii pp, 2001, ISBN 0-7334-1856-2

S67. P. Wang, "Applying Two-Dimensional Kalman Filtering Techniques to Digital Elevation Models for Terrain Surface Modelling", Unisurv S67, 175 + xi pp, 2001, ISBN 0-7334-1857-0

S69. C. Satirapod, "Improving the GPS Data Processing Algorithm for Precise Static Relative Positioning", Unisurv S69, 131 + viii pp, 2002, ISBN 0-7334-1901-1

S70. R. Mason, "Developing Australian Spatial Data Policies – Existing Practices and Future Strategies", Unisurv S70, 258 + xv pp, 2002, ISBN 0-7334-1941-0

S71. C. Ogaja, "A Framework in Support of Structural Monitoring by Real Time Kinematic GPS and Multisensor Data", Unisurv S71, 191 + xiii pp, 2002, ISBN 0-7334-1958-5

S72. L. Dai, "Augmentation of GPS with GLONASS and Pseudolite Signals for Carrier Phase Based Kinematic Positioning", Unisurv S72, 188 + viii pp, 2002, ISBN 0-7334-1975-5

S73. C. Roberts, "A continuous Low-Cost GPS-Based Volcano Deformation Monitoring System in Indonesia", Unisurv S73, 271 + xvi pp, 2002, ISBN 0-7334-1976-3

S74. V. Janssen, "A Mixed-Mode GPS Network Processing Approach for Volcano Deformation Monitoring", Unisurv S74, 199 + viii pp, 2003, ISBN 0-7334-2059-1

S75. Y. H. Lu, "Automatic Building Extraction for 3D Terrain Reconstruction Using Image Interpretation Techniques", Unisurv S75, 175 + xii pp, 2004, ISBN 0-7334-2141-5