Map-Image Registration Using Automatic Extraction of Features from High Resolution Satellite Images

by

Vijay Kumar Vohra

Thesis submitted for the degree

of

Doctor of Philosophy

of the

University of London

Department of Geomatic Engineering
University College London
Gower Street
London WC1E 6BT
United Kingdom

August 1999

Abstract

In every part of the world the rate of map revision is alarmingly low, when compared to the rate of change of many human influenced surface features. Map making is very time-consuming and often information used for updates has become history before the updated map is made available. There is therefore a requirement to regularly gather up-to-date information about surface features and to incorporate changes in maps both quickly and efficiently. Automation of two systems, i.e. the automation of map-image registration and then of change detection can fulfil the requirements of map revision. This thesis works on the first system.

The piece of work in this study has looked into a fast and an accurate solution to register high resolution satellite images to maps. This will allow changes in ground features to be used to update maps. Photogrammetric techniques used to update maps have previously shown good results, but they are tedious, time-consuming, and not beneficial for updating small changes at all.

Feature extraction methods were used in the present study. The system developed was designed for automatic extraction of suitable areal features in images. The emphasis was on areal features rather than point or linear features because they have a distinctive shape, and they are extracted easily from vector as well as raster data. The extraction of suitable polygons, as control information, from images was obtained by using two matching techniques. Patch matching to extract the conjugate map and image polygons, and dynamic programming to find the corresponding matched boundary pixels of the map and image polygons.

Some matched points were incorrect because of perspective, shadows and occlusions. A statistical model was developed to remove perspective distortion and large errors. The model demostrated the removal of erroneous match points, and selected the good match points and registered the images to maps with a sub-pixel accuracy.

A novel aspect of the study is that the automation is achieved with high accuracy in flat and moderate terrain areas without using height information, as it is essentially used in photogrammetric techniques.

Contents

List of Figures

List of Tables

Acknowledgements

I am feeling difficulty in expressing myself. I think words are seldom adequate and often look depleted when asked to convey the feelings of individuals. In trying to express my deep gratitude to Prof. Ian J. Dowman, for his guidance, by laudatory phrases, injustice would be done to truth, for he deserves much more than those often seen phrases. Infact, he has been 'the' hiden driving force taking me out of the hurdles and problems often arising in this thesis.

The author is grateful to The Natural Environment Research Council (NERC) for providing financial support (studentship) to conduct this study. He is thankful to the Space Department, DERA, Farnborough, UK and The Archangel Project, funded by EC 4th Framework, for providing high resolution satellite images for conducting the research.

My most sincere thanks to Alec Walker, Paul Dare, Masood Varshosaz and Diego Jimenez for stimulating discussions and valuable suggestions, and thanks to the staff members of the Department of Geomatic Engineering, UCL for friendly support and encouragement.

Very special thanks go to my parents, sister and her family for their constant care, support, and countless love.

Last, but not least, the author is indebted to all those who have directly or indirectly helped the author in his hour of need.

Chapter 1

Introduction

1.1 Motivation

The status of world topographic maps in the categories specified by the United Nations is good in developed countries and poor in most third-world countries. Moreover, the rate of map revision in every part of the world has remained alarmingly low (Konecny, 1992). It is well known that without an adequate map of an area, that is, a revised map or a current map of the area, plans for infrastructure development, resource exploration and exploitation cannot be efficiently executed. There is therefore an important need for an automated map updating system which could satisfy the user's requirements of updating their maps quickly, accurately and with cost savings. Advances in technology for converting analogue data to digital, and the availability of high resolution satellite images motivates us to analyse the problems and solutions related to automatic map revision.

1.1.1 Film to digital information

Since the First World War, images of the earth taken from aircraft have been widely used for defining the geographical framework. Stereoscopic use of such photographs formed the basis of much three-dimensional mapping and the drawing of contour lines, a laborious and time consuming exercise. Alternatives have been sought for a long time.

Observation of the whole earth or large sections of it first attracted public attention with the US Apollo space programme of the late 1960s and early 1970s. But the expense and difficulties associated with recovering film data from space led the technologists to process the film on board and then scan and transmit the image back to earth. This inspired them to develop digital image sensors, and now almost all satellites collect image data, store it and send it back to earth in digital form by telemetry. The image is composed of the amount of radiation reflected or emitted by a small area of the earth's surface and the size of the area is represented by a pixel (a picture element) and defines the basic resolution or detail of the system. Several pixels are generally required to recognize any 'real world' feature.

The first automatic imaging of the earth covering a wide-area, and available to civilians, came in 1972 with the launch of Landsat-1 satellite by the USA. This supported a multi-spectral scanner of 80 m ground pixel size offering four image bands and a continuous digital image swath of 185 km. Landsat's resolution improved considerably in 1982 when the Thematic Mapper scanner instrument was included in the Landsat-4 mission; it has seven spectral bands with all at 29 m resolution except the thermal band which has a 120 m resolution. Before the Thematic Mapper scanner, many satellites had delivered numerous images but none had been of sufficient spatial resolution to provide a suitable basis for mapping or for constructing and updating the geographical framework.

The digital domain is highly suitable for integrating and exploiting the benefits of digital data acquired from satellites. It offers a more direct interface with neighbouring disciplines such as digital photogrammetry, geographical information systems and image processing, resulting in improved performance and an extended range of applications. Nowadays, many applications for automation rely on digital information and this is currently an active research area.

1.1.2 Achievement of high resolution images

European technology advanced fast in the early 1980s - the French and their partners were the first to position a civilian 10 m digital sensor in space in 1986 with

the SPOT-1 satellite. In Britain, the Ordnance Survey researched the potential of satellite imagery for topographic mapping for many years (Hartley, 1991) but did not obtain acceptable results until spatial resolution improved with the launch of SPOT-1. After the launch of SPOT-1, the Ordnance Survey claims that it undertook the world's first extensive topographic mapping project in the civilian sector, using stereo SPOT hardcopy imagery to map over 25,000 square kilometes of north Yemen at scale 1:100,000 (Murray and Farrow, 1988).

In contrast to the digitally based developments in the Western world, a variety of photographic imaging systems was developed and deployed in 1970s and 1980s by Russia as military initiatives. Russia has recently, since the end of the Cold War, started marketing this photographic imagery and digitally scanned versions of it with 2m ground resolution, each image covering an area of 40 km by 40 km. The release of military technology into the civilian domain has led to the creation of a number of commerial organizations in the imaging field with global out-reach and ambitions (Fritz 1996). The most significant step forward for mapping and geospatial data gathering is the move to 3 m, 2 m, and 1 m digital imaging resolution. Such higher-resolution imagery has a number of implications. A recent UK Ordnance Survey assessment (Ridley et al, 1997) showed that while it may be possible to detect features in the 1 m spatial resolution imagery, precise interpretation and accurate locational positioning are severly limited for the core mapping activities of the large scale survey maps. The assessment also mentioned that a spatial resolution of 0.2 m or better would be sufficient to support the applications required to maintain the existing two-dimensional (2D) database of the UK which has 229,000 map tiles modelling urban areas at a scale of 1:1,250 for towns, 1:2,500 for rural areas, and 1:10,000 for mountain and moorland areas. Doyle (1982) claims that, for mapping, the relationship of image to map scale is a key issue, and it would be more realistic to assume that mapping scales in the order of 1:5000-1:15,000 could be possible from 1m space imagery in the future. The main satellite imaging systems relevant to mapping and maintaining an up-to-date geographical framework are set out in Table 1.1.

Table 1.1: Characteristics of imaging sensors and their relevance to mapping

Sensor	Launch date	Altitude (km)	Stereo capability	Pixel size	Image extent (km)	Temporal resolution	Relevance to mapping
Landsat 4 Landsat 5 Landsat 7	1982 1984 1999	705	No	29 m	185 (sq.)	16 day repeat	small-scale and simple topographic area
KFA-1000	1974	220-350	Yes (along track)	5-7 m	80 (sq.)	Mission	1:25,000 scale mapping; height accuracy weak
KVR-1000	1984	200-350	No	2 m	40 x 160	Mission	1:10,000 and smaller-scale
SPOT 1,2 SPOT 3,4	86,90 93,98	832	Yes (cross track)	10 m	60 (sq.)	26 day repeat	1:50.000 and smaller-scale
IRS-1C IRC-1D	1995 1997	817	Yes (cross track)	5.8 m	70 (sq.)	24 day repeat	Possibly 1:25,000 and smaller-scale
Earlybird	1997 (not operated)	470	Yes (along and cross track)	3 m	6 (sq.)	2-3 day revisit	Small scale 1:25,000 and smaller?
Space Imaging (Ikonos)	1999	682	Yes (along and cross track)	1 m	11 (sq.)	1-3 day revisit	Medium scale 1:10,000 and smaller?
Quickbird	1999 planned	470	Yes (along and cross track)	1 m	30 (sq.)	2-3 day revisit	Medium scale 1:10,000 and smaller?
Orbview 3 Orbview 4	1999 planned 2000	470	Yes (along and cross track)	1 m	64 (sq.)	< 3 day revisit	Medium scale 1:10,000 and smaller?
SPOT 5	2002	832	Yes (cross track)	3 m and 5 m	60 (sq.)	26 days revisit	1:25,000 and smaller-scale

Updating the framework in many of the more advanced countries has been by means of editing the database representation of only those features which change in the real world. Data collection systems and local intelligence, for example through planning systems, have ensured that this is a highly economical way to proceed. Rates of change of the features in the framework database rarely vary by more than 1 or 2 percent per annum, yet detecting these important changes is vital. This implies that repeatedly collecting images is of little value unless change can be automatically extracted and stored.

1.2 Background and objective

1.2.1 Background

For automatic map updating of a specific scale and a required resolution, its imagery has to go through two essential processes:

- automatic registration of image to map; and
- automatic change detection.

This can be simply shown by a diagram. Figure 1.1 clearly indicates two automatic systems to be developed for the automation of the map revision task. One is

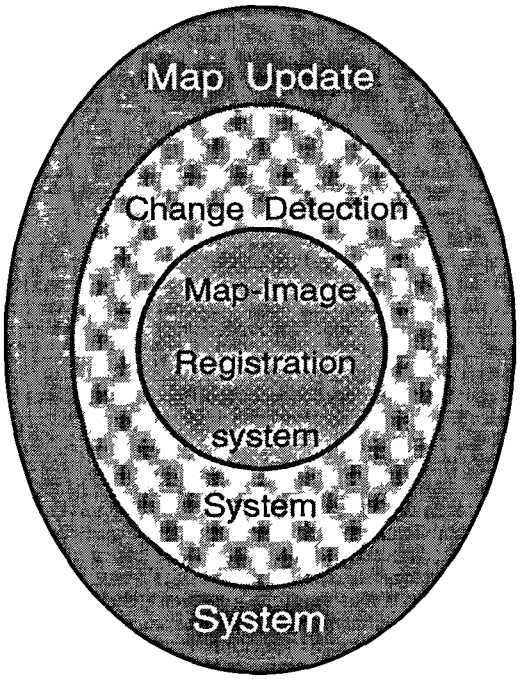


Figure 1.1: Components of map update system.

an automatic map-image registration system and the other is an automatic change detection system. The core of this system is the automatic map-image registration system. Its development is necessary for the development of the next system, that is, the automatic change detection system. The studies undertaken in this research are concerned with the core of the map updating system, that is, automation of the map-image registration system.

At present, the problems of automatic registration (absolute orientation) of an image to a map have not yet been solved in a flexible and general way, but considerable progress has been achieved in the automation of image orientation for photogrammetry and remote sensing over the last few years. Some work on automation of absolute orientation from aerial photography has been done and this has recently been comprehensively reviewed by Heipke (1997). Heipke notes that software packages are commercially available for automatic interior and relative orientation and automatic aerial triangulation (referring to Braun et al., 1996; Lue, 1996; Ackermann, 1996; Madani, 1996); but the situation is somewhat different for automatic absolute orientation. Encouraging research results and developments towards automatic absolute orientation exist for special cases, but a general solution is not available at present. The main problem lies with the automatic extraction of control information from the images. The solution to this problem will be tackled in this research. Automation would be of great help in eliminating the tedious and time-consuming task of identifying control information. The increasing need for satellite information for many applications related to updating maps, in a short time period, would also be met by automation.

In the literature, some researchers mention the automatic measurement of the orientation points used by absolute orientation, but none has achieved desirable results for fully automating the registration of images to maps. The problem is still under investigation. Morris et al. (1988) and Stevens et al. (1988) describe some techniques for finding the automation of map-image registration. Maitre and Wu (1989) show in their work how to register an image to a map using a combination of autoregressive modelling of the deformation and a dynamic programming algorithm. Schickler (1992) uses line segments for feature matching and Haala and Vosselman (1992) use relational matching showing good results. A technique described by Medioni and Nevatia (1984) uses line-based descriptors, and matches them by computing the most similar geometrical structures. Morgado and Dowman (1997) report a procedure for automatic absolute orientation of aerial photographs to a map by extraction of areal features. Their method shows improvements but needs a technique for the removal of high error match points, which is a critical

problem. A solution to the problem of the filtering of errors is included in this research.

1.2.2 Aims

In this study automatic extraction of areal control features from high resolution satellite images is investigated for the development of precise automatic map-image registration system.

1.2.3 Methodology

To solve the goal of auto-map-image registration, the general components of the method involved in this research are shown in Figure 1.2 and are discussed in detail in the following chapters.

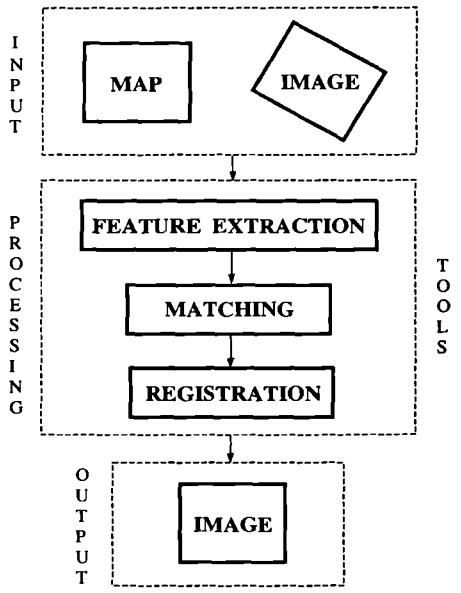


Figure 1.2: An outline of an automatic image to map registration system.

Automation of the registration processes is based on methods of feature extraction. In this study, areal features were considered in particular, because they have a distinctive shape and can be found in urban and non-urban areas. They can be extracted from images and digital maps automatically, and more easily than point and linear features. They also provide a large amount of redundancy.

Maps and images were brought into same representation for matching in this work, which needed preparation of both data sources. The preparation of map and image data was done to extract areal features from the map and the image to match their boundaries. Map preparation was done on the Arc/Info system and the corresponding image was prepared using several processes: smoothing, segmentation (region growing technique), clutter removing, patch matching, and integration of iterative region growing with patch matching. Previous workers [Abbasi-Dezfouli et al, 1994; Morgado and Dowman, 1997; Vohra and Dowman, 1997; Dowman and Ruskone, 1997] have used region growing and patch matching techniques to extract areal features in images, but they faced difficulties in obtaining all the matches correctly in patch matching and that hindered the automation of patch matching. In this work the integration of iterative region growing with patch matching was used. The integration process showed correct matches and extracted the areal control features and automated the method of extracting areal features without the user's interaction. Matching was done between the extracted boundaries of areal features of the map and image and conjugate matched points were obtained. Some match points were incorrect because of perspective distortion, shadows and occlusions. A statistical model developed in this work was used on match points and removed erroneous match points and obtained planimetric points. The filtering of erroneous matches showed high accuracy of map-image registration.

The system described in this work is designed for the automatic extraction of suitable polygonal features from images; automatically matching images to the corresponding polygons in the map of the same area, without any manual control points. Moreover, an additional automatic statistical analysis module is attached to the system to remove perspective distortion and large errors, in order to improve image registration accuracy and to validate the end results.

1.3 Thesis plan

This thesis consists of six chapters. It focuses on specific aspects of high resolution satellite information for automatic extraction of significant polygonal features in

images and the development of an automatic map-image registration system.

- Chapter 1 includes the introduction to the present research work. It mainly emphasises the aim of the research and the structure of the thesis.

- Chapter 2 describes the basic requirement for feature extraction, types of features and the techniques for feature extraction from images. This begins with the different levels of data representation in images, and describes how the role of image processing plays an important part in extracting suitable types of features. The chapter further deals with how the areal features can be represented by useful information such as chain code, area, and length of perimeter as intrinsic attributes to give a unique shape description value to each extracted areal feature in the image. At appropriate places, references to previous work are inserted as well as references to other chapters within this thesis.

- The preparation of map and image data for matching, and the matching process are described in detail in Chapter 3. The map is considered as reference data, and the preparation of a model from the map by semi-automatically tracing the boundaries of desirable objects on the user friendly Arc/Info geographic information system are shown. A shape descriptor technique was used for calculating the descriptive value for each traced object. For image data preparation, the techniques of smoothing, segmentation, and boundary detection and the shape parameters of segmented objects are mentioned. A region growing segmentation technique was used iteratively, with an increase of the gray level difference threshold between neighbouring pixels to grow the regions (in size) in the image by merging neighbouring pixels. Each iterative process was combined with a technique to extract a few desirable segmented objects in the image for symbolic representation and to match them to the prepared map model to extract the best shape of image polygons (control features). This is a new approach developed in this project to achieve the automatic extraction of control features in images using a map model. The approach was integrated with a combination of two well known matching

techniques to obtain correspondence of map and image at pixel level.

- Chapter 4 reports on image registration. Transformation parameters of matched pixels create misregistration of image to map when perspective distortion occurs due to tall objects and large terrain variations. A model was developed which describes the removal of distortion to obtain better transformation parameters for precise map-image registration.

- Automation of the image map registration system is covered in detail in Chapter 5. The use of a new data set of image and map to show the end to end result is described. A proper linkage of various modules developed in this research work discusses the achievement of automation, with the minimum interaction of the user, which is assumed to be an input to the system.

- The concluding chapter mentions the automatic extraction of significant image features and the automation of the map-image registration system developed. This chapter also mentions the novel aspect of the research work carried out, the limitations of the developed system and suggestions for future work.

1.4 Summary

A new approach for automatic registration of images to maps is described here. Its notable features are the achievement of automatic extraction of control features from images and precise map-image registration without using height information of objects. The extraction of significant features (using different landcover) from high resolution satellite images, matching them to their models, the elimination of high erroneous matches by a statistical model developed to select planimetric match points (control features) and the use of the points for precise map-image registration, is (all) achieved in this piece of research work. Only a small amount of user interaction is required to initialise the present system.

Chapter 2

Feature Extraction and Object Shape Descriptors

2.1 Introduction

It is well known that digital images are composed of pixels and each pixel corresponds to some object or group of objects in the real world. In general, many pixels may be required, depending upon how much area is represented by a pixel, to identify any 'real world' object feature in an image. It is not as easy to extract objects from images using computer algorithms as the human vision system can. The computer has to go through a great deal of processing before it can classify objects such as point, edge or region classes, and extract the shape of the objects in an image.

Feature extraction has been a very active area of research for the disciplines of digital photogrammetry and remote sensing for the last decade. It plays an important role in the extraction of meaningful structure from digital images. The structures extracted from the images can be useful for many different tasks, like automation of the relative orientation process (Haala and Novak, 1994 and Hellwich et al, 1994), automatic registration of images with maps (Lee et al, 1993; Dowman et al, 1996; Vohra and Dowman, 1996), and automatic detection of roads (Shahin et al, 1994). It is important to understand the basic processes that help extract the

meaningful structures from the images. What are these processes? This chapter gives a general overview of these processes. It includes related theories and builds a bridge between previous research work and research carried out in this piece of work, which is described in subsequent chapters.

2.2 Computer vision system

Computer vision is a science which has been fast-growing for the last two decades. It is the enterprise of automating and integrating a wide range of processes and representations used for visual perception. More importantly it includes techniques for geometric modelling and cognitive processing. In another words, it can be said to deal with the construction of explicit, meaningful descriptions of physical objects in images. Descriptions are a prerequisite for recognizing, manipulating, and thinking about objects.

2.2.1 Levels of image data processing

The aim of computer visual perception is to find a relation between an input image and models created of the real world. Broadly, computer vision systems can be divided into low-level and high-level processing systems. Computer vision is concerned with both the low-level (early processing) issues and with the high-level, i.e. "cognitive", use of knowledge. Vision requires many low-level capabilities that we often take for granted, for example, our ability to extract intrinsic images of *lightness, colour, and range.* Low-level processing capabilities are elusive as well as unconsious. They are not well connected to other systems that allow direct interpretation. For instance, our visual memory is quite impressive, but the quantitative verbal descriptions of images are relatively primitive in comparison to computer systems.

Knowledge is a high-level capability that can guide visual activities to achieve goals, and a visual system should be able to take advantage of it. Many tasks require judgment and large amounts of knowledge of objects in the world, how they look, and how they behave. Such high-level powers are so well integrated into

our vision system as to be effectively inseparable.

Where does vision leave off and reasoning and motivation begin? It is not known precisely, but it is believed that powerful, cooperative, rich representations of the world are needed for any advanced vision system. Without them, no system can derive relevant and invariant information from an input that is beset with ever-changing lighting and viewpoints, unimportant shape differences, noise, and other large but irrelevant variations.

Low-level image processing and high-level computer vision differ in terms of the data which is used. Original images represented by matrices composed of brightness values comprise low-level data, while high-level data originates in images as well, but only those data which are relevent to high-level goals are extracted, reducing the data content - for example object size, shape, and mutual relations between objects in the image. High-level data are usually expressed in symbolic form.

2.2.2 A range of data representations

According to Ballard and Brown (1982), image data representations can be stratified in four levels; however, there are no strict borders between them. These representational levels are ordered from signals at a low level of abstraction to the description that a human can perceive, and the information flow between them may be bi-directional. For some applications, a more detailed classification of the representational levels can be used and, for some specific uses, some representations can be omitted. These data representation levels are mentioned below.

- The lowest representational level comprises iconic images consisting of original data, that is, matrices with data about pixel brightness. Images of this kind are also outputs of pre-processing operations (e.g. filtration or edge sharpening) used for highlighting some aspects of the image important for further treatment.

- The second level of representation is segmented images. Parts of the image are joined into groups that probably correspond to the same object. It is useful to know something about the application domain while doing image

segmentation; it is then easier to deal with noise and other problems associated with erroneous image data.

- The third representational level is geometric representations holding knowledge about 2D and 3D shapes. The quantification of a shape is very difficult but also very important. Geometric representations are useful for doing general and complex simulations of the influence of illumination and motion in real objects.

- The fourth level of representation of image data is relational models. They give us the ability to treat data more efficiently and at a higher level of abstraction. *A priori* knowledge about the case being solved is usually used in processing of this kind.

Visual perception is the relation of visual input to previously existing models of the world. There is a large representational gap between the image and the models which explain, describe, or abstract the image information. To bridge that gap, computer vision systems usually have a loosely ordered range of represention between the input and the output to give a final description, decision, or interpretation of a scene. Computer vision involves the design of these intermediate representations and the implementation of algorithms to construct them and relate them to one another. These representations can remove some computational load by predicting or assuming structure, for example the grouping of image pixels of similar brightness into separate objects. Thus, the topic of extraction of features in the images comes into view, and it plays an important role in the extraction of meaningful structure in digital images.

2.3 Extraction of Features in Images

In digital images, extracted features can be points, linear or areal features. The point features are normally extracted using so called "interest operators" such as Moravec (1980) or Förstner (Förstner and Gülch, 1987) operators. Linear features

can be detected using a first or second order derivative operator. The most common segmentation technique for extraction of regions is region growing.

Low-level computer vision techniques overlap almost completely with digital image processing, which has been practiced for decades. For extracting features from digital images, the image processing technique of contrast stretching (grey level scale transformation), smoothing (averaging of neighbouring pixels) and gradient operation can be applied. Important features of these techniques are mentioned in the following subsections.

2.3.1 Contrast Stretching

The digital revolution has changed everything related to the darkroom of old, which has been replaced by the increasingly ubiquitous personal computer. Everyone can now use image-processing techniques, like histogram manipulation, to produce high quality results. The primary function of the display is to allow the human observer to understand and interpret the content of the image. The human eye has considerable acuity in discriminating fine detail (high-spatial-frequency information), although it is not particularly sensitive to low-frequency (slowly varying) information in the image. In some cases, then, it is helpful to match the display process to the characteristics of the human eye, which can be done by stretching the contrast of the distribution of the grey levels of the image. Some images may be more easily understood if they are displayed indirectly by using contour lines, shading, colour, or some other graphical representation.

Our eyes enable us to perceive an incredible range of brightness, from dim moonlight to full sunlight, and cameras and sensors can perform in the same way. But television screens or computer monitors have limitations in displaying all the brightness levels contained in an image; only a small range of pixel values of an image can be properly displayed at one time. Stretching is the most basic and important image-processing function. Simple stretching techniques can dramatically improve an image, which is difficult or impossible with conventional darkroom technology. Selecting different ranges of brightness to be stretched over the mon-

itor's optimum display capabilities can reveal different features in an image. Any stretching operation can be quickly visualized by looking at its effect on the histogram. Taking the stretching process to an extreme, we can force the histogram to have any shape.

- **Histogram:** A histogram is usually the only global information which is available about the image, and it is an important tool for image analysis. The histogram is a simple tabulation (usually displayed as a graph), giving the number of pixels of each brightness value in an image. It is used when finding optimal illumination conditions for capturing an image, grey scale transformations, and image segmentation of objects and background. Many images do not have a full brightness range and therefore cover a small portion of the display range available. These are low contrast images. Linear expansion of the histogram, of the low contrast images, can be applied to cover the full range of display. Expanding the brightness scale by spreading the histogram out to the full available range may improve the visibility of features and the perceived contrast in local structures. The histogram of a digital image typically has many local minima and maxima that may complicate its further processing. This problem can be avoided by local smoothing of the histogram, for which local averaging of neighbouring histogram elements is the base.

- **Histogram equalization:** A grey scale transformation for contrast enhancement is usually found automatically using the histogram equalization technique. Histogram equalization produces an equal number of pixels of each level of brightness in the image, and it makes the histogram flat. It enhances contrast for brightness values close to histogram maxima, and decreases contrast near minima. While this helps to display the full range of brightness at any one time, the image often appears unnatural and unattractive to the eye. The logarithmic grey scale transformation function is another frequently used technique. It simulates the logarithmic sensitivity of the human eye to the light intensity.

There are, of course, operations more complex than simple subtraction and multiplication which can be performed when stretching an image but they are beyond the scope of this thesis. Books written on image processing by Gonzalez and Wintz (1987) and Russ (1995) are recommended for advanced understanding of histogram manipulation.

2.3.2 Edge Preserve Smoothing

Smoothing is a process to suppress noise or other small fluctuations in the image that were incurred during image acquisition. It is obtained by using a filter (window) which passes over the image. A new value is calculated for each central pixel of the window on the basis of averaging its own and its neighbour's brightness values. In practice, convolution can be used to implement filters for smoothing or iconic feature extraction. The convolution consists of multiplications and additions, and it behaves like matrix multiplication except that it is commutative. Due to these properties, convolutions can be easily concatenated with one another or, inversely, convolutions can be separated. For example, if two one-dimensional convolution kernels (masks) as shown below:

$$R_r = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

and

$$R_c = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

are concatenated, then the two-dimensional convolution kernel will look like:

$$R = R_r * R_c = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

This convolution kernel is a linear filter which can be used for smoothing an image to reduce noise and it computes the mean of a 3 x 3 neighbourhood. Binomial

filters are discrete approximations of continuous Gaussian filters. The binomial filter of a 3 x 3 neighbourhood can be given as:

$$B = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Linear and binomial filters are not adequate for smoothing images, because they not only remove the noise but lead to the blurring of edges and therefore to a loss of information. The median filter is another type of filter which is a good non-linear filter and it smooths images and maintains edges but tends to round off corners. In order to overcome these deficiencies, a variety of filters have been proposed. Matsuyama and Nagao (1980) suggest a method of averaging, using a rotating mask, that avoids edge blurring by searching for the homogeneous part of the current pixel neighhourhood. This edge-preserve smoothing algorithm uses nine masks around a central pixel, as shown in Figure 2.1.

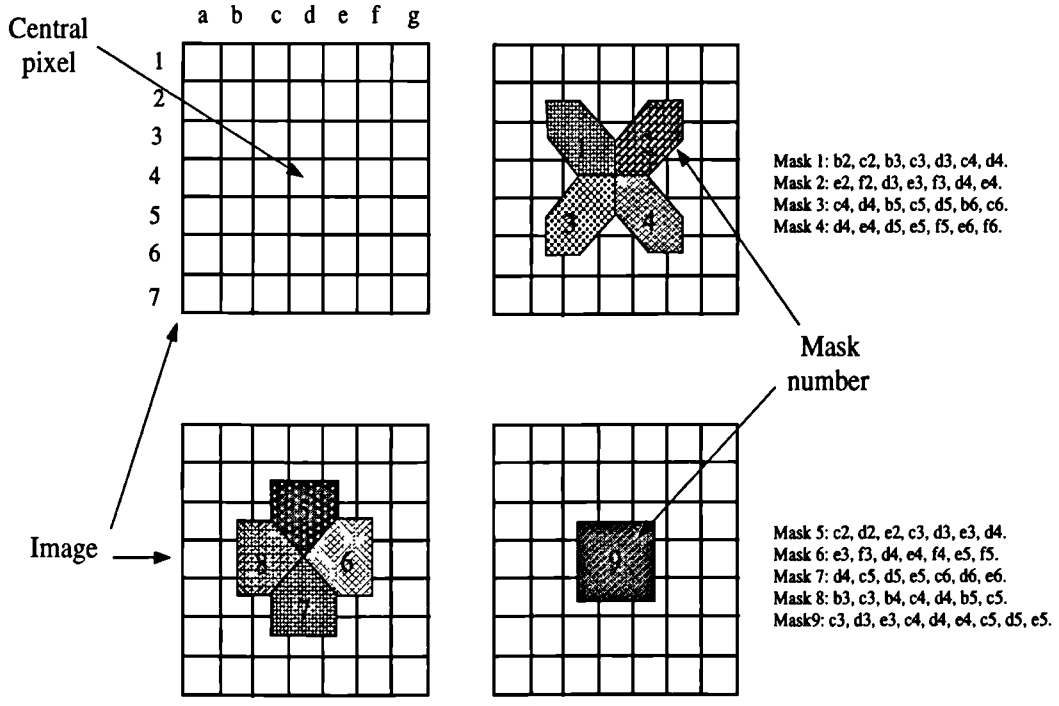


Figure 2.1: The nine masks of the edge-preserve smoothing routine.

This edge-preserving filter is applied prior to edge detection to strengthen the grey level discontinuities between different land cover types, and to reduce the

detection of edges in areas of texture that are internal to regions. A window with nine masks (filters), which is shown in the above figure, is passed over the image and the variance is measured in 9 orientations (masks) around the central pixel. The orientation of minimum pixel variance is determined, and its mean is given to the central pixel. The selected orientation should never lie across an edge. This is performed for each pixel in the image. This algorithm was used by Newton et al. (1994) on Thematic Mapper imagery to preserve grey level discontinuities and to reduce the detection of noise edges but enhance forest edges for the subsequent edge detection step. It was also applied by Vohra and Dowman (1996) to smooth high resolution satellite imagery and to preserve edges for the purpose of detecting the edges of large buildings in the image, as discussed later.

2.3.3 Gradient Operations

The homogeneity of the regions can be quantified, either based on derivatives of the grey value function or on texture criteria. In this study, criteria based on the grey values is considered. It is well known that the gradient operators are based on local derivatives of the image function. Derivatives are bigger at locations of the image where the image function undergoes rapid changes, and the aim of gradient operations is to indicate such locations in the image. The computation of gradients is done by processing finite differences of the grey values. Many different operators have been proposed for this purpose, and all can be represented in a simple form, as shown in equation 2.1

$$D_s = S * D \tag{2.1}$$

Where S is a smoothing and D is a differencing operator. The smallest differencing operators D for the computation of gradients are:

$$\partial_r^s = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and

$$\partial_c^s = \begin{bmatrix} 1 & -1 \end{bmatrix}$$

D_s is the combined effect of smoothing and differencing operators, which reduces the noise in the image and finds points, edges or regions by using different types of differencing operators. The exact composition of each convolution kernel is as

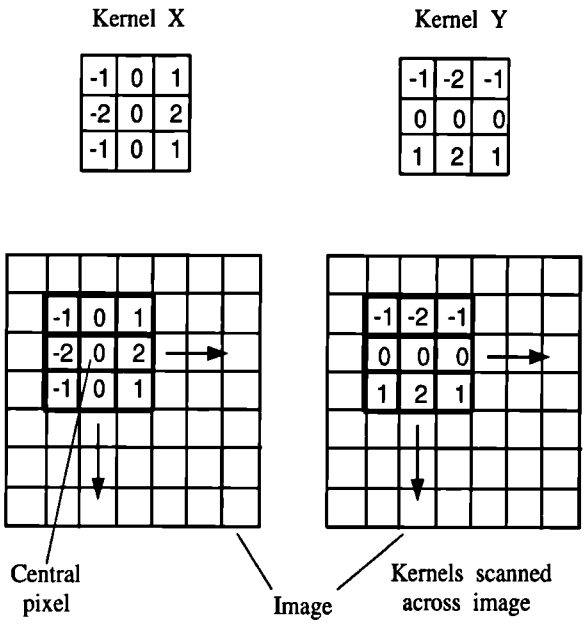


Figure 2.2: Edge enhancement by convolution of kernels with the image.

varied as the user's imagination, although the Sobel and Laplacian are particularly widely used (Mather, 1987). Two kernels shown in Figure 2.2 are used for the Sobel Operator. When these two kernels are passed across the image, for each pixel convolved with the image the program finds the strength of the edge at the central pixel of the kernel and its direction gradient. An edge is a property attached to an individual pixel and it is a vector variable that has two components, the magnitude of gradient (edge strength) and the direction. These two components, magnitude and direction, using the sobel operator, are given in the following equations 2.2 and 2.3 respectively.

$$\text{Strength} = \sqrt{[(\text{Result of convolving kernel X})^2 + (\text{Result of convolving kernel Y})^2]} \quad (2.2)$$

$$\text{Direction} = \tan^{-1}[(\text{Result of convolving kernel X})/(\text{Result of convolving kernel Y})] \quad (2.3)$$

There are several gradient operators used for edge detection, and the variations can be obtained by rotating their kernel's values. Here, it is not possible to discuss

the various operators and show their kernels because it is not relevant to this research work. In the next section, some of the important well-known operators will be mentioned with regard to their applications. The next section also gives an overview of the basic segmentation techniques for extracting points, edges and regions, and mentions the different operators (spatial filters) used for its extraction.

2.4 Feature Extraction for Segmentation

Image segmentation is often understood as deriving a symbolic image description: "primal sketch" according to Marr (1982) and "first primitive description" as Fischler and Bolles (1985) have mentioned. It may not necessarily be complete, but it does serve a special purpose by extracting meaningful structures from digital images, that is, extraction of geometric and structural features. This leads to a large number of approaches for detecting the basic features - points, edges and regions - which are mentioned in the following subsections.

2.4.1 Extraction of Points

Points are image objects, whose geometric properties can be represented by only two coordinates (x,y) and the symbolic description of points is given as a list containing geometric, radiometric and relational attributes. There are several types of points:

- Circular symmetric points
- Endpoints
- Corners
- Junctions

Hilltops, hollows, peaks and holes are circular symmetric points which are local inhomogenities in the interior of homogeneous image regions. Extracted circular symmetric points can be interpreted as region attributes (depending on the scale of the image) and are independent of the image structure. Endpoints are start

or end points of a line structure extracted from the image, and corners belong to the intersection of exactly two lines in the image. Intersections of more than two lines are junctions and are used for the geometric description of edges and region boundaries. In the literature, two point operators are frequently described when the topic of extraction of point features is dealt with, and they are mentioned below.

1. *Moravec Operator:* Moravec (1977) was first to propose an approach aimed at detecting points which can be easily identified and matched in stereo pairs of aerial photographs. He suggested measuring the suitability or "how interesting" an image point is by using a small window (5 x 5 or 9 x 9) around it, and calculating the variance along a direction (horizontal, vertical, or the two diagonals) as the average squared difference of neighbours. The minimum of these four variances is selected as the value of the operator. If the window variance exceeds a certain threshold, the pixel is chosen as an interest point. The operator identifies those places in an image where there is a high gradation in intensity level in all directions. However, it is time consuming if the calculations are performed for each pixel in the image.

2. *Förstner Operator:* Förstner and Gulch (1987) proposed another interest operator which can detect distinct points, like the corners and centres of circular features in an image. For each point in an image, an interest window (usually of size 5 x 5 or 7 x 7 pixels) is applied to form the following matrix N:

$$N = \begin{bmatrix} \sum g_x^2 & \sum g_x g_y \\ \sum g_x g_y & \sum g_y^2 \end{bmatrix} \tag{2.4}$$

where g_x and g_y are the first partial derivatives of the image function g(x,y). A corner (x_o,y_o) can be considered as the intersection of all the edge elements inside a window, using the error ellipse represented by the above mentioned matrix. The size (w) and the roundness (q) of the error ellipse are determined from the determinent and trace of the normal equation matrix N as follows:

$$w = \frac{trace(N)}{det(N)}$$

$$q = \frac{4\,det(N)}{trace^2(N)}$$

These values are compared with preliminary weights for w and q. The points whose w and q are larger than certain limits are selected as interesting points. The error ellipse should be small and normally as close as possible to a circle. The results of this operator are distinct points such as corners, the intersection of two or more edges, and centres of circular holes, disks or rings.

More information about interest operators can be found in the literature in Hannah (1980); Paderes et al., (1984); Förstner (1986); Deriche and Giraudon (1993). From the literature it can be concluded that the interest operator consists of two basic processing steps:

- selection of optimal windows, and
- accurate point location within these windows by least square estimation.

2.4.2 Extraction of Edges

An edge is an elongated image feature, where a certain image property, for example, brightness, depth, colour or texture, is changing rapidly perpendicular to the edge, and it is also assumed that on each side of the edge the adjacent regions are homogeneous in this property, as mentioned by Ballard and Brown (1982), Rosenfeld and Kak (1982), Kasturi and Jain (1991), Weidner (1995) and Forstner (1995). According to these characteristics, in general, edges can be classified into two types, step edges (edges) and bar edges (lines), as shown in Figure 2.3.

An edge, despite a boundary of a region or the intersection line of two regions, has no full extent in object space. It may occur as a small edge area due to image noise, smoothing or discrete image raster. This normally causes edge extraction to give incomplete segmentation of images, that is, edges do not build closed boundaries of the homogeneous image regions. On the other hand, lines either occur at a discontinuity in orientation of surfaces, or they really exist, but are small elongated objects like streets. Depending on the image scale, lines may be recognised as two-dimensional regions or, in small scale images, as one-dimensional boundaries.

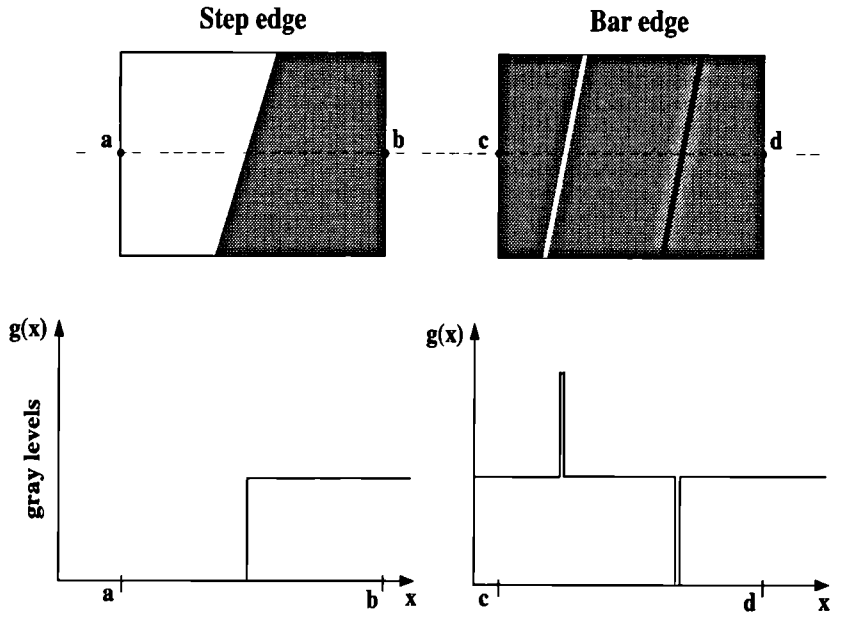


Figure 2.3: General types of edge representation in images.

Edges belong to some kind of boundary and are normally defined by an intensity discontinuity in the image (Lemmens, 1996). However, this can be due to various physical events in the scene, such as discontinuities in surface-normal, in surface reflectance, in depth and illumination (Nalwa, 1993). This leads to the problem that not all boundaries correspond to relevant object outlines and wrong boundaries can be detected and, therefore, the feature extraction of edges has been an active research topic for many years. A most interesting definition of edge detection is given by Attneave (1954):

> "One purpose of edge detection in an image is to strip away some of the redundancy of sensing to encode and describe the information contained in the image in a form more economical than that in which the information impinges on the sensor."

In general, edge detection is obtained by performing smoothing, local edge detection, thresholding, thinning and edge linking processing steps. Local edge detection is based on some form of differentiation of the local grey level value function. Since differentiation is a mildly ill-posed problem (Torre and Poggio, 1986) smoothing is often applied beforehand for regularization purposes. Thresholding

is a decision process in which the label edge or non-edge is assigned to each pixel, based on the response of the local edge detector. The response is usually tested against one or more pre-specified thresholds. These thresholds may be determined on a heuristic basis or by a quantification of image disturbances such as noise. Due to the spatial extent of local edge operators, the intial edge map is generally not one pixel thick. Thinning is required to obtain one pixel thick outlining of object boundaries, which can be obtained by using, after thresholding, a skeletonizing algorithm to erode the thick edges. To obtain higher localization precision one may use, before thresholding, non-maximum suppression to exclude pixels located perpendicular to its gradient direction. The disadvantage is that junction pixels may be deleted too. Finally, the edge pixels are linked to form a boundary of connected pixels.

Several edge operators have been implemented during the last two decades in an attempt to create the "optimal" edge detector. In spite of all the efforts, various factors have affected the realisation of this dream, mainly the trade-off between accuracy in detection and localisation of the edges and also the over-simplistic assumption that surfaces are imaged with uniform brightness (Horn, 1986). According to Gülch (1991), classical edge feature extraction is performed with low level image processing methods, and some of the most representative edge detectors are listed below:

- Mask search (Sobel, Kirsch, Prewitt,...)
- Gradient methods (Canny operator)
- Zero Crossings (Marr, Grimson,...)
- Edge fitting (Hueckel)

It is well known that derivatives and their discrete approximations emphasise noise. The higher the order of derivative, the more pronounced the effect (Nalwa, 1993). An operator such as the Marr-Hildreth operator, which computes the second derivative across and along the edge, is more prone to enhance noise than the Canny's, which computes the second derivative only across the edge. The Sobel

operator has been shown to be superior to other small support operators, like the 3 x 3 Prewitt operator. Lyvers and Mitchell (1988) showed that the magnitude error of the Sobel operator may reach 7.93% and the angle error 2.90 degrees for ideal step edges, while the magnitude error of the Prewitt operator may even reach 12.87% and the angle error 7.43 degrees. Many authors (Brooks, 1978; Marr & Hildreth, 1980; Haralick & Waston, 1981; Hildreth, 1983; Canny, 1986; Nalwa & Binford, 1986; and Torre & Poggio, 1986) have attempted to put edge detection in a more rigorous mathematical framework, but no coherent theory has been developed. No general algorithm exists which can be applied successfully to all types of images. Lemmens (1996) mentions that one of the main reasons for failing to detect edges may be that local edge detectors can not discriminate among the many types of features that may be present in the image.

2.4.3 Extraction of Regions

Segmentation aims to divide an image into parts that have a strong correlation with objects of the real world contained in the image and so it splits up an image into regions. Regions are image areas which fulfill certain similarity criteria, hold properties distinct from their neighbours, and are generally approached from two points of view: by the detection of edges that seperate regions, as mentioned in the last subsection, or by the extraction of regions directly. Both being the opposite sides of the same coin, it would be expected that each approach would yield the same results. But Low (1991) claims that this is rarely true, and a choice between the two must be made.

Region extraction can be simple: Low (1991) outlines some of the most basic algorithms, based on grey level thresholding. Intuitively, neighbouring pixels are more likely to be part of the same region than two pixels far apart. Region growing, which is a sub-branch of region extraction, seeks to incorporate this heuristic. The classic approach to region growing is to pre-define a grey level difference threshold that will define if neighbouring pixels are members of the same region or not. Starting at a pre-designated pixel, all neighbours are considered and, depending

on the grey level difference with respect to the pre-defined threshold grey value chosen, they are captured by the region. The definition of this threshold is vital, as the resulting regions depend solely upon it. When a region can grow no more, a new region is started (Kai and Muller, 1991).

An image is divided into separate regions that are homogeneous with respect to a chosen property such as brightness, colour, reflectivity, context, etc. The brightness of a pixel is a very simple property which can be used to find objects in images. If an image of a complex scene is processed, for example an aerial photograph of an urban scene, a set of possibly overlapping homogeneous regions may result giving a partial segmentation. The partially segmented image must then be subjected to further processing if necessary, and the final image segmentation may be found with the help of higher level information. Totally correct and complete segmentation of complex scenes can not usually be achieved in the lower level of the processing phase. An immediate gain in using low level processing is the substantial reduction in data volume and the achievement of partial segmentation which may be good enough to solve a particular task like the registration of two different image data sets of the same area. In the literature, many authors have reviewed the topic of region segmentation (Haralick and Shapiro, 1985; Pal and Pal, 1993; La Moinge and Tilton, 1995), and recent work by Dowman and Ruskone (1997), based on a region growing segmentation technique to extract polygon features from a SPOT panchromatic image to register with a map, has shown encouraging results.

According to Marr (1975), boundaries of objects are perhaps the most important part of the hierarchy of structures that link raw image data to their interpretation. It seems that a better approach to obtaining boundaries of real world objects in images could be a region growing technique. Using this technique, regions which have been extracted from images can be converted to vector polygons (Gahegan and Flack, 1996), which lead to object shape descriptors in the images. This is described in the next section.

2.5 Object Shape Descriptors

In images, 3D objects are represented in a 2D plane and the shape properties of objects in it are computed in two dimensions. The two dimensional shape representation is sufficient in the majority of practical applications but, if 3D object reconstruction is the objective, then the object description becomes a much more difficult task. This section is limited to 2D shape features, which concerns the present research work and proceeds under the assumption that the described objects result from the image segmentation process. Later, chapter 4 discusses how to remove the part of the segmented objects which results from their heights, and select the planimetric portion of the segmented objects. The planimetric segmented portion will show a similar result to the 3D information of the objects for the registration of the images to their maps.

The traditional ways of representing image data structures, such as matrices, chains, graphs, lists of object properties, and relational databases, are important not only for the direct representation of image information, but also as a basis for more complex hierarchical methods of image representation. Moreover, in computer vision, chains are usually used for the description of object boundaries. One element of the chain is a basis symbol. For image data, chains can be arranged as a sequence of symbols (pixels) to represent objects in the image.

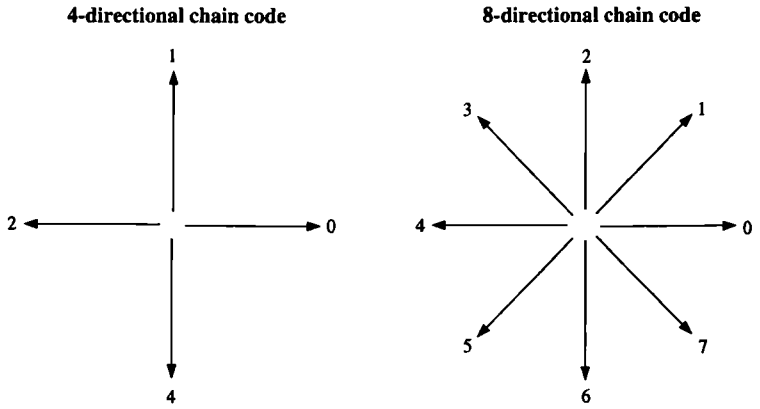
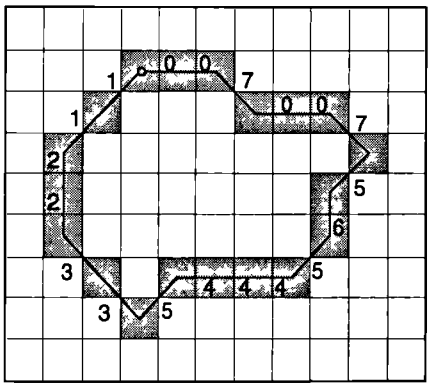


Figure 2.4: Directional representation for chain code.

Chain code was first developed by Freeman (1961) for the description of object boundaries or other one-pixel wide lines in images. It consists of a sequence of

connected straight-line segments with a certain length and direction connecting neighbouring pixels. The chain code uses either four or eight directional adjacency (see Figure 2.4) to identify neighbours in a rectangular grid. For example, as shown in Figure 2.5, a boundary of an object on a rectangular grid with each element (pixel) is shown with its directional number. The first derivative of the chain code (differences in direction) can be used instead of the code itself, which is rotation independent. The difference is computed by calculating (in a clockwise manner) the number of directions that separate two adjacent elements of the code. The first element of the difference is computed using the transition from the last to the first component of the chain code. Both the chain code and its differences in direction code are shown in Figure 2.5.



Chain code: 0070075654445332211
Frequency: 4 2 2 2 3 3 1 2

Differences: 1017012711007201010
Frequency: 7 7 2 0 0 0 0 3

Figure 2.5: Chain encoding of boundary using 8-directional chain code.

The chain code and its first derivative both have the disadvantage of being too long, sensitive to slight changes in the boundary and very sensitive to the chosen starting point (Gonzalez and Wintz, 1987). To remove this drawback, Abbasi-Dezfouli and Freeman (1994) propose the use of frequencies of the direction segments in the chain code and frequencies of change in direction to characterise the shape of polygonal objects in images. These frequencies are independent of the starting point and imply the use of a reduced set of numbers, viz, 4 or 8. Both the frequency of the chain code and the frequency of its differences in direction of

the above mentioned example are shown in Figure 2.5. This helps in matching two similar shape objects of the same size in two image data sets.

Abbasi-Dezfouli and Freeman (1994) also mention that the frequencies of direction in a chain code and its first derivaties do not always uniquely identify a shape, but they will provide a way of distinguishing shapes that may be indistinguishable using other shape attributes. However, these other attributes which characterise a polygon can be used besides these frequencies, eg the perimeter, the area within the boundary of the polygon, and the width and height of the corresponding minimum bounding rectangle, to compensate for possible uncertainty derived by frequencies of the chain code direction segments and change in direction. These four measures of the shape are rotation invariant but extremely dependent on the scale.

In the present work, the automatic determination of scale difference between two data sets to be matched is considered, which eliminates the effect of scale dependence for the rotation invariant shape measure attributes. This is described in the next chapter. Dowman et al. (1996) and Morgado and Dowman (1997) have also used shape characteristics of polygons and showed satisfactory results for matching areal features.

2.6 Application of Feature Extraction

Extraction of points, edges and areal features in images is of immense importance for reducing the volume of data and for handling the images for further processing. The extracted feature information from the images can be utilised to solve problems of a specific application. A close look at two applications of feature extraction are considered here, and are mentioned in the following subsections.

2.6.1 Feature extraction for matching

Two basic approaches are categorised in the literature for matching images: area-based and feature-based matching (Freeman and Abbasi-Dezfouli, 1993). In the area-based matching method the intensity values of pixels of two images of the same scene are directly compared. Matched pairs are decided on the basis of the

similarity measurement of the intensity values. This method calculates a correlation value over a predefined window and selects the pixel in the other image which gives the maximum correlation value within the search window (Barnard and Thompson, 1980; Ackermann, 1984; Gruen, 1985; Day and Muller, 1989; Hanaizumi et al., 1990). Feature-based matching is recognized as a scheme for matching features derived from original images, such as interest points, edges, or region segments. The area-based correlation method is not appropriate for matching images taken from different acquisition systems, because the grey-level values of the images to be matched can vary from sensor to sensor, and can result in unreliable correlation measures (Fonseca and Munjunath, 1996). In such cases, feature-based methods are more reliable, require less stringent approximations, and are believed to be a better approach to automatic image registration than area-based methods (Tseng et al., 1997). In this study, area-based matching is not considered because matching is taking place between a map and an image.

The use of feature-based matching with point features, or so-called interest points such as corners or centres of circular features in the field of digital photogrammetry, was initially reported by Förstner (1986). By comparing the descriptive parameters of interest points, conjugate point features are identified in images, but actually the descriptive parameters of different points are commonly not very distinctive. The global separability of a point feature is usually not good enough, and it needs certain approximation values to match its conjugate features globally. Hahn and Förstner (1988) achieve success using this technique by including known epipolar geometry for the automatic acquisition of digital elevation models.

Linear features are commonly used in the field of pattern recognition and computer vision for feature-based matching. The similar geometric properties of extracted linear features are compared to obtain conjugate linear features. In digital photogrammetry, the matching of linear segments (Greenfeld and Schenk, 1989) and zero-crossing curves of Laplacian of Gaussian (Schenk et al., 1991) have been used to automate the relative orientation of a pair of aerial photographs. Use of linear features provides more reliable matching results than using point features,

but they have the major disadvantage of suffering from segmentation and scale differences between images. Areal features, on the other hand, have an advantage over linear features because the segmentation differences of conjugate areal features are usually smaller than those of linear features and also areal features provide richer shape information than linear features (Tseng et al., 1997). Moreover, Dowman and Ruskone (1997) in their recent work mention that a benefit of using areal features for matching is that common polygonal features can be identified easily in different image types of the same scene.

In the next chapter, the shape information of areal features extracted from images is considered in detail. Attributes of polygonal features extracted from an image are used to match with the attributes of polygonal features in the map to find corresponding matched polygons. These corresponding polygons lead to conjugate points, using a dynamic programming matching algorithm for further processing for the map image registration.

2.6.2 Automation of map-image registration

An important question that comes into the picture here is which class of features (points, edges or regions) should be considered for the automation of registering images to images or images to maps. On the basis of the two arguments mentioned below, a class of feature is chosen for the automation of registration of images to maps.

- Not all abrupt intensity changes in images necessarily correspond to the relevant edges of objects (Lemmens, 1996). They may be due to noise, texture or shadows. It may be possible that real boundaries may be missing and may not show up as intensity changes due to low contrast or occlusion. Moreover, it is known that the perfect edge operator for all situations has not been developed and no single operator is likely to perfectly segment an image. Considering this, Ballard and Brown (1982) comment that:

  > "efforts are usually best spent in developing methods that can use or improve the measurements from unreliable edges rather than in

search for the ideal edge detectors".

In such cases, the extraction of edges to be used for automation of registration of images to maps, without the user's interaction, might produce incorrect results.

- The extraction of areal features from images for registration with maps could be the best approach for the automation of map image registration. It also seems that if automatic map image matching is to be successful, a large redundancy must exist for matching, which can possibly be obtained by extracting areal features, but not by extracting points or edges. According to Dowman and Ruskone (1997), it is easy to identify and extract common areal features for matching in different types of images of the same scene by using a region growing segmentation technique. Dowman and Ruskone (1997) also indicate that the extraction of areal features from images would generate a large redundancy for matching. Ackermann (1996), in a recent discussion on matching, coined the phrase "replace intelligence by redundancy", which can be obtained by the extraction of areal features from images.

On the basis of these two arguments, it has been decided that areal features will be considered for automation. The subsequent chapters explain the stepwise building of the automatic procedure for registering images to maps.

2.7 Conclusions

This chapter discussed previous studies on image registration related to the present work. After assessing the strengths and weaknesses of different approaches, it was decided to develop a new method for the automatic registration of images to maps. The method allows the automatic matching of maps and images of different scales, an approach which has not previously been tried. The approach will establish a relationship between the shape of the objects and the segmentation process in order to extract areal control features in images which belong to real world objects. This approach is tested in the next chapter.

Chapter 3

Matching the Map Model to Image Polygons

3.1 Introduction

There is always a loss of information in a two-dimensional (2D) image of a 3D scene. The 2D image may also contain ambiguous information due to perspective distortion, shadows or occlusions. This raises a classification problem concerning images, which can be overcome with the help of *a priori* knowledge in the form of constraints and context assumptions. In this study, a map model of objects in 2D is taken as *a priori* knowledge to solve this problem.

In the last chapter, previous work on feature extraction in images was discussed. The extraction of points, edges and regions was discussed for image registration. On the basis of the arguments developed therein, it was decided that the extraction of areal features from the image is more suitable for the automation of map-image registration. Areal features are considered because they can be more reliably extracted in the maps and images. Moreover, the areal features can provide a large redundancy for matching.

Previous workers have used region growing techniques for the segmentation of areal features in images. No work has been found in the literature which integrates iterative region growing with patch matching to extract areal features which most

closely resemble their map model counterparts. This combination would allow the extraction of a higher number of accurate boundary pixels from the image's areal features which will precisely match real world objects. The new work is introduced in this chapter. In this work, a higher number of accurate image boundary pixels are extracted, with the use of patch matching, by reducing the orientation and scale differences between map and image. The centre of gravity of an areal feature is used to find translation differences between map and image which allows the identification of the correct conjugate polygons in patch matching. The use of the centre of gravity in this way for image to map polygon matching has not been exploited before. The sensitivity of polygon attributes in patch matching caused by orientation and scale difference was described in the previous chapter.

In this chapter, methods are described which automatically extract significant polygonal features from images and match them, using shape attributes, to the appropriate corresponding polygonal feature of the map model of the same scene. The chapter also deals with determining the conjugate point pairs using matched corresponding polygons. Using a statistical model, the matched conjugate pairs are used to eliminate significant errors. Furthermore, they are used for the automatic registration of images to maps with high accuracy. This is discussed in the next chapter.

There are nine sections in this chapter. Section 3.2 gives information about the two assumptions upon which the experimental work is based. An overview of the methodolgy is given in Section 3.3. Three test areas are considered for the experiment and they are described in Section 3.4. Section 3.5 consists of general information about topographic maps and the preparation of a map model of objects in digital raster format by use of a semi-automatic Arc/Info system. Section 3.6 is concerned with obtaining the boundaries of significant areal features (objects) from images through an automatic process of smoothing and region segmentation. In Section 3.7 an algorithm is described which automatically determines the shape attributes of objects of the map model and of the extracted areal features from images. This section also describes an algorithm which matches the attributes of polygonal features extracted from the image with the attributes of polygonal

features of the map model to find corresponding matched polygons. Further, it explains the use of the corresponding matched polygons and the application of a dynamic programming algorithm to find the conjugate match points. Section 3.8 deals with matched values of the corresponding matched polygons to find the correct shape and size of extracted image areal features for each object in the map model. The concluding remarks for this chapter are given in Section 3.9.

3.2 Assumptions

In this research work, two assumptions are made:

- The best extracted segmented objects in an image, using the shape attributes, will match some parts of the corresponding object of the map model; and
- those parts of the matched objects which belong to their boundaries, will contain some planimetric points which coincide with the boundaries of real world objects.

This chapter considers the first assumption and finds matched boundary segments. The next chapter deals with the second assumption. A statistical model is developed which eliminates large error match points and selects planimetric match points automatically. These planimetric points are used for registering the image to the map with a high degree of accuracy.

3.3 Overview of the methodology

Maps and images both represent the Earth but in different forms. Information about the Earth is frequently available nowadays through satellite images and digital rasterised maps. Satellite images can be used for updating maps by a matching process. To match the map and the image it is essential to extract object boundaries or edges from both data sources. This needs preparation of the map and image (i.e. feature extraction) for matching. Matching maps and images of high resolution with a high accuracy is the main objective of this thesis.

Feature extraction and feature matching techniques for the preparation and matching of map and image are described in this chapter. The overall methodology for matching is shown in figure 3.1. Figure 3.2 shows a more detailed procedure for the preparation of map and image for matching.

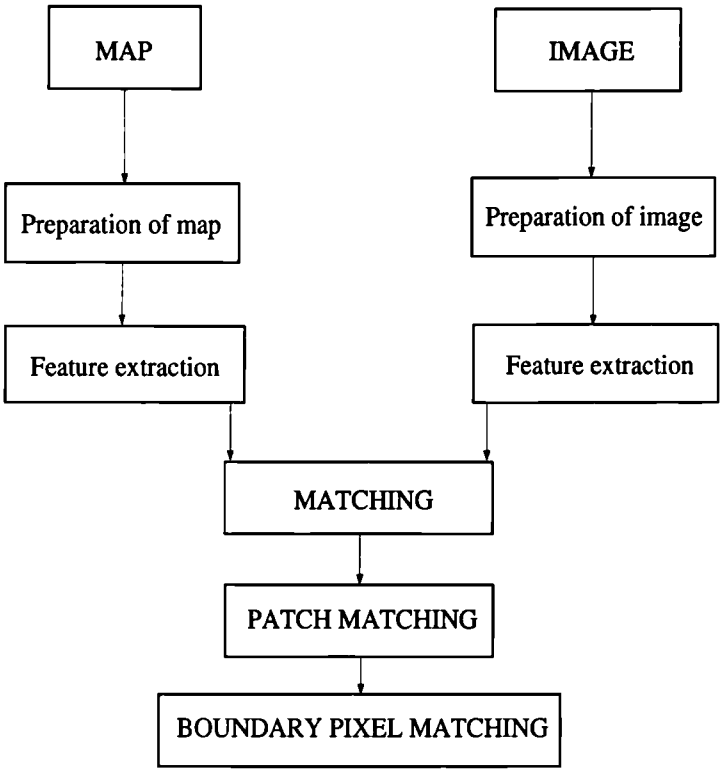


Figure 3.1: An overview of the matching process.

Map preparation is comparatively easy compared to image preparation. The preparation of map data differs according to the source. If the data is available in vector form, from a data-base, then the process of extracting features of a given type from the attributes attached is straightforward. Otherwise, manual methods are required for map preparation.

In this study, maps in raster format were used. The steps used for map preparation are shown in Figure 3.2. The first step is to select areal features and digitise them. The selected digitised areal features are given a separate unique grey value to distinguish them from each other. An algorithm then assigns to each areal feature its areas, centre of gravity, perimeter and other shape attributes. These attributes of map polygons are used with the attributes of image polygons in patch matching

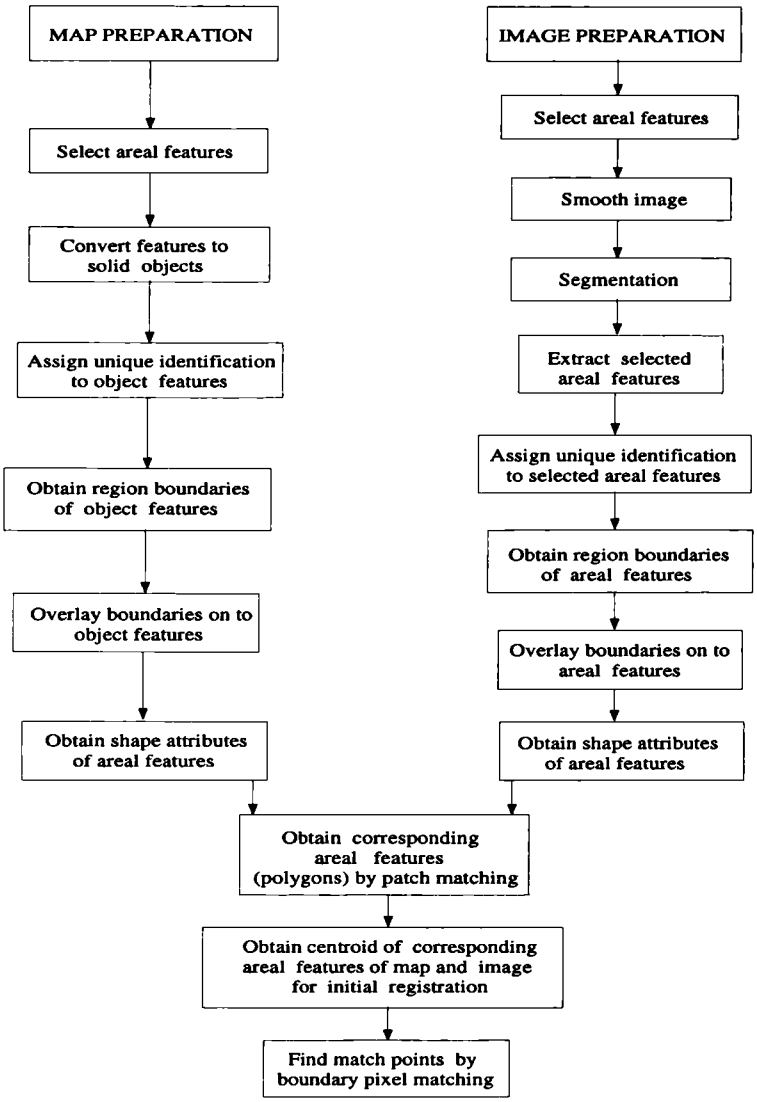


Figure 3.2: Map and image preparation for matching.

to find conjugate polygons of map and image.

In the image preparation process the extraction of suitable features in images is not an easy task. The extraction of suitable features from images is the main problem [Morgado and Dowman (1997); Heipke (1997); Dowman et al., (1996)]. A proper strategy is required to extract image features from images and to match them to their map models.

A strategy to solve the problem of extracting suitable areal features from images by a region growing and patch matching technique is developed in this work. This is shown in Figure 3.3 and the sequence of processes involved is explained below.

The basic concept for both these techniques was described in detail in the previous chapter.

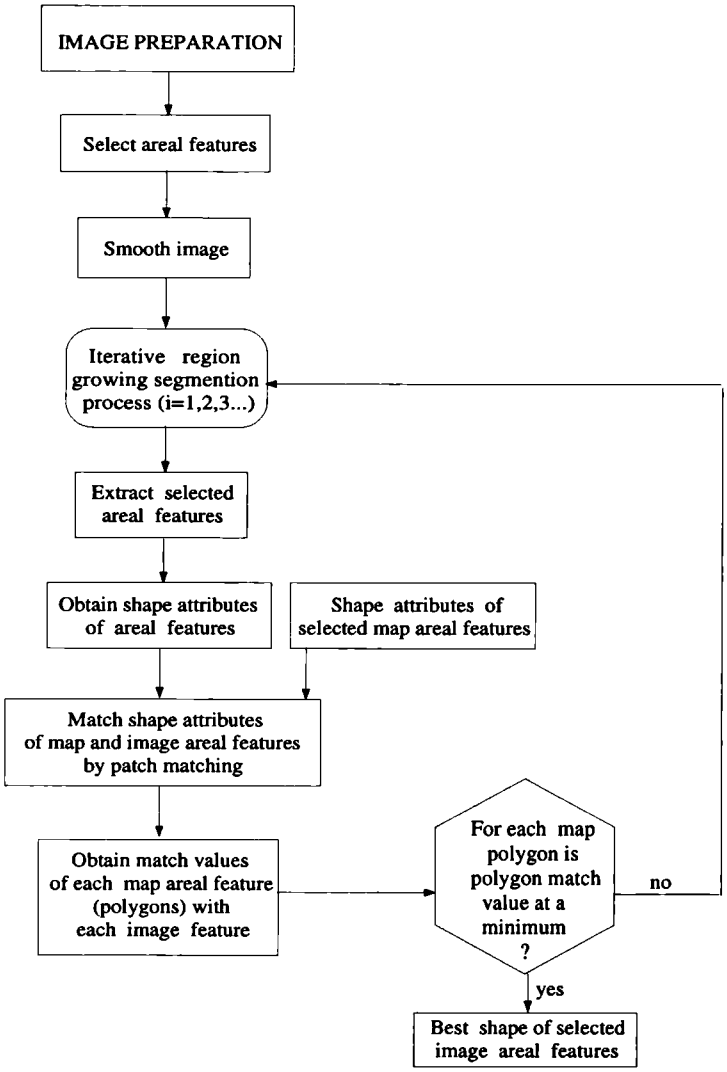


Figure 3.3: An algorithm for best shape extraction of selected areal features in images by combination of region growing and patch matching techniques.

Before describing the processes for extracting suitable areal features from images, it is important to mention that the current limitation of patch matching techiques is that they are only used for images of the same scale and orientation. In this work, a technique is developed for matching the map and image patches of *different* scales with similar orientation. The processing steps which are applied to achieve this target are as follows.

- determining shape attributes of map and image polygons.

- finding approximate scale ratio between map and image by use of the centroid of map and image polygons.
- Applying the scale ratio to the attribute values of map and image polygons to obtain new attribute values.
- determining approximate translation differences between the map and image co-ordinate system using the new centroid values of map and image polygon.
- using the translation difference to bring both the map and image co-ordinate systems close to each other in order to correctly match the conjugate centroids of map and image polygons.

This new approach to finding scale and translation differences has two benefits in patch matching. First, it allows the matching of conjugate centroids of map and image polygons. Secondly, the use of iterative region growing integrated with patch matching allows the extraction of conjugate polygons of map and image. The use of an iterative region growing process grows the size of polygons in segmented images and the patch matching plays a role in finding the cost function of matching. In patch matching, each map polygon is matched with each extracted polygon of the region growing segmented images. The pair of map and image polygons which show minimum matching cost value are considered the conjugate polygons. This process of combined region growing and patch matching finds the best shape and size of selected areal features from the iterative region growing segmented images. The advantage of the strategy used in this work is that it extracts the boundaries of image polygons whose maximum boundary pixels belong to real world objects.

Processes of extracting areal features from images: The following steps are used to extract areal features from images.

- The areal features in the image are selected by clicking on the selected features on the displayed image. These features are those which were selected for map preparation.

- The co-ordinates of the selected points, which belong to the selected areal features, are stored in a text file.

- The region growing process is used iteratively for image segmentation. The threshold grey value is increased in each iterative step. The threshold value is important to merge neighbouring pixels, as discussed in the previous chapter.

- The co-ordinates of the stored points are used to find the pixels with the grey values associated with these points and to extract them in each segmented image. These extracted pixels are in the form of patches and are the selected areal features. These patches have a different shape and size from the selected areal features in each segmented image due to the use of a different threshold value.

- Extracted patches containing less than nine pixels are not considered as polygons in this study. They are a product of insignificant growth of selected features due to a lower threshold value applied in the iterative region growing process. Such patches are considered to be clutter information and are removed by an algorithm from the segmented images.

A number of algorithms[1] were used to fulfil the requirements of the methodology described above. Some algorithms were developed in this work and others were taken from the departmental software database library. Given below are brief descriptions of the algorithms, image processing and graph drawing packages.

3.3.1 The image processing environment and software

The processing of the maps and images was performed on a SUN SPARC unix operating system work station. The algorithms were written in the C programming language and the C shell (*csh*) which shares much of the C language structure. Awk was used for analyzing and manipulating text files. These tools greatly aided automation of the processing. To process images, the routines of the Human Information Processing System (HIPS) were used and to digitize map object boundaries

[1]Algorithms are introduced with their authors name(s).

the Arc/Info system was used to prepare the map model for matching. Both these systems process images in their own internal formats and the format of the images subsequently needs to be converted.

The following algorithms were used for the preparation of maps and images, and for matching.

Algorithms for map preparation

The following steps were performed to extract areal features from raster maps.

- The Arc/Info system was used to digitize selected areal features and to convert them to solid objects by the use of the following commands.

  1. Converting raster maps to the IMAGINE format (used in the ERDAS IMAGINE image processing system).
  2. Converting the IMAGINE format to the Arc/Info GRID format using the command IMAGEGRID.
  3. Digitizing selected features on the Arc/Info system using ArcTools software.
  4. Converting digitized features to solid polygons using the Arc/Info command POLYGRID.
  5. Transferring the image with solid polygons to the IMAGINE format with the command GRIDIMAGE and then using the IMGHIPS command to change to a HIPS format.

  The steps outlined above transform each selected areal feature into homogeneous object polygons.

- The algorithm **boundaries** (Vohra) was used to determine the boundaries of homogeneous objects. This program first looks for homogeneous (solid) objects with different grey values and then extracts the boundaries of the objects. The algorithm **unique_id_number** (Vohra) was used to give the unique grey level value to each homogeneous object.

Algorithms for image preparation

The following steps were performed to extract areal features from images.

- Map selected features were chosen in corresponding images for matching. An algorithm **points** (Paramananda) was used to collect the coordinates of a point lying inside each selected feature. Using these coordinates was useful in picking regions in the segmented images.

- Images were smoothed to reduce the noise effect in the images. Algorithms **EdgePreserveSmoothing_SUN** (Ruskone) and **edge_preserve_smooth** (Newton) were used to smooth images.

- On smoothed images, an algorithm **regseg** (Kai) was used to segment the images into regions. The point coordinates of selected areal features were used on the segmented images to extract selected features by means of their grey values. The grey level values of the point coordinates of the selected features were obtained using the algorithm **getpixel** (Allison). These grey values were used to pick the selected features from the segmented images by using the algorithm **pixVal** (Vohra). Any small extracted features, of less than nine pixels, were removed using the algorithm **threshold_size** (Vohra).

- To find the boundaries of objects (selected features) and to give a unique grey value to each object, the same algorithms that were used for the maps were applied to the images.

Algorithms for matching map and image

Matching between map and image was performed using the boundary information of the extracted map and image polygons. The above algorithms described the preparation of the map and image and extraction of the boundaries of the map and image polygons for matching. Two matching techniques were used to perform the matching between map and image polygons in this work: patch matching and boundary pixel matching. The use of patch matching allows the extraction of the best shape and size of image polygons as well as the centroid of polygons in the

map and image for the first approximation for matching boundary pixels. The boundary pixel matching technique with the use of a first approximation produces a number of match points automatically.

The following operations were performed to match map to image to obtain match points.

- The shape attributes of map and image polygons are needed in patch matching. To determine the shape attributes of selected areal features in map and image the algorithm **shape_attributes** (Morgado and Vohra) was used. The selected areal features for separate grey values overlaid with boundaries were required to determine the shape attributes. An algorithm **overlay** (Vohra) was used to define the boundaries around objects.

- To match the shape attributes of map and image the scale factor, translation difference and rotation between the map and the image were used. The algorithm **size_diff** (Vohra) was used to find the scale factor, **trans_xy** (Vohra) to find the translation difference and **geometric**[2] to apply approximate rotation between the map and image.

- In patch matching, the shape attributes of the map and image polygons were used to find the corresponding patches. This matching was performed for each iterative region growing a segmented image. In each segmented image, a grey level threshold value (ThD value) is increased to merge neighbouring pixels into regions. For example, if the grey level threshold value is 2 (ThD 2), then any neighbouring pixels having a grey difference of 2 with the point region selected to grow, will be merged to the point region. In this way, each selected areal feature is grown in each segmented image and the patch matching technique applied. To find the best corresponding image polygon matched to each map polygon, the algorithm **match_poly** (Vohra) was used. This algorithm includes the scale factor, the translation difference and the rotational compensation algorithms to obtain the best image corresponding

[2]HIPS subroutine

polygons. The effect of scale and rotational differences in patch matching was discussed in the previous chapter in Section 2.5. The best match is chosen by a numerical matching value. The map and image polygons showing the minimum match value are the best conjugate polygons. The best matched image polygon for each map polygon is the product of a single image from the iterative region growing segmented images. The algorithm **mixture** (Vohra) was used to pick the best shape corresponding image polygon for each map polygon as well as projecting them into a single image frame and extracting the centroid of conjugate polygons of map and image.

- The boundary pixel matching technique was performed between the map and image polygons using their centroids to obtain a large number of match points. The algorithm **map_image_match** (Newton) was used to determine match points.

- A statistical model algorithm **planimetric** (Vohra) was applied to the match points to remove match points with large errors and to obtain a map-image registration of a high accuracy.

XVGR and IslandDraw were used for plotting graphs and drawing figures.

3.4 Test data

Three test experiments were performed using two different data sets to prove the assumptions made in Section 3.2. These data sets were used to develop methods to generate conjugate match pairs automatically.

1. The first test uses a section of a very high resolution ($\approx 2m$ ground resolution) Russian DD5 satellite image of the Regent's Park area, London (UK) along with an Ordnance Survey 1:10,000 map in raster format. In this test, a few control points are selected manually for matching in order to achieve a large number of conjugate match points.

2. In the second test, a section of a SPOT panchromatic scene (10 m ground resolution) of Elchingen-Grosskuchen in Baden-Wuerttemberg (Germany), along with 10 m resampled raster map data derived from an ATKIS 1:5,000 vector data-base, are used. The map and the image have the same scale. No manual control points are selected in this test; a translation difference between map and image is obtained automatically. Polygons are matched using their attributes and a large number of conjugate match points subsequently obtained.

3. The third test uses a subsection of the London Regent's Park area from the first data set around Lord's Cricket Ground. Here, also, no manual control points are used. Finding differences of scale and translation between the map and the image automatically, and matching corresponding polygons, ensures that numerous conjugate match points can be obtained automatically.

3.5 Generation of the map model

This section describes how a model of map objects is prepared and how it is used to match with the areal features extracted from the images. Three subsections are included in this section. They include general uses of geographical maps, a consideration of maps as reference data, and preparing the map model of the three data sets considered for the experiment.

3.5.1 The general utility of geographical maps

The graphical representation of ground surface in the form of maps is very natural, easily understandable, and has a long historical tradition. For us, a map is our main source of information about the environment, about close surroundings and far regions. The elements making up the major part of the environment are, for example, woods, agriculture, vegetation, water bodies, dwellings, and industrial complexes. These are usually the decisive elements making up the character of a region. It seems logical to represent these elements in maps without deformation

(Marsik, 1996). A map lets us see the broader spatial relations that exist over large areas and allows us to understand the interrelation of features.

Maps are smaller than the regions they portray. Each map has a defined dimensional relationship between reality and itself. This relationship is called a map scale, and because of the relative 'poverty' of map space, the scale sets a limit on the information that can be included. All maps use signs (symbols) to represent elements of reality. A detailed map of a small region, depicting its landforms, drainage, vegetation, settlement patterns, roads, geology, or a host of other detailed distributions, communicates the relationships necessary to plan and carry out many types of work. Building a road, a house, a flood-control system, or almost any other construction, requires prior mapping or a current map of that area. These types of maps are called large-scale topographic maps for general reference purposes. Less detailed maps of larger areas, showing floodplain hazards, soil erosion, land use, population character, climate, and so on, are indispensable to understanding the problems and potentialities of an area. These types of maps are called large to medium-scale topographic maps. Maps of the whole world indicate generalizations and relationships of broad earth patterns by which we may study the course of past, present, and future events. They are small-scale maps.

Uses of advancement in map-making: We know that each scientific discipline has its own selection of cartographic documents. These special maps serve the needs of various environmental professions. Mineral exploration, for example, relies heavily on geological maps, but complex problems such as species survival, enviromental contamination, and quality of life could not be addressed adequately by breaking the environment into parts watched over by specialists. The environment could best be understood and managed by viewing it as a system of interrelated processes. A change in one component or process is seen to lead to changes in others as well.

At all times, there has been a close relationship between mapping and the prevailing state of technological development. Cartographers have usually been quick to borrow and adopt technological innovations. One reason for this close

correlation between technological advances and cartographic achievement is that each succeeding generation of cartographers has had to face the same two problems. Society has been unrelenting in its demand for maps that are more up to date, accurate, and complete, and at the same time there has been a continual demand for greater accessibility to lower-cost maps. These dual forces have been at play regardless of cartographic developments in previous generations. Indeed, to a large extent it has been the constant struggle to meet these goals that has led to the kinds of maps we know today.

Today, there are many different kinds of mapmaking methods, and their objectives and techniques also seem very different, but all have the same goal of communicating spatial relationships. Photography, first applied to mapmaking in the mid-nineteenth century, rapidly gained great importance. Nowadays satellite images have given cartographers a dramatic new map form, the imagemap. The variety and uses of this image-based map form now seem unlimited and are highly in demand for the revision of maps.

3.5.2 Maps as a reference data source

Images have to be related to the ground coordinate system when an absolute orientation of images is required. A map is sometimes used to provide the ground coordinates of the points needed for absolute orientation. Considering maps as reference data is beneficial for two important applications. The first is registering different images of an area to a map (a common local geodetic coordinate system), and then fusing these images together to enhance the image information of the area. It is well established that more information about the terrain can be gathered with increased interpretation capabilities and reliable results by combining data from different sensors of the same area instead of using only one single satellite sensor (Pohl and Genderen, 1998). This leads us to the second application which is change detection for map revision. As is the case with interpretation capabilities, combining data from different sensors gives us change detection results superior to ones derived from only one sensor.

Up to date maps contain all the pertinent information available in an area. In spite of that, they are not considered to be an ideal reference. The reason is that in the map production process, at each stage of the drawing, the tools and methods used introduce inaccuracies in object positioning. The magnitude of inaccuracies depends on the map scale and on the accuracy required technically by the map producer. For instance, the same accuracy is not required for a 1:25,000 and a 1:100,000 scale map. There is also a distortion due to the constraints of the cartographic language. In order for the map to keep its legibility, object representation must respect constraints induced by human morphology, such as visual sharpness, alignment acuity, separation threshold and colour appreciation level. For this reason, dimensional specifications are required, such as a minimum deviation between objects or a minimum stroke width. For example, a road 7 meters wide on the ground would be 0.3 mm wide on a 1:25,000 scale map, but it may be plotted approximately between 0.3-0.5 mm wide for a better perception of the information. This type of distortion is known as map generalisation. This type of geometrical distortion in maps is unavoidable and so maps are not considered to be ideal reference data sources for real world objects. They can be treated only as reference data, a tool for understanding real world objects.

3.5.3 Models of map objects

A model of real world objects can be obtained from a map database. If a map database is not available for an area, then a digital raster map, nowadays widely available, can be used instead.

This subsection describes the use of the Arc/Info system to digitise a raster map on a computer screen in a user-friendly environment. Using this system, a map model of three data sets is prepared and matched with corresponding images. This process is described in Section 3.7.

Use of the ARC/INFO system

Arc/Info is a powerful geographic information system (GIS) with tools for the automation, analysis, display, and management of geographic information. ArcTools (a menu-based interface tool) is used in the present work to digitize raster maps on a computer screen. The following paragraphs describe how ArcTools is used to digitize a raster map.

ArcTools are functionally grouped into four tool sets: Map Tools, Edit Tools, Grid Tools, and Command Tools, as shown in Figure 3.4. In order to digitize a

Figure 3.4: ArcTools menu in the Arc/Info system.

raster map in the Arc/Info system, it is necessary that it should be brought first into a format which can be used on Arc/Info for display. In the present study, the raster maps in HIPS format were converted into IMAGINE format. This allows the use of the command IMAGEGRID on Arc/Info to convert the raster map into a suitable format.

In the Arc/Info system, digitisation of maps can be performed quite easily. A click of the cursor on the boundary of the selected object starts digitising the boundary of that object. When the digitisation process encounters a sharp turn on the boundary of the object, the software requires an input from the keyboard. The digitising preceeds leftwards by typing a number 2 or rightwards by typing number 3. This procedure can trace the boundaries of objects semi-automatically. The same procedure was applied in this work to digitize raster maps.

Preparation of a map model of three data sets

Three data sets of digital maps and images are tested in this chapter. In this section, the three map data sets are considered and the image data sets are dealt with in the next section. Three map models are prepared from the three map data sets of desirable objects using the Arc/Info system, as described above.

- **London Regent's Park area (UK):** A section of 1000 x 2000 pixels, in raster format, is extracted from an Ordnance Survey 1:10,000 scale map. This extracted section contains London Regent's Park area which is shown in Figure 3.5 (top). The boundaries of three objects were digitized from

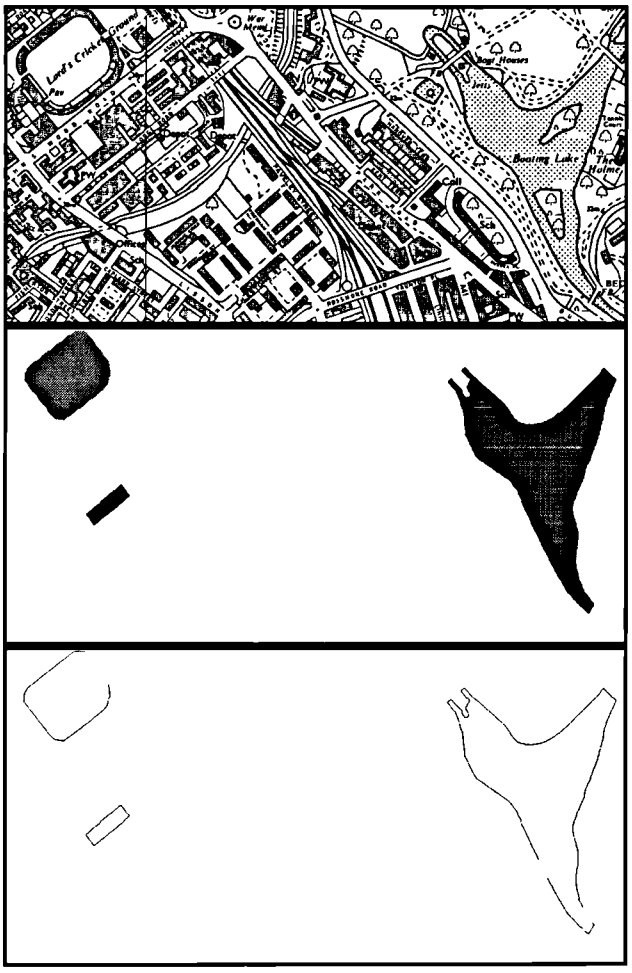


Figure 3.5: Ordnance Survey 1:10,000 raster map of London Regent's Park area (top), with polygons of interest (middle), and their boundaries (bottom).

the extracted map section using the Arc/Info digitization software, saved as

solid polygons and transferred to the HIPS system. In the HIPS system with the use of an algorithm **unique_id_number** a separate grey value is assigned to each solid polygon [Figure 3.5 (middle)]. This algorithm scans the map space from top left to bottom right and looks for separate solid polygons with digital number (DN) value zero while scanning, and assigns them a higher DN value in an order between 0-255 as it encounters them. The outcome of this algorithm is used as an input to an algorithm **boundaries** which detects the boundaries of the solid polygons [see Figure 3.5 (bottom)]. Later in this chapter, from these map boundaries and their corresponding image boundaries, pairs of manual control points will be selected for the initial transformation for matching the map and the image to achieve large redundancy.

- **Elchingen-Grosskuchen in Baden-Wuerttemberg (Germany):** A 10m resampled raster map created from an ATKIS 1:5,000 vector map was selected as the second data set. This map data has the same scale and orientation of the SPOT image data used for the test. A section of 450 x 590 pixels is extracted from this raster map. The extracted area contains the Elchingen-Grosskuchen area which is shown in Figure 3.6 (top). Five well distributed polygons were selected from this extracted map section. The boundaries of these polygons were digitized using the Arc/Info system, saved as solid polygons, and transferred to the HIPS system, where the use of the **unique_id_number** algorithm gave a unique grey value to each solid polygon. Applying the **boundaries** algorithm to these solid grey value polygons detected their boundaries and the **overlay** algorithm overlaid the boundaries on the polygons as shown in Figure 3.6 (bottom). This stage completes the map preparation stage. Later, this chapter describes how this map prepared information is used to obtain the shape attributes of the selected polygons for matching. In patch matching, as discussed earlier in Section 3.3, the shape attributes of map and image polygons are used to match and identify conjugate polygons.

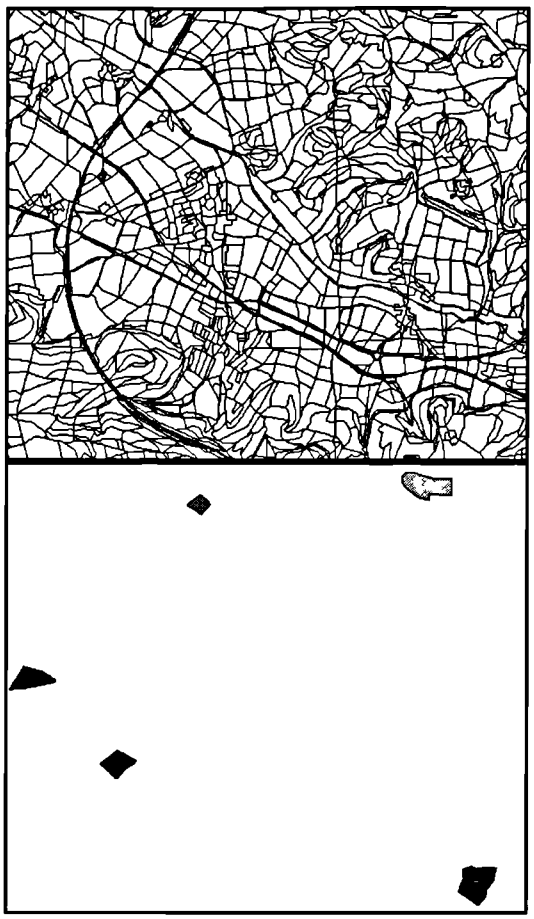


Figure 3.6: 10 m raster map of the Elchingen-Grosskuchen area created from an ATKIS 1:5,000 vector map (top) and polygons of selected areal features with their boundaries (bottom).

- **London Lord's Cricket Ground (UK):** A section of 1088 x 1818 pixels is extracted from an Ordnance Survey 1:10,000 scale map in raster format. This extracted section contains the area around Lord's Cricket Ground which is shown in Figure 3.7 (top). The boundaries of seven objects were digitized from the extracted map section, using the Arc/Info digitization system, saved as solid polygons, and transferred to the HIPS system. All solid polygons were given a seperate DN value and their boundaries were detected using the algorithms described above for the two test areas. Using the **overlay** algorithm, the boundaries of the polygons were overlaid on the solid polygons, as shown in Figure 3.7 (bottom).

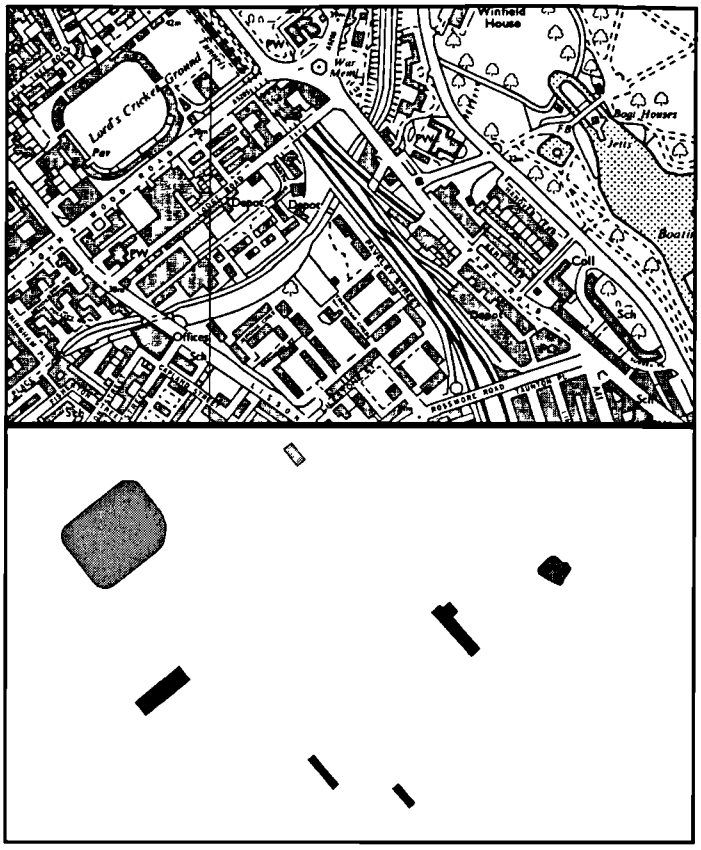


Figure 3.7: Ordnance Survey 1:10,000 raster map around London Lord's Cricket Ground (top), and polygons of selected areal features with their boundaries (bottom).

In this test, the map data has a different scale and orientation to the DD5 image data used for the experiment. A solution to such a problem is considered in this study. Later in this chapter, an algorithm (discussed earlier in Section 3.3.1) is described which uses the centre of gravity (centroid) of extracted polygons from the map and the image to find the differences of scale and translation between the map and the image automatically, after adusting the rotation differences of the map and image. These findings will be used in patch matching to identify conjugate polygons of the map and the image. Their centroids will be used as control points for the initial transformation for matching the map and the image to achieve numerous conjugate match points automatically.

3.6 Extraction of image areal features

This section describes the extraction of desirable areal features from images to match with the corresponding map model prepared in the last section. Two subsections are included in this section. They include the smoothing of images and the extraction of desirable areal features, using region segmentation, from the three image data sets considered for the experiment.

3.6.1 Smoothing of images

In satellite images, real world objects are not represented as homogeneous objects, i.e. they are not shown by unique grey level (DN - digital number) values. Instead they comprise a cluster of a few close DN values; and smoothing and segmentation techniques can be used to bring these DN values together to achieve homogeneous objects. In Section 2.2.2, the smoothing technique was discussed in detail, in particular how to suppress noise in the images while smoothing and preserving the edges of the areal features. Here, the text describes the observed effects of smoothing on the three image data sets used for the experiment.

Two smoothing algorithms were tested on the three image data sets. The first algorithm **edge_preserve_smooth**, moves over an image with a window of 5 x 5 pixels which has nine masks, as shown in Figure 2.1. Similar to this algorithm, **EdgePreserveSmoothing_SUN**, uses an additional four corner filters (masks) passing through the central pixel. Both algorithms work on the principle developed by Nagao and Matsuyama (1980). After selecting the mask which shows minimum pixel variance, the mean of this mask is assigned to the central pixel. This mask never lies across an edge. This process is performed for each pixel in the image. It is noticed in the present work that the second algorithm has the advantage over the first in that it works quite fast and that the size of the window can be changed. Increasing the size of the window takes more computational time and does not show a considerable change in the smoothing effect. Nevertheless, the second algorithm **EdgePreserveSmoothing_SUN** is considered for the development of the automated system on the basis that it works faster and has an option for chang-

ing the window size. The following points were observed using these smoothing algorithms.

- The grey values of images were not smoothed enough to achieve complete homogeneity of any areal features, either by smoothing once or iteratively.
- Iterative smoothing showed sub-regions developed in almost every region, and this created difficulties in obtaining images of real world objects after segmentation.
- Breaking of the boundaries of areal features and merging with neighbouring features was produced using iterative smoothing. The boundaries of mixed merged objects created problems in matching with the map model.

Images showing urban areas, for example with small buildings and/or groups of trees, with poor surrounding contrast, have a tendency to merge together on iterative smoothing to build large polygons of their own. These large polygons were found difficult to remove automatically from the scene. Thus, in this work, it was decided to smooth images once instead of iteratively and to rely on the segmentation method which is discussed in the next subsection to find a solution making the areal features homogeneous.

3.6.2 Extraction of desirable areal features using region growing segmentation

For the registration of images to their corresponding maps, it is not necessary to extract a large number of areal features, but it is important that they are well distributed in a suitable pattern over the image (Dowman et al, 1996). In the present work, taking this information into account, three tasks were considered important for the extraction of well distributed areal features. These tasks, with emphasis on their applications, are mentioned below.

1. **Selection of desirable areal features for extraction from images:** The areal features selected from the displayed image on a computer screen have

to be those for which the map model is available or prepared. The user has to click a cursor approximately on the centroids of the selected areal features to store the co-ordinates of the clicked points. These clicked points belong to the selected areal features.

2. **Extraction of features with the correct shape from the images:** The second task is to determine the grey level values (DN) of selected points which belong to selected areal features. This process is performed for each iterative region growing the segmented image. In sections 2.4.3, 2.5 and 3.3, the methods of extracting the patches of selected areal features using grey values of the points as well as determining the shape attributes of patches was described. On matching the shape attributes of these extracted areal features with the shape attributes of the prepared map polygons, the best shaped (correct) feature will be selected from one of the iterative region growing segmented images for each selected map polygon. The best extracted areal features will represent the shape of its real world object and will be sufficient to solve the problem of the registration of the image to its map model. Marr (1982) and Fischler and Bolles (1985) have stated that a symbolic image description from a segmentation process may not be complete, but it does serve a special purpose by extracting meaningful structures from digital images, that is, extraction of geometric and structural features which are termed the best shaped areal features in this study.

3. **Removal of non selected features and clutter from the images:** All features except the selected desirable areal features must be removed from the image. This task of removal of unwanted features will enable us to extract only the desirable areal features in a well distributed pattern over the image to avoid any confusion in the matching process in order to obtain a large number of match points.

Region growing segmentation of three data sets

The region growing technique was described in detail in Section 2.3.3. Here, the same technique is used for the region segmentation of images. An algorithm **regseg** is used iteratively on the smoothed images of the three data sets. This algorithm uses two loops to define seed points, the first in image coordinate space and the second in grey level. Each iterative step is increased by a grey level difference threshold value (predefined local difference threshold, i.e. ThD value) to grow regions from the seed points. The region grows from each seed points in all directions and merges neighbouring pixels in the region until the absolute difference of the grey level of seed pixel and neighbouring pixel is above a predefined threshold grey value.

In this study, the predefined threshold grey values (ThD) between one and ten were sufficient for the iterative use of the **regseg** algorithm to find the best shape of desirable areal features for the three data sets. Using thresholds of 1 and 2 did not show useful growth of the selected regions. So, here, the results of ThD 3 to ThD 10 only were used for the interpretation.

To extract the best shape of desirable areal features and their boundaries, the above three tasks were performed on the three image data sets.

- **London Regent's Park area (UK):** A high resolution Russian DD5[3] ($\approx 2m$ ground resolution) satellite image is used for the first test. A section of 850 x 850 pixels is extracted from the image which contains London Regent's Park area and is shown in Figure 3.8 (top-left). Three areal features are selected from the image section whose map model was prepared in the last section. Here, a selection of features is performed by clicking the cursor close to the centroids of the selected features in the image displayed. The coordinates of the clicked points are stored using the algorithm **points**. These clicked points belong to the selected features. The processing of the image is done by first smoothing the image; then the **regseg** algorithm is applied to it for region segmentation. The **regseg** is used iteratively for growing

[3]A DD5 image is resampled product of KVR-1000 (0.75m ground resolution) satellite image.

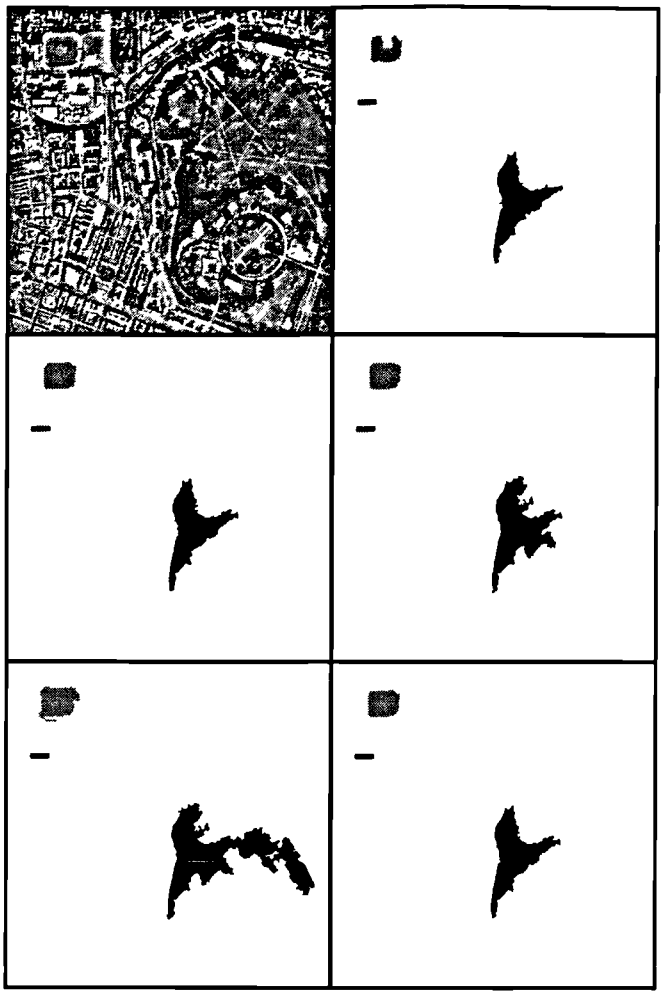


Figure 3.8: A section of DD5 image of London Regent's Park area (top-left), its segmented result with removal of clutter at grey level threshold difference (ThD) of 4 (top-right), ThD of 5 (middle-left), ThD of 7 (middle-right), ThD of 8 (bottom-left), and best shape extracted from desirable areal features (bottom-right).

the selected regions. In each iterative region growing segmented image, the grey values (DN) of the stored points are determined using the **getpixel** algorithm. These grey values are extracted in the form of patches from each segmented image (using the algorithm **pixVal**). Non-selected features (other features) in the image are removed automatically; they are considered to be clutter information. The new images of the segemented images contain only useful information, i.e. selected areal features. The iterative region growing process used here increments a grey level threshold (ThD) in ten steps, from

1 to 10, and extracts polygons of selected features in the segmented images. Comparison of the extracted image polygons is done with the extracted map polygons. Each map polygon is compared visually with the extracted image polygons of the segmented images to select the best shape image polygon, and the rest of the image polygons are masked out, using a **add_poly**[4] program. These masked frames (segmented images) are merged together onto a single image to obtain the best segmented result (using the algorithm **overlay**), which is shown in Figure 3.8 (bottom-right).

In this test, the selection of the best image features is performed manually from the region growing segmented images, but for two other tests this manual step is automated. Figure 3.8 shows the images top-right, middle-left, middle-right and bottom-left which are the result of region growing segmentation with ThD of 4, 5, 7 and 8 respectively. Figure 3.8 (bottom-right) shows the best extracted shape of desirable areal features without any clutter information. Using the **boundaries** algorithm, the boundaries of the best extracted areal features are detected, and they are shown in Figure 3.9.

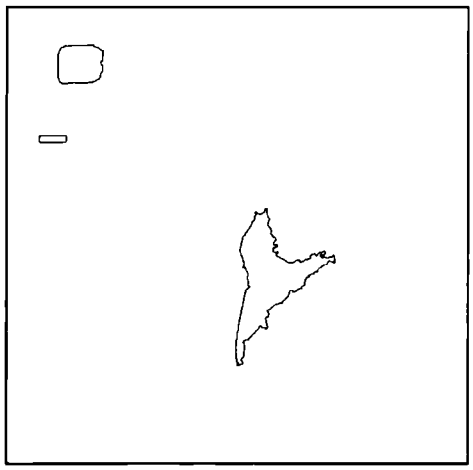


Figure 3.9: Boundaries of the best extracted desirable areal features of London Regent's Park area.

Later in this chapter, pairs of manual control points will be selected from these image boundaries and their corresponding map boundaries for the ini-

[4]HIPS subroutine

tial transformation for matching the map and the image to obtain large redundancy.

- **Elchingen-Grosskuchen in Baden-Wuerttemberg (Germany):** From a SPOT panchromatic image of 10m ground resolution, a section of 480 x 600 pixels is extracted for the second test. The extracted area contains the Elchingen-Grosskuchen area shown in Figure 3.10 (top-left). This image data

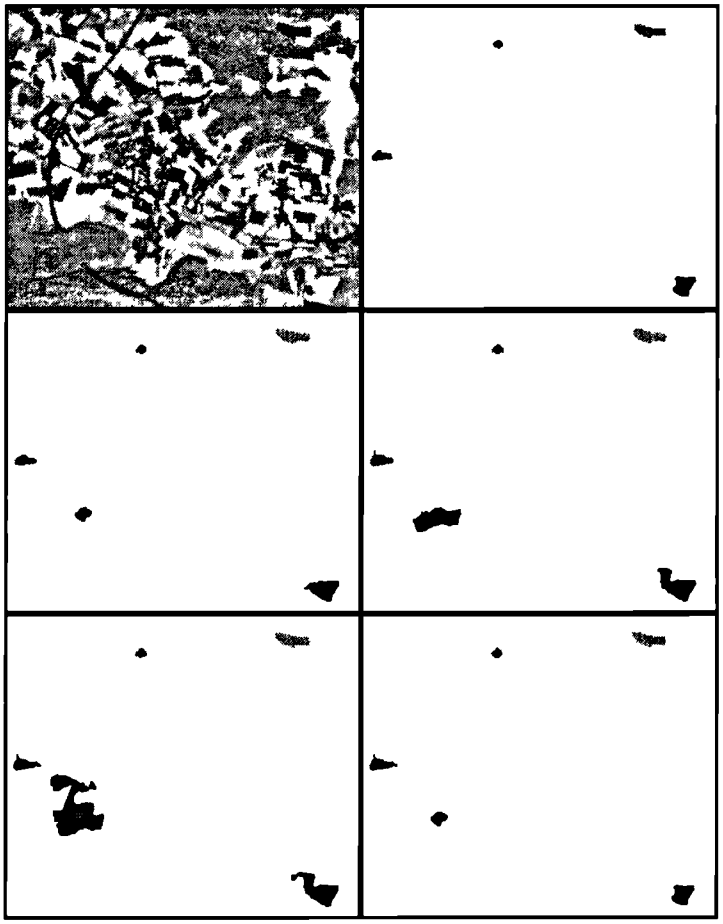


Figure 3.10: A section of SPOT image of the Elchingen-Grosskuchen area (top-left), its segmented result with the removal of clutter at the grey level threshold difference (ThD) of 4 (top-right), ThD of 6 (middle-left), ThD of 9 (middle-right), ThD of 10 (bottom-left), and best shape extracted from desirable areal features (bottom-right).

has the same scale and orientation as the corresponding map data. Five areal features are selected from this image section. The process of feature

extraction for this test is similar to the first test, from the click of the cursor on the selected features to the ten region growing iterative process. But the process of extracting the best shape for each feature from these iterative frames (segmented images) is different. Here, it is not selected manually on the basis of a visual check and masking other features. A code **mixture** is used which automatically finds the best match of desirable features from the segmented images to their map model, and then extracts and merges them to a single frame using the **pixVal** algorithm. A selection of best shape features is based on a matching value. The method of patch matching described in Section 3.3 uses the attributes of map and image polygons to identify the best matched image polygon to each map polygon. The pairs of map-image polygons showing the minimum match values are considered to be the best matched polygons. Figure 3.10 shows the images top-right, middle-left, middle-right and bottom-left, which are the result of region growing segmentation with ThD of 4, 6, 9 and 10 respectively with clutter removed and with only selected areal features retained. Figure 3.10 (bottom-right) shows the best shape extracted polygons of the five desirable areal features. Applying the algorithm **boundaries** to these extracted solid polygons (areal features) detected their boundaries as shown in Figure 3.11 (top). To overlay the boundaries on the solid polygons, the **overlay** algorithm is used and the result is shown in Figure 3.11 (bottom). This completes the stage of image preparation. This information will be used later in this chapter to find the attributes of the extracted image polygons. The centroid of the image polygons will be used with its corresponding centroid of map polygons as a control point for the initial transformation for matching the map and the image to achieve large redundancy.

- **London Lord's Cricket Ground (UK):** A section from the first test area of DD5 image is selected for the third test. This test is carried out to automatically find the scale and translation differences between the map and image. The map and image of this test area were viewed on a screen to

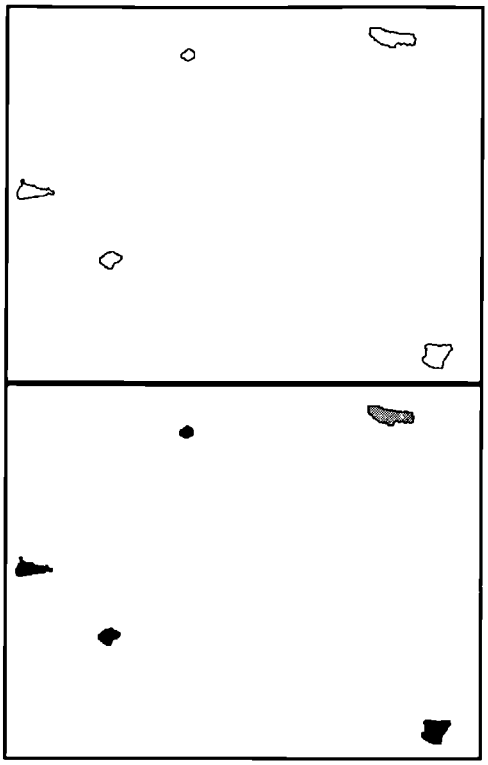


Figure 3.11: Boundaries of the best extracted desirable areal features of Elchingen-Grosskuchen (top), and boundaries overlaid on its solid polygons (bottom).

see the orientation difference between them. It was noted that the image is oriented approximately 30 degrees anti-clockwise to the map. The algorithm **geometric** was used to align the image to the map. The approximated rotation angle is input manually as it is needed to use the algorithm. The rotated image is shown in Figure 3.12 (top-left) with smoothing effect (using **EdgePreserveSmoothing_SUN**). It consists of 412 x 555 pixels and contains the area around the Lord's Cricket Ground. Seven areal features are selected from this image section while the map model was described in the last section. The process of the extraction of features for this test is the same as the second test described above, from the click of the cursor on the selected features to the last step of extraction of the best shape of desirable selected areal features on to a single frame.

The result of this test is shown in Figure 3.12. The images shown are the outcome of a region growing segmentation with ThD of 3, 5, 6 and 10, which

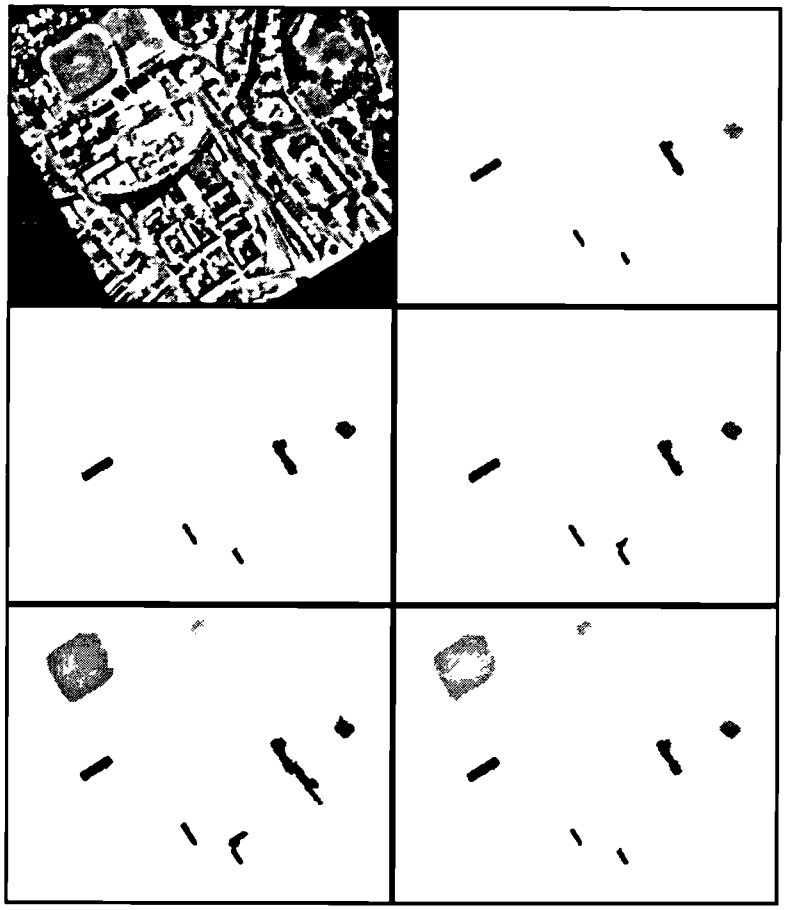


Figure 3.12: A section of DD5 image around London Lord's Cricket Ground area (top-left), its segmented result with the removal of clutter at grey level threshold difference (ThD) of 3 (top-right), ThD of 5 (middle-left), ThD of 6 (middle-right), ThD of 10 (bottom-left), and best shape extracted from desirable areal features (bottom-right).

shows only selected areal features. In Figure 3.12 (bottom-right), the best shape extracted from the seven selected areal features is shown, with its boundaries, in Figure 3.13 (top). Figure 3.13 (bottom) shows the boundaries overlaid on the solid polygons which will be used to find the shape attributes of the polygons in the next section, as well as finding the scale and translation differences between the map and image. These findings will be used further in the patch matching to identify the conjugate polygons of the map and image.

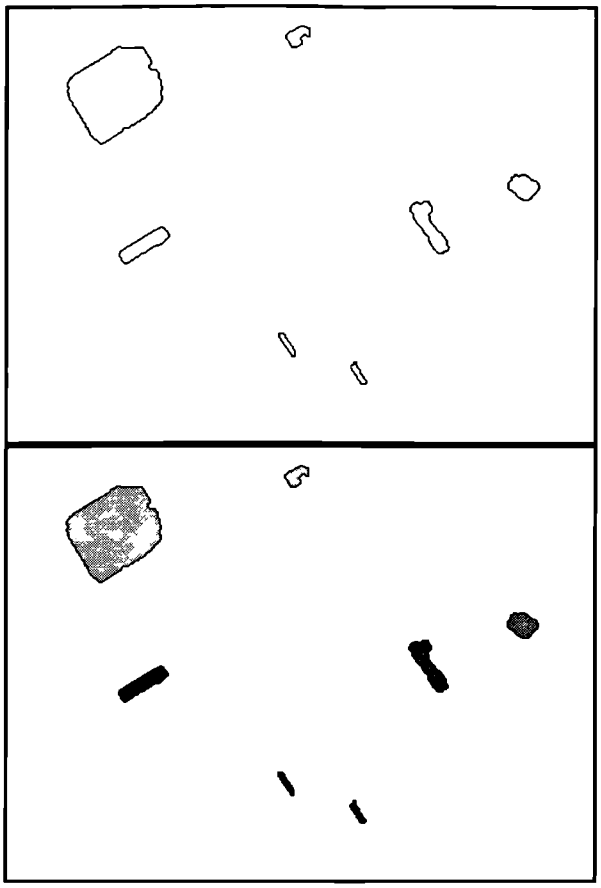


Figure 3.13: Boundaries of the best extracted desirable image areal features around Lord's Cricket Ground (top), and boundaries overlaid on its solid polygons (bottom).

3.7 Matching the map model to image features

This section deals with the automated measurement of images and the corresponding ground co-ordinates. Two matching techniques are involved in the measurements. One is patch matching which determines conjugate polygons based on the shape of the extracted patches of map and image. The other is dynamic programming which finds the corresponding points of the map and image. Both techniques are interconnected to obtain match points, as shown in Figure 3.14. Each step shown in the figure is discussed in the current section. The match points are obtained for registering the image to the map.

Both methods are illustrated with examples and the results obtained from the map to image matching are presented. This section includes the methods of match-

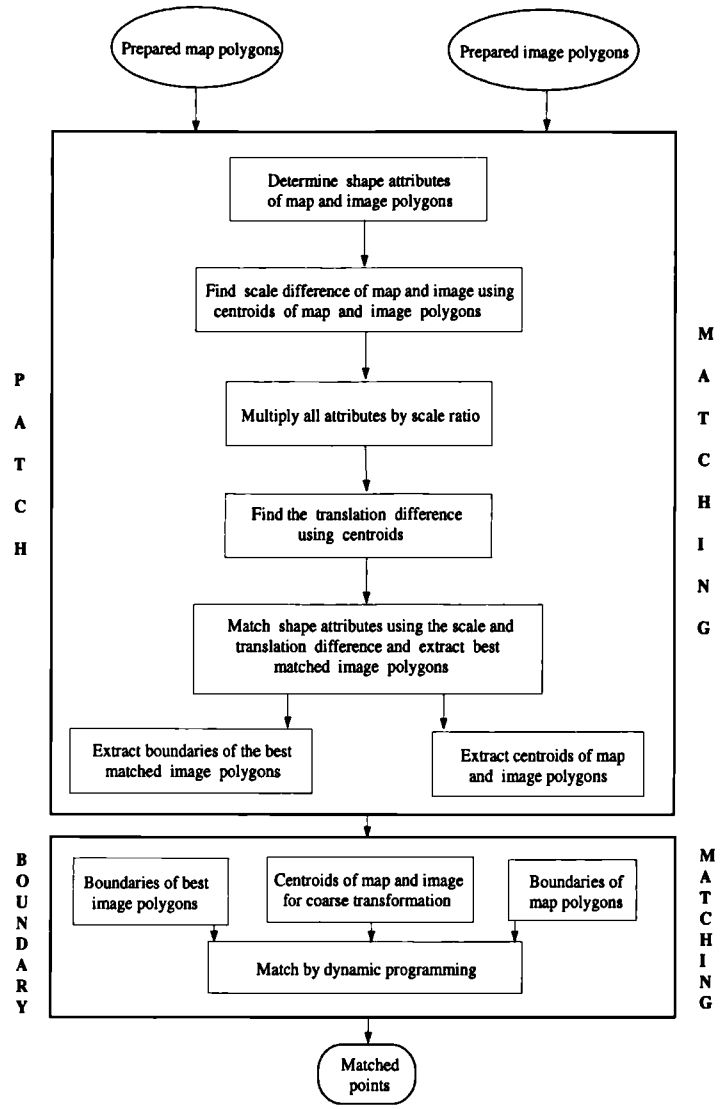


Figure 3.14: Schematic diagram showing a strategy for combining two matching techniques to obtain match points for map-image registration.

ing the map model to the image patches, using dynamic programming to determine conjugate match points, and combining the two matching methods for automation.

3.7.1 Matching map-image patches

The last two sections were concerned with preparing the map and image polygons for matching. This subsection deals with the relationship between the map and the image polygons in order to match them on the basis of their shapes. The extracted polygons of the map and the image of the two data sets, i.e the second

and the third test areas, were considered here for the experiment. It was stated before in Section 3.4 that the map and image of the second test have the same scale and orientation, and the scale and orientation of the third test are different. The problem of different scale and orientation was discussed earlier in Section 2.5 for matching patches. A solution, matching the patches correctly, is given here. Four components are involved in the patch matching process, and they are described below.

1. **Shape attributes of polygons:** It is known that maps and images show different representations of any terrain. This is the reason why a map and an image of an area cannot be used in stereo-image matching methods for selecting common points. However, a technique based on matching the shape of common features of the map and the image can be used to provide conjugate points for computation of the first approximation of the transformation between the map and the image. For using this shape-based matching technique, the shape attributes of the extracted features from the map and the image are needed [Abbasi-Dezfouli and Freeman (1994) and Dowman et al. (1996)]. These parameters are:

   - frequency of direction in the chain code (code);
   - frequency of the first derivative of the direction in the chain code (differences);
   - number of pixels that compose the boundary of the polygon (perimeter);
   - area of the polygon;
   - width and height of the bounding rectangle; and
   - centre of gravity of polygons.

   The benefit of using shape-based matching, i.e. using the above mentioned parameters, is that it provides the centre of gravity (centroid) of polygons for the first approximation for matching. Using centroids as control points eliminates the tedious traditional method of matching and also supports the procedure for the automation of matching. The drawback of this technique

is that it is only successful with images which have the same scale and orientation. Two examples of previous work are discussed below to show the application of this technique.

Abbasi-Dezfouli and Freeman (1994) have shown that the frequencies of direction in a chain code and its first derivative do not uniquely identify the shape of an object. Adding other shape attributes (perimeter, area, width and height of bounding rectangle of the objects) to it may distinguish the shape for matching patches between two images of the same resolution. Abbasi-Dezfouli and Freeman considered two SPOT images of the same area (both having the same scale and orientation) taken at different times, and showed a good result for matching corresponding patches using shape attributes. Morgado and Dowman (1997) showed satisfactory results using the same method to find the correspondence between polygons extracted from an aerial photograph and a map of the same area. They reduced the image 8 times to bring its scale nearly to that of the map used and with almost the same orientation. From these two pieces of research, it seems that the shape attribute matching technique is suitable when the two data sets of an area have the same scale and orientation. But, in the present study, the patch matching technique is developed further and it can be used for different scale data sets. This is explained later in this section.

The algorithm **shape_attributes** was used to determine the attributes of the extracted map polygons as well as the image polygons of the region growing segmented images of Elchingen-Grosskuchen and London Cricket Ground test areas, obtained in the last two sections. This algorithm finds the parameters listed above. The recorded parameters of the extracted map polygons and the best extracted shaped image polygons of the two test areas are shown in tabular form, Table 3.1 to Table 3.8. The information in these tables is used in the patch matching for identifying conjugate polygons of the map and image of the two test areas. In the tables, frequencies of chain code and the frequencies of the first derivative of the direction in the chain code show

Polygon	Code								Differences							
Number	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
1	47	4	7	18	38	1	22	6	105	18	2	1	0	0	0	17
2	8	10	1	9	7	10	2	8	19	19	2	0	0	0	0	15
3	22	16	8	0	44	8	2	14	57	28	3	0	0	0	0	26
4	18	13	1	14	10	17	1	10	31	29	1	0	0	0	0	23
5	31	9	18	13	11	19	18	3	66	27	2	1	0	0	0	26

Table 3.1: Chain code and first derivative frequencies of map polygons of the Elchingen-Grosskuchen area (units in pixels).

Polygon Number	Perimeter (pixels)	$\sqrt{Area}$ (pixels)	Bounding Rectangle (pixels)		Centre of Gravity (pixels)	
			width	height	X	Y
1	144	31.43	57	28	475.27	20.10
2	56	16.00	26	20	212.36	39.07
3	115	25.21	52	24	26.92	215.20
4	85	23.64	41	28	124.02	297.47
5	123	32.35	43	39	530.70	416.54

Table 3.2: Attributes of map polygons of the Elchingen-Grosskuchen area.

magnitude (number of pixels) in 8 directions (for detail see Section 2.5) for defining the shape of the polygons; and they also show the square root of area instead of the area of polygons to keep the same dimensional units (in pixels) for all attributes in patch matching in order to identify the correct matches.

As mentioned earlier in this section, the map and the image of the Elchingen-Grosskuchen test area have the same scale and orientation, whereas the Lon-

Polygon	Code								Differences							
Number	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
1	47	2	15	15	36	7	16	9	82	19	12	0	0	0	1	33
2	7	6	6	3	11	3	8	4	20	12	4	0	0	0	0	12
3	29	6	18	5	35	7	10	12	67	20	8	0	0	0	1	26
4	17	6	11	5	15	7	11	4	28	13	10	0	0	0	0	25
5	28	7	16	11	17	11	19	4	57	24	6	0	0	0	2	24

Table 3.3: Chain code and first derivative frequencies of image polygons of the Elchingen-Grosskuchen area (units in pixels).

Polygon Number	Perimeter (pixels)	$\sqrt{Area}$ (pixels)	Bounding Rectangle (pixels)		Centre of Gravity (pixels)	
			width	height	X	Y
1	148	28.65	58	25	483.55	37.11
2	49	12.40	17	15	222.43	57.92
3	123	22.18	46	25	31.26	233.79
4	77	18.24	27	22	127.34	319.42
5	114	27.22	37	31	539.90	440.71

Table 3.4: Attributes of image polygons of the Elchingen-Grosskuchen area.

Polygon	Code								Differences							
Number	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
1	7	15	10	36	5	17	8	36	77	28	3	0	0	0	0	26
2	68	129	55	100	56	120	85	79	400	150	0	0	0	0	0	142
3	28	31	15	32	28	28	21	29	94	60	2	0	0	0	0	56
4	7	34	11	106	2	34	16	101	238	35	4	0	0	0	1	33
5	22	95	8	27	24	94	8	28	187	60	1	1	0	0	0	57
6	5	11	11	67	1	14	9	66	129	27	3	0	0	0	0	25
7	3	13	6	44	2	13	7	43	94	21	1	0	0	0	0	15

Table 3.5: Chain code and first derivative frequencies of map polygons of Lord's Cricket Ground area (units in pixels).

Polygon Number	Perimeter (pixels)	$\sqrt{Area}$ (pixels)	Bounding Rectangle (pixels)		Centre of Gravity (pixels)	
			width	height	X	Y
1	135	39.03	58	61	754.93	68.61
2	693	222.23	276	284	287.68	278.54
3	213	64.07	88	78	1447.38	374.02
4	312	75.49	129	144	1187.00	525.54
5	307	69.21	145	130	411.20	691.08
6	185	43.74	82	89	835.45	901.70
7	132	36.09	59	63	1045.65	960.98

Table 3.6: Attributes of map polygons of Lord's Cricket Ground area.

Polygon	Code								Differences							
Number	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
1	11	9	6	7	10	7	11	4	32	15	4	0	0	0	1	13
2	44	32	30	30	28	35	40	17	110	68	6	0	0	0	0	72
3	14	5	10	10	10	9	6	10	34	21	2	0	0	0	0	17
4	12	6	17	27	7	6	22	22	61	30	2	0	0	0	0	26
5	18	24	5	5	17	25	4	5	36	36	1	0	0	0	0	30
6	4	0	10	12	3	0	11	11	24	13	3	0	0	0	0	11
7	4	2	9	10	4	1	11	9	21	14	3	0	0	0	0	12

Table 3.7: Chain code and first derivative frequencies of image polygons of Lord's Cricket Ground area (units in pixels).

Polygon Number	Perimeter (pixels)	$\sqrt{Area}$ (pixels)	Bounding Rectangle (pixels)		Centre of Gravity (pixels)	
			width	height	X	Y
1	66	15.16	22	20	269.74	26.18
2	257	72.63	89	92	101.38	78.30
3	75	21.18	29	24	480.40	165.85
4	120	26.03	37	49	395.30	204.65
5	104	24.35	47	34	126.05	222.76
6	52	9.84	15	22	258.81	313.63
7	51	9.85	15	21	324.65	340.00

Table 3.8: Attributes of image polygons of Lord's Cricket Ground area.

don Cricket Ground test area has a different scale and orientation between the map and the image. The effect of both cases can be observed by comparing the corresponding attributes of the map and the image polygons from these Tables. It can be seen, from the Tables, that each corresponding attribute value of the map and image polygons is close and the differences are low for the Elchingen-Grosskuchen test area, but, in the case of the Lord's Cricket Ground test area, the situation is different. The differences of corresponding attribute values are high due to the scale difference between the map and image. In the last section, while preparing the image of the Lord's Cricket Ground test area for matching, the image was rotated to get aligned with the map; but the scale difference was not compensated for. This section considers how to compensate for the scale difference. The scale difference between map and image of this test area is not adjusted manually by a resampling process as performed by Margado and Dowman (1997). Instead, a method was developed which finds and uses the scale ratio. The attributes of the map and image polygons are multiplied by the scale ratio to bring the attribute values to a comparable state. This decreases their differences which will help in finding the corresponding best matched polygons. How to find the scale difference factor is described below.

2. **Scale difference between map and image:** A basic geometric method can be used in finding the scale difference between two data sets of an area. The distance between the centroids of two or more real world objects in two image data sets, of a different scale, will show a scale ratio between them. For example, consider four areal features in two images as shown in Figure 3.15.

   In Figure 3.15, the value of A, B, C, D, E and F, which are the distances between the pair of centroids of the areal features, can be calculated by the following equations.

$$A = \sqrt{[(X_a - X_b)^2 + (Y_a - Y_b)^2]} \tag{3.1}$$

$$B = \sqrt{[(X_b - X_c)^2 + (Y_b - Y_c)^2]} \tag{3.2}$$

$$C = \sqrt{[(X_a - X_c)^2 + (Y_a - Y_c)^2]} \tag{3.3}$$

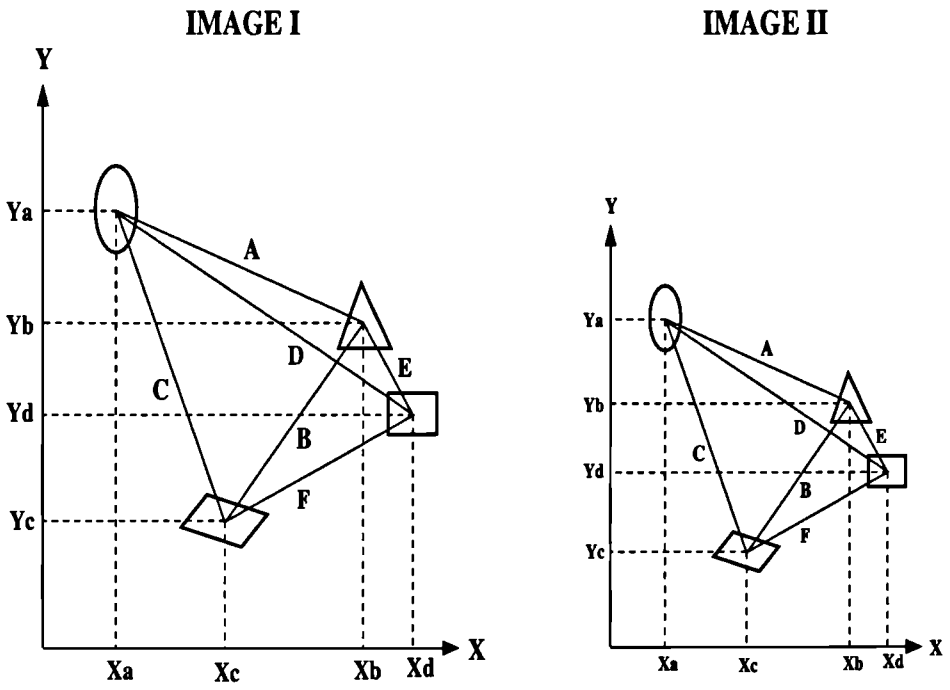


Figure 3.15: Determination of the scale difference factor between two image data sets.

$$D = \sqrt{[(X_a - X_d)^2 + (Y_a - Y_d)^2]} \tag{3.4}$$

$$E = \sqrt{[(X_b - X_d)^2 + (Y_b - Y_d)^2]} \tag{3.5}$$

$$F = \sqrt{[(X_c - X_d)^2 + (Y_c - Y_d)^2]} \tag{3.6}$$

The scale difference factor (S_{diff}) between two image data sets IMAGE I and IMAGE II can be determined, assuming a scale constant over each image, by following equation 3.7.

$$S_{diff} = \frac{(A + B + C + D + E + F) \text{ of IMAGE I}}{(A + B + C + D + E + F) \text{ of IMAGE II}} \tag{3.7}$$

On the basis of equation 3.7 an algorithm **size_diff** was developed to find the scale factor between two image data sets of an area. This algorithm is applied on the extracted map and image polygons of the Lord's Cricket Ground test area. The centroids of the map and image polygons, given in the above Tables, are used to find the scale factor between them. A scale factor of 2.97 is obtained, and is used to multiply with the attributes of the

image polygons in the patch matching process to find the best matched image polygon to each map polygon.

Verification of the validity of this algorithm was performed by applying it to the Elchingen-Grosskuchen test area. The map and image of this test area have the same scale. The result showed a value of 0.99 scale difference, which satisfies the validity of the algorithm.

3. **Translation difference between map and image:** In this study, the scale and translation difference between a map and an image are used in the patch matching process as additional parameters to find the best matched image and map polygons. The scale ratio is used for multiplying all the attributes of the polygons, whereas the translation difference is used only to change the value of the centroid of the polygons, because other attributes are invariant to the translation in patch matching. Neither of these parameters actually changes the scale or translates the map and the image; they are only used to change the value of the attributes in order to identify correct conjugate polygons.

   To determine the translation difference between a map and an image, the centroid of the same areal features in the map and the image can be used. The map and the image should have similar orientations and a known scale difference. Using the scale ratio on the centroid of the polygons of the map and the image will determine the translation difference between them. This can be illustrated by a diagram. Consider four areal features in two image data sets, as shown in Figure 3.16. Their centroids can be used in finding the translation difference between the image data sets in X direction (T_X) as well as in Y direction (T_Y), as shown in the following equations.

$$T_X = \frac{S_{diff} * (X_a + X_b + X_c + X_d) \text{ of IMAGEII} - (X_a + X_b + X_c + X_d) \text{ of IMAGEI}}{\text{Number of polygons}} \quad (3.8)$$

$$T_Y = \frac{S_{diff} * (Y_a + Y_b + Y_c + Y_d) \text{ of IMAGEII} - (Y_a + Y_b + Y_c + Y_d) \text{ of IMAGEI}}{\text{Number of polygons}} \quad (3.9)$$

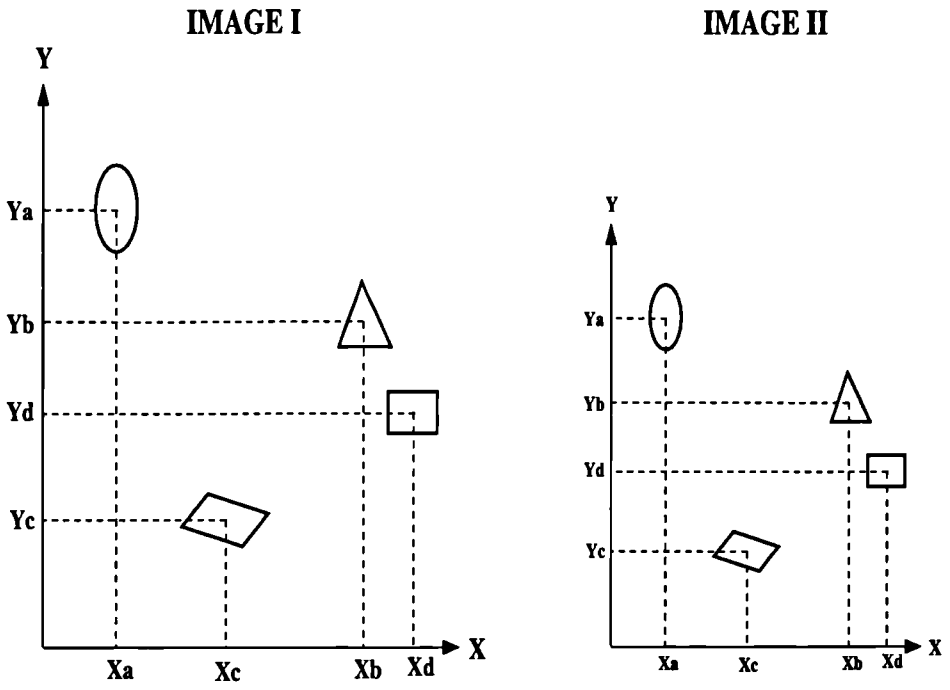


Figure 3.16: Determination of translation difference between two image data sets.

Using these two equations, an algorithm trans_xy was developed to determine the translation difference between two image data sets. This algorithm is applied on the map (IMAGE I) and the image (IMAGE II) of the two test areas, Elchingen-Grosskuchen and Lord's Cricket Ground. The centroid of the extracted map and image polygons are used, as given in the above Tables (Table 3.2, Table 3.4, Table 3.6 and Table 3.8). The T_X and T_Y values for the Elchingen-Grosskuchen test area are 4.24 and 17.95 pixels and for the Lord's Cricket Ground test area -22.70 and 30.45 pixels respectively. These values will be added to the centroid of the map polygons of the respective test areas to compensate for the translation difference between the map and image.

4. **Matching shape attributes of map and image to extract best matched image polygons:** The patch matching technique can be used to compare polygons of two images or of an image and a map of the same scene to find the conjugate polygons. Size, shape and geometric position of polygons are the basic measurements which identify the individuality of an extracted polygon

in the image to match it to its corresponding polygon in the map. These can be calculated by finding the shape attributes of polygons [see Abbasi-Dezfouli and Freeman, 1994; Dowman et al, 1996; Morgado and Dowman, 1997; and Vohra and Dowman, 1997], with the additional parameters of scale and translation difference of the map and image which were discussed above.

In this work, two algorithms were used for patch matching. The first algorithm **shape_attributes** was used to find the attributes of the extracted map polygons and image polygons of segmented images produced by iterative use of region growing segmentation. The second algorithm **match_poly** was used to calculate a matching value for each map-image polygon. The matching value is determined by adding the value of the absolute differences of the corresponding attributes of each map polygon with each image polygon. The pairs of polygons which show the minimum matching value are considered the best matched polygons. This algorithm includes the parametric values of translation difference and scale factor between the map and the image. By this method, each map polygon finds its best matched image polygon (BMIP) if the orientation of the map and image is approximately similar. Figure 3.17 shows a map model consisting of five polygons and the image of the same objects. Both map and image are prepared for matching. The image has gone through ten iterations of region growing segmentation and each segmented image is shown in the figure. Consider that this prepared map and image is of the Elchingen-Grosskuchen test area. By patch matching, each map polygon in this figure will find its best matched image polygon in one of the segmented images.

The following equations are used for patch matching in this study to find the best matched image polygons.

$$Cx_{diff(i,j)k} = |(M_{cx(i)} + T_X) - (S_{diff} \star I_{cx(j)k})| \tag{3.10}$$

$$Cy_{diff(i,j)k} = |(M_{cy(i)} + T_Y) - (S_{diff} \star I_{cy(j)k})| \tag{3.11}$$

$$MV_{(i,j)k} = \sum |M_{a(i)} - (S_{diff} \star I_{a(j)k})| + Cx_{diff(i,j)k} + Cy_{diff(i,j)k} \tag{3.12}$$

$$\text{Total number of Match Values} = i * j * k \tag{3.13}$$

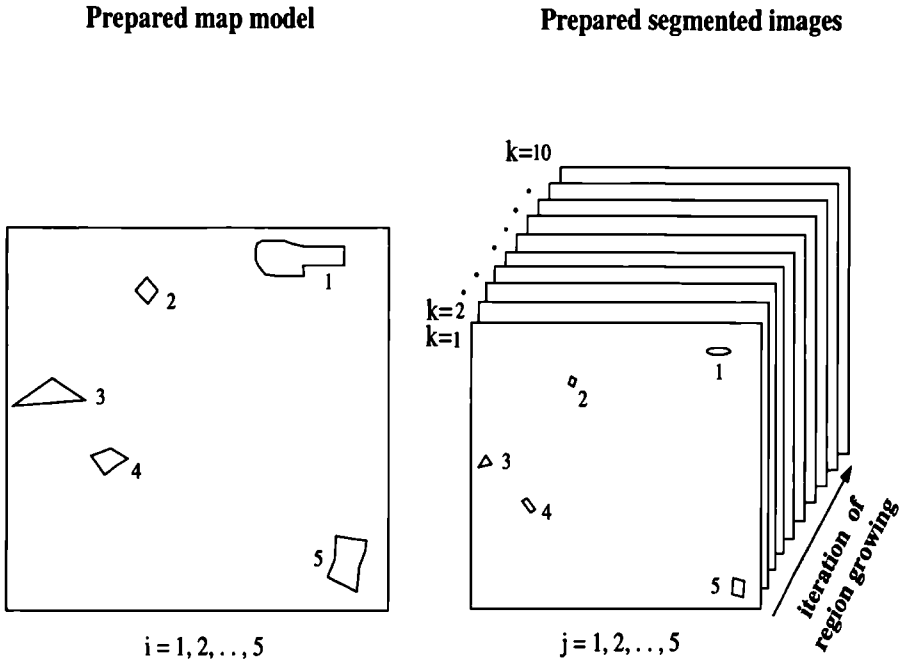


Figure 3.17: Prepared map and image polygons for patch matching.

$$\text{BMIP} = \text{Minimum}[\text{MV}_{(i,j)k}] \tag{3.14}$$

Where:

i denotes map polygons; j image polygons ; and k refers to the segmented image iteration

M_{cx} and M_{cy} denote x and y coordinates of centroid of map polygon

I_{cx} and I_{cy} denote x and y coordinates of centroid of image polygon

Cx_{diff} is the difference of x coordinates of map and image centroid

Cy_{diff} is the difference of y coordinates of map and image centroid

T_X and T_Y are the translation difference between map and image co-ordinates in the X and Y direction

S_{diff} is the scale difference between image and map

BMIP denotes best matched image polygons

M_a and I_a denote the corresponding shape attributes of map and image polygons, where a = 1,2,...,22 attributes, which are shown in the above Tables (3.1-3.8) for the map and image of both test areas.

For each map polygon, all attributes (i.e. 22 in total) are used to find the best matched image polygon. Table 3.9 and Table 3.10 show the result of the best matched image polygons of the areas. In section 3.6, the results were shown in a diagramatic form. Therein, Figure 3.10 (bottom-right) shows the best matched image polygons of the Elchingen-Grosskuchen area and Figure 3.12 (bottom-right) of the Lord's Cricket Ground area. These images were obtained by use of the algorithm **mixture**. The algorithm, using the DN values, extracted the patches of the best matched image polygons from the respective region growing segmented images (frames) and merged them to a single frame image. The DN values of the patches with respective segmented images are listed in Tables 3.9 and 3.10. The centroids of the map and

Map polygon number	BMIP in segmented frame	Minimum matched value	Matched polygon DN value	Centroid of map polygon (pixels)		Centroid of image polygon (pixels)	
				X	Y	X	Y
1	ThD 9	0.771521	51	475.27	20.10	483.55	37.11
2	ThD 6	2.052954	102	212.36	39.07	222.43	57.92
3	ThD 10	0.920889	153	26.92	215.20	31.26	233.79
4	ThD 6	1.800120	204	124.02	297.47	127.34	319.42
5	ThD 4	0.789884	241	530.70	416.54	539.90	440.71

Table 3.9: Best matched map-image polygons of Elchingen-Grosskuchen area.

image polygons shown in these tables will be used, later in this chapter, for the initial transformation to match the map and image of the two test areas.

Analysis of automation of patch matching: A combination of four components, which are described above, are used in automatic patch matching. These components are:

1. shape attributes of polygons of map and image,

Map polygon number	BMIP in segmented frame	Minimum matched value	Matched polygon DN value	Centroid of map polygon (pixels)		Centroid of image polygon (pixels)	
				X	Y	X	Y
1	ThD 10	0.086367	36	754.93	68.61	269.74	26.18
2	ThD 8	0.023993	51	287.68	278.54	101.38	78.30
3	ThD 5	0.046626	85	1447.38	374.02	480.40	165.85
4	ThD 5	0.034232	127	1187.00	525.54	395.30	204.65
5	ThD 10	0.035027	182	411.20	691.08	126.05	222.76
6	ThD 3	0.040202	204	835.45	901.70	258.81	313.63
7	ThD 5	0.075692	241	1045.65	960.98	324.65	340.00

Table 3.10: Best matched map-image polygons of Lord's Cricket Ground area.

2. scale difference between map and image,

3. translation difference between map and image co-ordinates, and

4. matching attributes of map and image polygons using scale and translation difference.

The combination of these components showed the automatic identification of correct conjugate map and image polygons. In this study, the two shape attributes were given different weights to identify correct matched conjugate polygons as well as to automate the patch matching. These two attributes are *area* and *centroid* of the polygons. Different weights were assigned to these two attributes so that correct polygons are matched automatically. The square root of the polygon area was used instead of the area on its own so that this attribute should not play a dominating role in patch matching and lead to incorrect matches. Using the area of the polygon results in a higher difference value in patch matching than other attributes, even by a slight change in the shape of the polygon. So, the square root of the area was considered to reduce the dominant character of the area in patch matching.

The polygon centroid was considered as a special parameter in patch matching. In spite of the fact that the centroid of a polygon is not actually a shape attribute of that polygon, it identifies the position of the polygon in the co-ordinate system. This character of the centre of gravity of polygons is exploited in this study. The

centroid of polygons is in this work to find the scale and translation difference between map and image. The centroids of polygons alone, in conjuction with the scale and translation difference, in a map and image of similar orientation, identified the correct conjugate centroids. This raised confidence in the centroid difference of map and image polygons compared to other attributes in patch matching and, further, gave a reason to assign the weights to the centroid difference, so that the conjugate polygons cannot be missed. The centroid difference was multiplied by a factor of ten (an arbitrary number) in patch matching and this executed the correct matches.

This study involved two important steps. First, it identified correctly matched image-map polygons using the centroid of polygons and second, it extracted the best matched image polygons to their map model. Both these steps played different roles in patch matching:

- centroids of polygons find the matched conjugate polygon, using scale and translation difference, in all the segmented images (frames) irrespective of their extracted shape or size for each map polygon, and
- other attributes combined together, using scale difference, find the best shape of the matched conjugate polygons, from the segmented images.

3.7.2 Determining conjugate match points

Once the corresponding polygons are established, the next problem is to match them in order to extract the conjugate match points. To optimize the results, a method of dynamic programming can be used. This is a solution strategy for combined optimization and involves a sequential decision-making process, expressed as a recursive search [Bellman and Dreyfus (1962) and Ballard and Brown (1982)]. In the present work, dynamic programming described in detail by Newton et al. (1994) is used. This technique owes much of its inspiration to Maitre et al. (1989). The algorithm was originally developed to detect changes in forests by Newton (1993), and the same algorithm is used here to find conjugate match points of areal features extracted from maps and images. This subsection is divided into two. The

first part includes information about the dynamic programming algorithm and the second part shows the results of the map-image matching of the three test areas.

1. **Dynamic programming matching technique:** The core of the dynamic programming algorithm used is a routine for matching map and image object boundaries or matching the detected edges of two images. This algorithm works on an affine transformation. To start matching, at least three control points are required in a well distributed pattern to obtain an initial loose transformation between them. By use of the initial transformation, each map boundary pixel is projected to predict its position in the image. This allows the matching to proceed by tracing along a map boundary, pixel by pixel. The projected map boundary pixel seeks to find the nearest edge pixel. By defining a search area around this predicted edge pixel position, as shown in Figure 3.18, the nearest edge pixel is looked for. The predicted position is considered as the centre of a circular search area with a certain radius which is a function of the residuals obtained from the map to image transformation. All edge pixels falling inside this area are considered as potential matches. The distance between the predicted edge pixel and the edge pixel under consideration defines the matching, which is a cost function. Only one edge pixel of the image can be matched to one map boundary pixel. The edge pixel under consideration with minimum cost is considered as the best match. All boundaries of the map, once matched with the lowest cost of matches, are recorded and used as new control points to update the transformation of the map to the image. The matching procedure starts again, excluding the previous selected boundaries, and continues until all matched boundaries are selected again and the matches recorded. The matching and refinement of the transformation proceeds thus side by side till the conjugate matched points are selected.

2. **The conjugate points of three test areas:** For matching map and image, three sets of information are required to obtain conjugate points. These are the boundaries of the map and image polygons and the control points. Prepa-

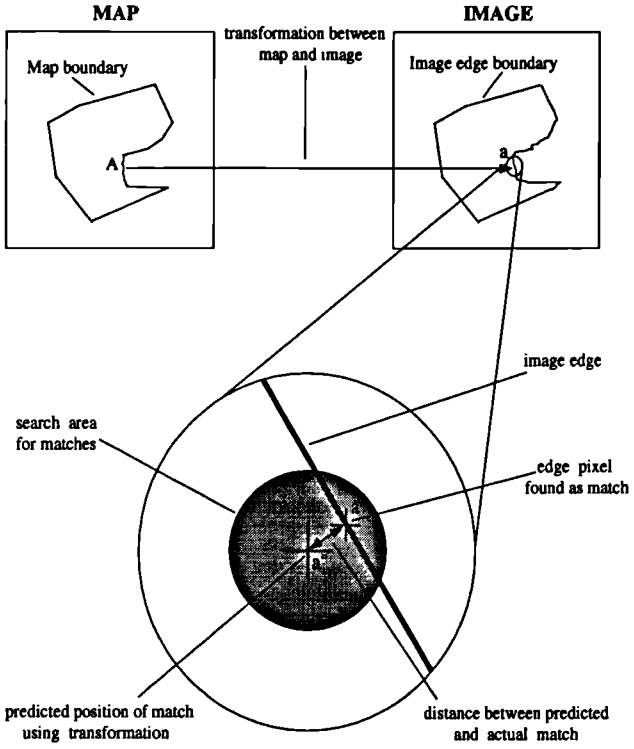


Figure 3.18: Matching map boundary pixels to image edge boundary pixels (after Newton et al., 1994).

ration of the boundaries of the map and image polygons of the three test areas were shown in Sections 3.5 and 3.6. The control points for the first test, from the map and image of London Regent's Park area were selected manually and are shown in Table 3.11. For two other test areas, Elchingen-Grosskuchen and the Lord's Cricket Ground area, the centroids of the extracted map and image polygons were chosen as the control points, and are shown in Tables 3.9 and 3.10. The algorithm **map_image_match** was used for matching the bound-

Map X	Map Y	Image x	Image y
186.00	56.00	85.00	92.00
516.00	370.00	239.00	108.00
188.00	1430.00	380.00	457.00
900.00	1904.00	665.00	437.00

Table 3.11: Control points of map and image of London Regent's Park test area (units in pixels).

aries of the map and image polygons to obtain the conjugate points. The resulting conjugate points in these three test areas are shown in Figure 3.19, Figure 3.20 and Figure 3.21. The matching algorithm also created a text file of the match points, which is used for further processing to remove the matches of large errors and to obtain a high accuracy of image registration.

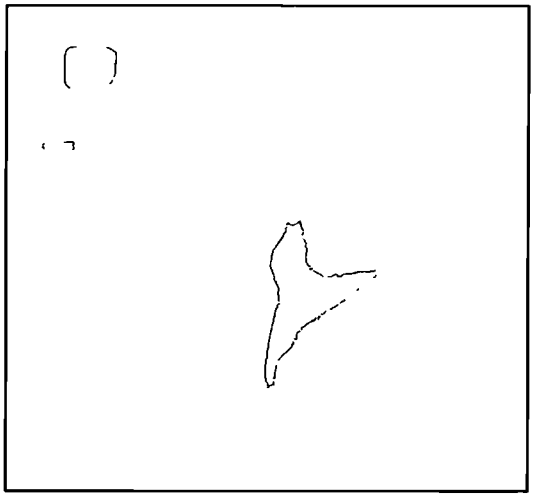


Figure 3.19: Matched points (865) of the London Regent's Park subscene.

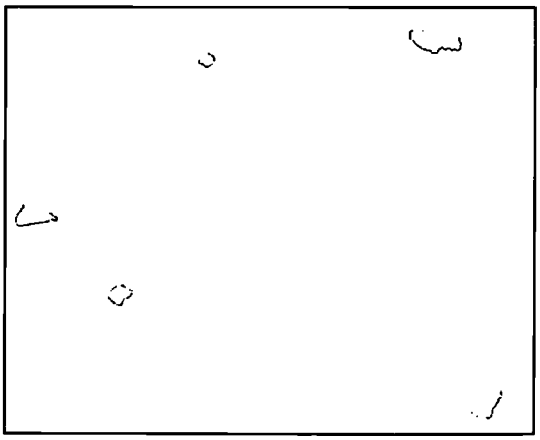


Figure 3.20: Matched points (269) of the Elchingen-Grosskuchen subscene.

3.7.3 Analysis of the combined effect of two matching techniques

To automate the measurements of the coordinates of corresponding control features on the map and the image, the combination of two matching techniques were used.

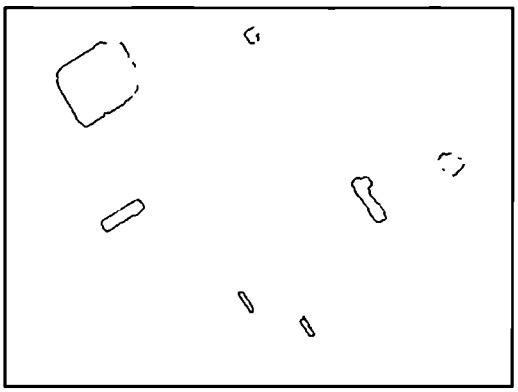


Figure 3.21: Matched points (1039) of the Lords's Cricket Ground subscene.

- Patch matching technique: shape-based to extract conjugate polygons.
- Dynamic programming matching technique: edge-based to extract conjugate points.

These techiques were discussed in detail in the last two subsections. Both techniques operated independently, but complemented each other to automate the process for measurement required for registration. The patch matching technique provided the location of corresponding polygons and restricted the area for dynamic programming to perform efficiently. It also provided the points for the first approximate transformation between the map and the image. The dynamic programming performed a refinement of the transformation. The combined effect of both techniques produced the end result of automatic extraction of a list of conjugate map and image coordinates, which will be used in the next chapter for automation of the registration of the images to their corresponding maps.

3.8 Result of best matched image polygons with their map model

In the last section, the patch matching technique integrated with region growing segmentation was shown. Therein, the combination of both these processes extracted the best matched image polygon (BMIP) for each map polygon. This section is concerned with the assesment of the quality of the best extracted image

polygons. The matching results of the two test areas, Elchingen-Grosskuchen and Lords's Cricket Ground are shown here in Table 3.12 and Table 3.13 respectively.

Map polygon number	Matching value of conjugate image polygon in region growing frame number							
	3	4	5	6	7	8	9	10
1	1.155973	1.031303	0.993355	0.909030	0.948099	0.844345	**0.771521**	0.794134
2	2.382606	2.379784	2.247189	**2.052954**	2.309897	2.298441	2.586882	2.618578
3	2.240531	1.463285	1.532215	1.358915	1.340347	1.296611	1.333367	**0.920889**
4	23.298459	22.985949	1.828143	**1.800120**	1.850853	1.825730	7.044445	19.298873
5	1.075941	**0.789884**	1.343449	1.638953	1.248596	6.809837	3.652149	5.444261

Table 3.12: Best matched map-image polygons of Elchingen-Grosskuchen area.

Map polygon number	Matching value of conjugate image polygon in region growing frame number							
	3	4	5	6	7	8	9	10
1	0.652735	0.597061	0.590272	0.588750	0.588913	0.588557	0.586674	**0.086367**
2	0.088574	0.024507	0.024345	0.024470	0.024181	**0.023993**	0.024017	0.024892
3	0.049051	0.048961	**0.046626**	0.048350	0.067456	0.049986	0.047967	0.0.50152
4	0.035318	0.035018	**0.034232**	0.039190	0.038949	0.068349	0.106908	0.110699
5	0.037394	0.036389	0.035694	0.036140	0.035713	0.035588	0.035575	**0.035027**
6	**0.040202**	0.041953	0.044735	0.040818	0.041281	0.041082	0.096910	0.042442
7	0.077615	0.075871	**0.075692**	0.077986	0.079521	0.097877	0.095933	0.095033

Table 3.13: Best matched map-image polygons of Lords's Cricket Ground area.

In these tables, a matching value for each map polygon matched with its corresponding image polygon is shown for each segmented image. In the last section this was illustrated by a diagram (see Figure 3.17), which can be compared with Table 3.12 in this section. It can be seen that each map polygon has a minimum match value (shown in bold) in one of the segmented images. The corresponding image polygon for each map polygon which shows a minimum value is the best matched image polygon.

Insufficient growth of selected areal features in the region growing process sometimes causes their non-extraction in a few segmented images. This is because the neighbouring pixels are not merged to the growing conjugate polygon due to the

low threshold value applied (for details see Section 2.4.3). In such cases, that is in the absence of the conjugate image polygon, the map polygon is matched with another polygon in the image. Such matches show high matching values, as shown in the above Tables (underlined matched values).

Remarks: In this study, for each map polygon the best matched image polygon was extracted by patch matching. That allowed the extraction of a large part of the boundary pixels from the boundaries of the best extracted image polygons on matching with the map model boundaries (as shown in Figure 3.19, Figure 3.20 and Figure 3.21). The extraction of a large number of boundary pixels is important for registering the image to the map with high accuracy.

If extracted matched polygons are not the best matched polygons, as described above, then in such a case the boundaries of the image polygons may not be correctly matched to the map model boundaries and will not provide a large redundancy for matching. This may also cause a higher degree of error in registering the images to their map model.

3.9 Conclusions

In this chapter, the extraction of real world objects from images by automatic matching with a map model was investigated. The extraction of control features from images is the main problem. The steps to solving this problem were studied and the results are summarised below.

- extraction of corresponding objects in map and image
- determination of the shape attributes of extracted objects (areal features)
- finding the scale and translation difference between map and image, using the centroid of extracted objects
- finding of conjugate map-image polygons using the shape attributes with the help of scale and translation difference

- extraction of corresponding matched points from the boundaries of conjugate polygons

These steps performed the preparation of map and image for matching, and final matching. Two matching techniques were combined to achieve the correspondence between map and image. The first finds the corresponding polygons applying a patch matching procedure. The second extracts the conjugate points from the corresponding polygons using a dynamic programming. Both these techniques have been used by previous researchers, but no solution has previously been found to extract control features from images automatically. In the present work, this was obtained by the use of the integration of an iterative region growing segmentation of images with patch matching.

In the first instance, each extracted feature in the segmented images was given a chance to be a control feature candidate. The selection of the control feature was decided by the patch matching on the basis of a cost function. The shape and size of the extracted features were used in the patch matching to compare with those of the map model to select the control features. The feature with the closest resemblence (in shape and size) to the map model were selected the control features. This was illustrated with figures in this chapter.

The role of the centroid has played an important part in patch matching by identifying conjugate polygons of map and image. The centroid alone enabled the determination of the scale and translation differences between the map and image. Both these differences were used in patch matching and showed correct matches and automated the patch matching. Previous workers faced difficulties in obtaining all the matches correctly in patch matching and that hindered the automation of patch matching. Besides this, they used patch matching in a limited way, only for images of similar scales. But the new developement in this study has enabled the advances in patch matching to be used for images of different scales.

Another advancement in patch matching achieved in this work is that it can identify objects of the same shape, size and orientation in two images after initial selection of approximate positions. The position of objects can be located and identified in images by the use of their centroids. This is of real benefit in high

resolution images of urban areas where many buildings may have the same size and orientation. In such images, the method described in this study would be able to correctly match buildings automatically.

The techniques discussed in this chapter showed satisfactory results in the automatic extraction of areal control features from images as well as determining conjugate matched points for the registration. The key features achieved in this study are:

- patch matching can be used for images of different scales, and
- for the automatic identification of conjugate polygons.

It can be concluded that unless the extracted features of the image have a close relationship (matching value) in shape, i.e. shape similarity with their corresponding map features, it is difficult to extract the large number of conjugate points which is necessary for high accuracy map-image registration.

Chapter 4

Elimination of Errors for Image Registration

4.1 Introduction

Updating maps is performed by finding differences in ground features. Satellite images and aerial photographs can be used for this purpose. The correct differences in surface features can only be gained by registering images to maps with precise geometric corrections. This chapter deals with geometric errors in images and suggests a method to remove such errors completely or minimise them for precise map-image registration.

In photogrammetry, rectified photos are treated as truly vertical photos, free from tilt displacements. However, they contain image displacements due to the topographic relief. These relief displacements can be removed in a process called differential rectification or orthorectification, and the products obtained are called orthophotos. This process needs an experienced operator and, moreover, it is time-consuming. Nevertheless, the method is used widely, although it is not economically beneficial for updating maps because it involves a large amount of work, even for making *small* changes. A different method, which is automated and economically viable for registering images to maps, is developed in this study. It is based on a statistical approach.

When matching maps and images, the relief objects produce large errors. It is necessary to remove these errors to achieve precise image registration. The statistical method developed in this work removes the large errors. In the last chapter, match points were derived by matching extracted image polygon boundaries to their map model boundaries. Some of the match points belonged to real world objects and others are mismatched or distorted by the perspective view, shadows or occlusions in the images. Separation of the two types of match points is the focus of this work, as well as the selection of match points which belong to real world objects. The new work is introduced in this chapter.

The issues discussed in this chapter are: the statistical approach to remove errors, the strategy for precise image registration, the process of image registration, the elimination of errors, and results obtained. Finally, concluding remarks are made on the precise map-image registration process.

4.2 Use of sample data in statistics

In statistics, a population is the whole set of measurements about which we want to draw a conclusion, whereas a sample is a sub-set of a population, a set of some measurements which comprise the population. The properties of normal distribution have a very important use in the statistical theory of drawing conclusions from sample data about the populations from which the samples are drawn. Choosing samples and drawing conclusions about the population have two benefits:

- it saves time, money and effort; and
- even if all the information about the population is not available then the sample data can be useful in drawing conclusions about the population.

A graphical represention of the normal distribution has a bell-shape distribution with most values concentrated in the centre and a few at the extremes. The distribution is unimodal with symmetry, as shown in Figure 4.1. It has two parameters, mean (μ) and standard deviation (σ).

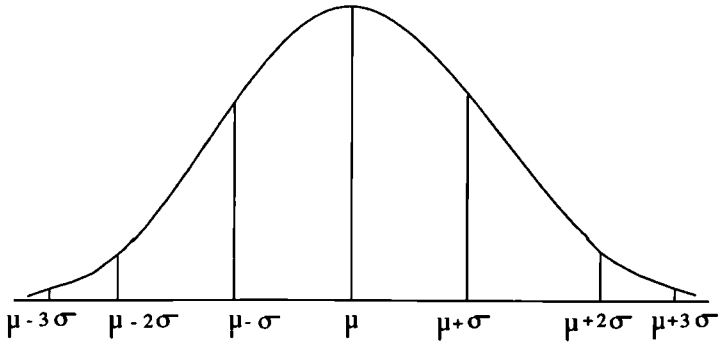


Figure 4.1: A statistical model for the removal of large errors.

There are a number of related properties of normal distribution which give us a better understanding of the meaning of standard deviation as a measure of variation because, for example, approximately 68% of the area of a normal distribution lies within one standard deviation of the mean. In Figure 3.1, this is shown between the vertical lines drawn at $\mu - \sigma$ and $\mu + \sigma$, which is roughly two-thirds of the total area. The two standard deviations and the three standard deviations of the mean contain an area which is nearly 95% and 99.7% of the total.

The methods used to calculate the mean, standard deviation, and other basic statistical parameters, are not discussed in this section because it is not felt necessary. They can be found in any basic book on statistics, such as Rowntree (1981) and Rees (1989).

Here, in this chapter, the main concern is whether the sample data chosen belongs to a single population or not. If the sample data is not from a single population, then the question arises how to separate a sub-set from the sample data which belongs to a single population. For example, a list of the radii of five thousand lemons is given in which the radius of five hundred cherries and five hundred apples are included in error (mistake). A statisical approach of mean and standard deviation can be used to find sample data (sub-set) which contain only lemons, i.e. to find a sub-set which belongs to a single population. Assume that the average size of a lemon is greater than a cherry and smaller than an apple. In the normal distribution curve, the radius of lemons will concentrate in the centre, the cherries at the left-tail and the apples at the right-tail portions of the curve.

To remove the cherries and the apples from the sample, deselect part of the sample data, say 2.5% from both ends and select $\pm 2\sigma$ of the mean. However, there is still a chance of having a few cherries and apples in the selected sample sub-set. Again, the same process of selecting a sub-set from the selected sample sub-set will remove both, the cherries and the apples, and increase the probability of selecting a sample set of a single population. This idea is adopted and used iteratively in this study to remove errors for a high accuracy of map-image registration and is explained in detail with experiments in this chapter.

4.3 Strategy for precise image registration

Objects in images show geometric distortions, which change their shape and size. The causes, in general, of geometric distortion of objects in images are sensor tilt and terrain relief. In the last chapter, it was shown that matching images to their model produces a large number of match points. Not all matches are true object match points; they include points which are geometrically distorted. The distorted match points are the product of perspective distortion, shadows and occlusions. On registering the images to the map model, the distorted match points produce large errors. The main concern of this chapter is the removal of such distorted match points in order to obtain precise image registration.

Some distorted match points may be associated with tall objects like buildings; some may be the result of occlusion by objects like trees, and some may be the product of shadows; they are all distorted match points and will produce large errors. These errors are designated as systematic errors and tend to show patterns. In fact, they are mismatch points and do not exist in the map model. They belong to a different population than that contained in the model.

The magnitude of errors of the match points can be considered as sample data, and treated statistically, as discussed in the last section, to eliminate match points with large errors. To remove the distorted match points, a strategy is outlined in Figure 4.2.

The method works to select a sample set of match points which are of the map

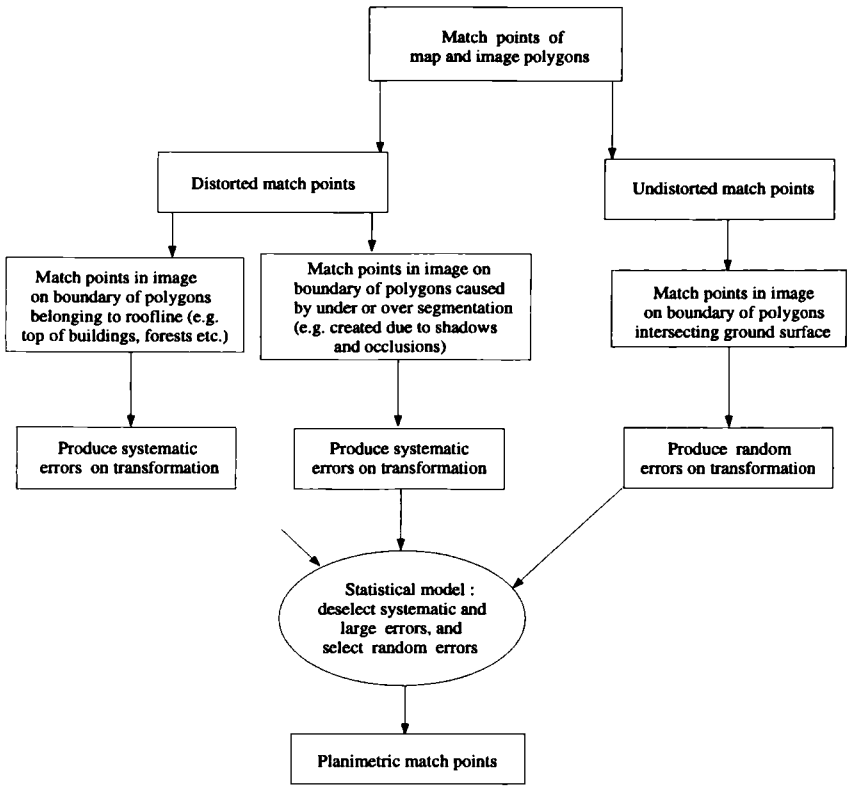


Figure 4.2: Schematic diagram for the behaviour of distorted match points.

model population. To remove the errors, the mean and the standard deviation of the errors are obtained and large errors removed by selecting a sample set of $\pm 2\sigma$ of the sample data. This process is repeated until the sample set selected belongs to the single population, i.e. the standard deviation of the current and the preceding process becomes the same. The selected sample set is left with random errors only and each measurement in it has the same probability of being selected as a good match point and truly belonging to the map model population. Method of obtaining a sample of random errors is targeted in the present work.

4.4 Image registration

To obtain absolute image orientation or map-image registration, the image coordinates are related to the ground coordinates. High quality maps are required for this purpose. Digital elevation models (DEMs) and sensor orientation parameters

(camera model) are essential if the terrain contains tall objects or if it presents high relief. According to Toutin (1995), photogrammetric techniques originally developed for the rectification of aerial photographs, using a DEM and camera model, can generally be extended to satellite images. But, instead of using a digital elevation model (DEM) and a camera model in this study, a statistical model was used to remove the effect of heights in order to obtain precise image registration.

Before discussing the model, it is better to explain the basic methods which are used for image registration. Two subsections are included in this section. They deal with image registration of high and low relief areas.

4.4.1 Image registration of high relief areas

A digital elevation model (DEM) of the ground with the orientation parameters of the sensor are required for rigorous geometric corrections for the image registration of high relief areas. To model the corrections it is first necessary to understand the effect of distortion due to high relief terrain on satellite images.

The effect of high relief on images In practice, the orientation of the sensor is never parallel to the reference surface. Moreover, the surface may not be flat. In such a situation, the relief effect causes a change in the scale throughout the image. For this reason, positions of the image object are displaced from the position at which they should appear if there was no such relief. In aerial photographs, the displacements of the objects are radial from the nadir and are maximum at the border of the photograph. In satellite images, the displacements depends on the look angle of the sensor, and the orientation of the objects to it. The apparent effect of the displacement in satellite images is less than in aerial images due to a great difference in the altitude of their platforms. For example, in Figure 4.3, a point P on the ground surface appears at point p on an aerial photograph due to high relief. The same point will appear at point p'' if the relief is absent. The displacement due to relief can be calculated by the following equation.

$$pp'' = pn * \frac{h}{H} \tag{4.1}$$

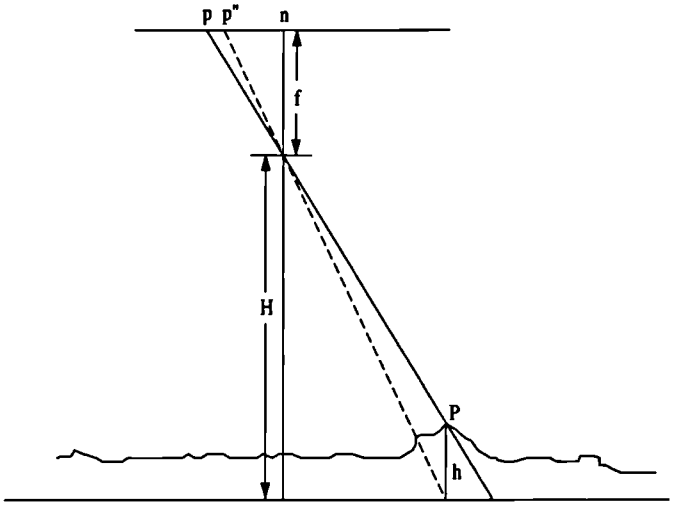


Figure 4.3: Displacement of object on an image due to relief.

On satellite images, the effect of the displacement of objects is less because the altitude of a satellite, H in equation 4.1, is large. In general, in vertical high resolution images, the geometry is strong enough for heights to show perspective distortion, but, in low resolution images, the effect of relief distortion is negligible (Forshaw et al., 1983). For example, a tall building, like the British Telecom Tower in London, will show some effect of displacement on DD5 images, but not on Landsat images. That is the reason high resolution images need corrections for high relief terrain.

In terms of stability, satellite platforms are more stable than lower altitude platforms. This makes it easier to orientate any imagery taken from a satellite as there are fewer effects to consider from the attitude of the platform than from an aerial platform [Paderes and Mikhail (1983)]. This allows a reduced number of orientation elements to reconstruct the satellite platform movement at the time of imagery acquisition. Moreover, satellite orbital parameters also can be used to describe the position and movement of satellites (Lee et al., 1993).

Models to correct the effect of relief For registering high resolution images or aerial photographs to high relief ground, the process of extracting three dimensional co-ordinates from images is required. The measurement of three dimensional co-ordinates in such images must include the determination of the ori-

entation and the position of the sensor in order to remove the distortion caused by the attitude of the sensor. In photogrammetry, inner, relative and absolute orientation processes are used to correct the attitude distortion (Morgado, 1996). Therein, the position of points on the image is related to the points on the ground through a mathematical model using stereoscopic images for the third dimension. The same technique is used for satellite images. The mathematical model takes into account orientation parameters to produce orthoimages and for this it needs few measurements, i.e. the correspondence between map and image to solve the problem of absolute orientation (registration) of an image.

To perform absolute orientation, i.e. a 3D conformal transformation, requires at least 3 sets of xyz model coordinates (from a stereo-model) with corresponding XYZ ground coordinates. The equation for three-dimensional coordinate transformation (absolute orientation) is represented by equation 4.2.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = sM \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \tag{4.2}$$

Where X, Y, Z are the ground coordinates; x, y, z are the model coordinates; s is the scale factor between both coordinate systems; and M is the rotation parameter between the two coordinate systems which can be expressed by following equation.

$$M = \begin{bmatrix} \cos\kappa\cos\phi & \cos\kappa\sin\phi\sin\omega + \sin\kappa\cos\omega & -\cos\kappa\sin\phi\cos\omega + \sin\kappa\sin\omega \\ -\sin\kappa\cos\phi & -\sin\kappa\sin\phi\sin\omega + \cos\kappa\cos\omega & \sin\kappa\sin\phi\cos\omega + \cos\kappa\sin\omega \\ \sin\phi & -\cos\phi\sin\omega & \cos\phi\cos\omega \end{bmatrix} \tag{4.3}$$

The equation 4.2 has seven unknowns which are independant transformation factors: three rotation angles, omega (ω), phi (ϕ) and kappa (κ); three translation factors T_x, T_y and T_z; and a scale factor s. The detailed explanation of the three-dimensional conformal transformation is given by Wolf (1983); Ghosh (1988); Moffit and Mikhail (1980).

Nowadays, map products derived from satellites are proving to meet most of the required specifications [Gugan and Dowman (1988); Hartley (1988)] as well

as improvements in cost saving for mapping [Hartley (1991); Hoffmann (1992); Lenzen (1992)]. Examples are the mapping of Djibouti at 1:50,000 and 1:200,000 scales [Veillett (1992)] and of North East Yemen at 1:100,000 [Murray and Newby (1990)].

4.4.2 Image registration of low relief areas

When terrain is flat or nearly flat and relief can be ignored then two basic steps of image transformation are necessary for map-image registration. They are described below.

2D transformation: This procedure requires polynomial equations to be fitted to the control point data, using least-squares criteria, to model the corrections directly in the image domain without explicitly identifying the source of the distortion. For moderate distortion, as in a flat or nearly flat area, a six-parameter affine transformation is sufficient to register the image to the map. This type of transformation (affine) can model six kinds of distortion in the remote sensor data: translation in x and y, scale changes in x and y, skew and rotation (Billingsley, 1983). These six operations, combined into a single expression, can be written as:

$$x' = a_1 + b_1 x + c_1 y$$

$$y' = a_2 + b_2 x + c_2 y$$

where x and y are positions in the map, and x' and y' represent corresponding positions in the transferred image. The use of these six coordinate coefficients will transform (relocate) pixel values from the original distorted image [$x_{original}$,$y_{original}$] to the grid of the map [x,y]. However, before applying the rectification to the entire image data set, it is important to determine how well the six coefficients derived from the least-squares regression of the control points accounts for the geometric distortion in the original distorted image. Any discrepancy between the values, such as (x'-$x_{original}$) and ($y' - y_{original}$), is called $x_{residual}$ and $y_{residual}$ respectively for each point. This discrepancy represents the image geometric distortion which is not corrected for each point by the six-coefficient coordinate transformation.

A simple way to measure such distortion is by computing the vector of residual

for each control point and calculating total RMS_{error} using the following equations:

$$Residual_{vector} = \sqrt{[(x' - x_{original})^2 + (y' - y_{original})^2]} \tag{4.4}$$

where $x_{origional}$ and $y_{original}$ are the original row and column coordinates of the control points in the image, and x' and y' are the computed or estimated coordinates in the transferred image. Calculating $Residual_{vector}$ for all control points will show which points exhibit large errors. They will also represent a measure of accuracy for each point.

The total RMS_{error} is calculated by first computing the X RMS_{error} and Y RMS_{error}.

$$R_x = XRMS_{error} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{residual})_i^2} \tag{4.5}$$

$$R_y = YRMS_{error} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{residual})_i^2} \tag{4.6}$$

$$Total\,Root\,Mean\,Square\,error = RMS_{error} = \sqrt{R_x^2 + R_y^2} \tag{4.7}$$

where n is the number of control points.

Intensity interpolation: Resampling this process involves the extraction of a brightness value from x', y' location in the transferred image and its relocation to the appropriate x,y coordinate location in the registered output image. The practice of brightness value interpolation is commonly referred to as *resampling*. Most of the time, the x' and y' coordinates to be sampled in the transferred image are real numbers (i.e., they are not integers). To solve such a problem there are several methods of brightness value interpolation that can be applied, including nearest-neighbour, bilinear interpolation, and cubic convolution. It has been seen that the quality of the output images increases from nearest-neighbour to cubic convolution, but there are increases in relative computation time. Bilinear interpolation seems to be a good compromise for a wide range of applications.

4.5 Elimination of errors

In the last chapter it was stated that a large number of match points (conjugate points) can be produced by matching a map and an image of an area. It is possible, however, that many of the match points may show large errors on applying a transformation technique. The large errors are mostly produced by the perspective distortion of elevated objects, shadows and occlusions in images. A precise registration of the images to the corresponding maps is not possible unless the large errors are removed from the images.

Five subsections are included in this section. The first subsection describes how perspective distortion in the high resolution images is created and produces large errors when matching images to map models. The second subsection introduces a technique for the elimination of errors. The third subsection shows a model for the removal of perspective distortion to select planimetric match points for map-image registration. The fourth subsection deals with the validation of this model by the application to the match points of the three test areas obtained in the last chapter. The last subsection makes a comparison between the photogrammetric methods and statistical model for image registration.

4.5.1 Perspective distortion in images

In 2D images, a three-dimensional real world object always exhibits a loss of information as well as incorrect geometric information due to perspective distortion. The magnitude of distortion is related to the camera (sensor) look angle and height of object. This can be illustrated by a diagram. In Figure 4.4, the projection of a 3D object into a 2D image shows a shaded area which is created due to perspective distortion. This particular shaded portion of the image will produce systematic errors when matched with its map model.

The effect of distortion on matching: After matching images to their maps, it is important to understand the effect of the perspective distortion on the extracted match points. In Figure 4.3, it was shown that higher elevated objects showed

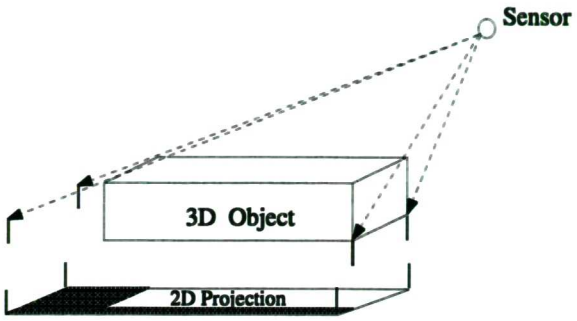


Figure 4.4: A model of the perspective distortion.

larger distortion in the images. A small test is performed to illustrate this. A Farnborough area was chosen for the test. The area contains large buildings whose

Figure 4.5: A section of a DD5 image of Farnborough area.

height information is available, varying between 3.2 to 9.5 metres. A DD5 image and Ordnance Survey raster map data were selected for the test. A section of 230 x 180 pixels from the DD5 image (see Figure 4.5) and a section of 900 x 500 pixels from the Ordnance Survey 1:10,000 map (see Figure 4.6) were used for the test. Large buildings from the image and the map were extracted for matching [for detail see Vohra and Dowman (1996)]. The matching produced 393 match points. Using matched points for transformation showed a different magnitude of errors for different heights of buildings. This is shown as a graph in Figure 4.7, where the magnitude of the errors (residual vectors) of the match points (which belong

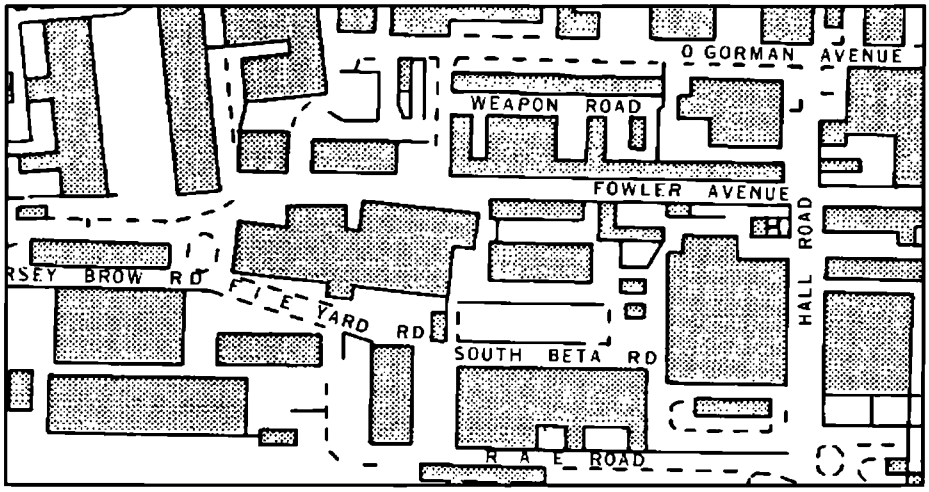


Figure 4.6: Ordnance Survey 1:10,000 map of the Farnborough area.

to buildings of different heights) are plotted against frequency. The magnitude of errors was checked with the known building heights and it was noted that the large errors are associated with the tall buildings [Vohra and Dowman (1996)]. The differences in building heights caused clustering in the errors. The clustering is caused by the roof line of the buildings being extracted from the image and matched with the building line of the map model. In the Figure, a portion of the graph shows a stepwise curve reflecting the clustering effect. A model is required to remove the match points of the roof line (which causes perspective distortion), shadows and occlusions, and to select only the points of the building line, i.e. planimetric match points, in order to achieve a high accuracy of map-image registration.

4.5.2 An overview of a technique for the elimination of errors

The feature extraction and the matching processes were conducted on the three test areas, as described in the last chapter. Both processes were performed before any geometric correction had taken place. The distortion due to terrain variation as well as due to the height of the objects were still present in the matching result, which will cause errors in the image registration. This could lead to complete mis-registration. The removal of errors is the interest of the present work.

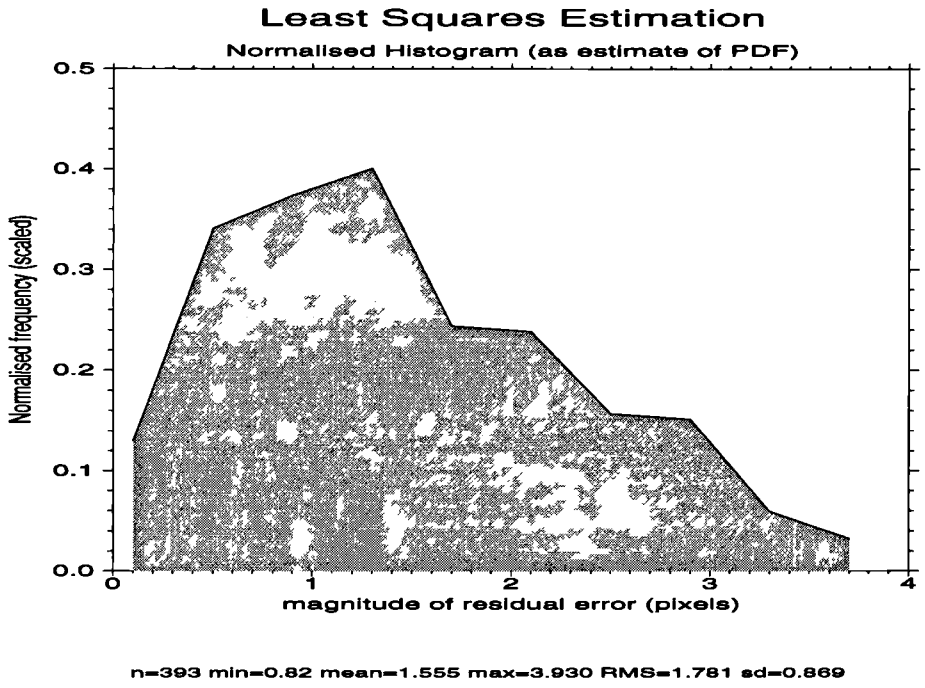


Figure 4.7: Affine transformation residual vectors vs. frequency.

An idea is developed in this study to collect correct information from the images and to remove information which produces large errors. The idea came after viewing the match points of the undistorted and distorted (false) edges of the extracted segmented objects in images. The idea can be illustrated by a diagram to select match points of the undistorted edges in images. Figure 4.8 shows segmented objects after segmenting a satellite and an aerial images. It can be seen that the undistorted edges of the segmented objects are related to the position of the sensor. A process which selects the match points on the undistorted edges shows a good result for image registration. The method of selecting undistorted match points will deselect automatically the match points which belong to distorted edges. This approach is considered in this study.

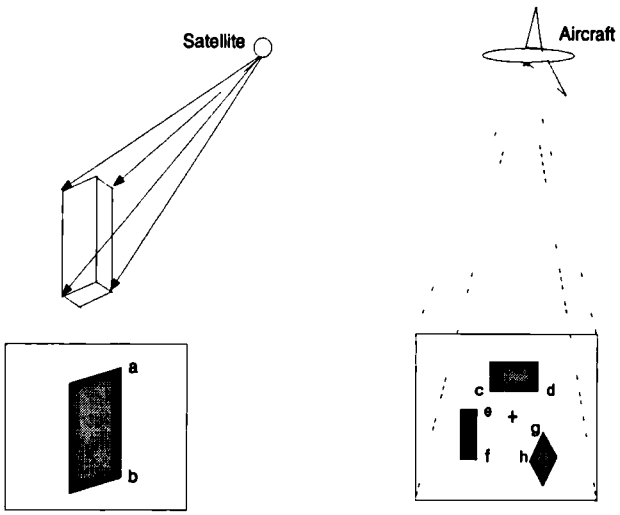


Figure 4.8: A model for the removal of perspective distortion.

4.5.3 A statistical model to select planimetric points in images

A statistical model that rejects distorted match points, and selects the match points of undistorted edges automatically, is developed in this study.

The match points which were obtained in the previous chapter are tested in this section. Each match point shows a displacement (residual) error in X and Y directions on applying a transformation. The residuals have positive as well as negative real number values, as shown in Figure 4.9. The method of the normal distribution explained in section 4.3 is applied on the residual errors. The mean value of both the X and Y residuals shows zero; and thus the RMS (root mean square) error and the standard deviation [sigma (σ)] have the same values. In such conditions, 95.45 percent of the sample population will lie within $\pm 2\sigma$ [Harshbarger (1971) and Triola (1989)]. The RMS error of the match points in X and Y directions and the total RMS error can be calculated by using equations 4.5, 4.6 and 4.7 respectively.

As discussed before in section 4.3, the model developed rejects match points

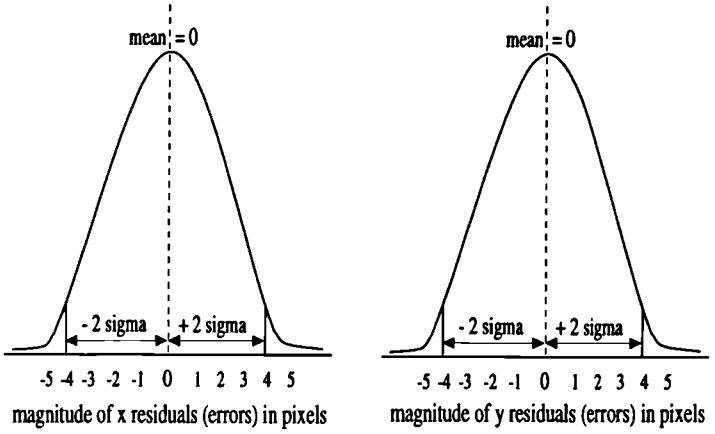


Figure 4.9: A statistical model for the removal of large errors.

which produce residual error values in the X or Y direction greater than their $\pm 2\sigma$ on transformation. The selected match points with errors within $\pm 2\sigma$ in the X and Y direction are used for transformation again and the RMS error and the standard deviation are recalculated. This process is performed iteratively until the RMS error difference of the current and the preceding selected match points becomes negligible.

The model is formulated with the following equations.

$$X_{error} \geq -2\sigma \; and \; X_{error} \leq +2\sigma \tag{4.8}$$

$$Y_{error} \geq -2\sigma \; and \; Y_{error} \leq +2\sigma \tag{4.9}$$

$$RMS_i - RMS_{i-1} = 10^{-4} \tag{4.10}$$

X_{error} and Y_{error} are the magnitude of the displacement (residual) errors of each match point in X and Y directions obtained by applying a transformation. RMS_i is the root mean square error of the match points of the iterative i^{th} (i=1,2,3,....) transformation. In equation 4.10 the value 10^{-4} is chosen so that the difference between the standard deviation of the last two iterations is negligible. This situation will be a situation where the remaining match points will belong to a single population, signifying the removal of systematic errors as explained in Section 4.3. These three equations were used to develop the algorithm **planimetric** for the statistical model for the removal of large errors.

4.5.4 Validation of the statistical model

For the accuracy of measurements in the fields of photogrammetry and surveying the terms of gross error, systematic error and random error are often used. In this study the relevant terms, i.e. systematic and random error, are only used. Various automatic methods for dealing with gross error within the least square adjustment have been determined. Such methods, often termed 'robustified LSE' are detailed in Caspary (1987) and Cooper and Robson (1996).

The statistical model algorithm was validated on the match points of the three test areas which were derived in the previous chapter. The results and analysis of these test areas are discussed below.

London Regent's Park area (UK)

Three features were extracted from the map and the image of Regent's Park for matching: a cricket ground, a lake and a building. Matching these polygons produced 865 match points. The planimetric algorithm was applied to these match points and works iteratively. Each iterative step of the algorithm removes those match points whose X or Y residual errors are greater than their $\pm 2\sigma$ i.e. match points with large errors. In the first iterative step, the algorithm uses all match points without removing any match point and in further steps it removes the match points with large errors. In all iterative steps, the algorithm calculates statistics. The results are shown in Table 4.1. The Table shows that 24 iterative steps were conducted on the match points and, in the final step, the algorithm selected 376 match points. The last iterative step shows that the difference between its RMS error and the preceding one is zero and there is no change in the number of match points. This indicates that further application of the algorithm on the last iterative step will not remove any match point. These match points were used for obtaining the transformation parameters to register the image with map.

To illustrate the content of Table 4.1, a graphical diagram is shown in Figure 4.10. Figure 4.10 (top-left) shows the range of X residual values (with upper and lower value limits as a vertical bar) of each iterative process which reduces

Iteration No.	X MIN	X MAX	X RMS	Y MIN	Y MAX	Y RMS	T RMS	RMS_{i-1} - T RMS_i	Match points
1	-10.240	2.660	1.001	-2.840	2.550	0.962	1.388320	1.388320	865
2	-1.980	2.060	0.774	-1.890	1.990	0.779	1.098140	0.290180	767
3	-1.530	1.530	0.653	-1.570	1.570	0.704	0.960221	0.137919	694
4	-1.300	1.340	0.587	-1.380	1.400	0.659	0.882525	0.077696	635
5	-1.180	1.200	0.524	-1.320	1.310	0.639	0.826376	0.056149	593
6	-0.990	1.090	0.470	-1.290	1.260	0.623	0.780403	0.045973	555
7	-0.900	0.940	0.433	-1.230	1.250	0.606	0.744799	0.035604	525
8	-0.860	0.860	0.414	-1.200	1.230	0.592	0.722399	0.022400	508
9	-0.810	0.830	0.403	-1.160	1.230	0.573	0.700527	0.021872	490
10	-0.790	0.820	0.392	-1.120	1.190	0.551	0.676214	0.024313	472
11	-0.770	0.780	0.383	-1.070	1.120	0.529	0.653093	0.023121	456
12	-0.740	0.760	0.374	-1.030	1.060	0.515	0.636475	0.016618	442
13	-0.740	0.750	0.369	-1.000	1.030	0.503	0.623835	0.012640	432
14	-0.730	0.740	0.365	-0.980	1.010	0.493	0.613412	0.010423	423
15	-0.730	0.730	0.358	-0.980	1.000	0.488	0.605234	0.008178	416
16	-0.710	0.690	0.350	-0.970	0.970	0.477	0.591632	0.013602	406
17	-0.700	0.680	0.346	-0.940	0.970	0.466	0.580407	0.011225	396
18	-0.700	0.670	0.341	-0.930	0.940	0.458	0.571004	0.009403	390
19	-0.690	0.680	0.337	-0.890	0.900	0.446	0.559004	0.012000	382
20	-0.660	0.660	0.335	-0.880	0.900	0.444	0.556202	0.002802	379
21	-0.670	0.660	0.334	-0.880	0.880	0.442	0.554004	0.002198	378
22	-0.670	0.660	0.333	-0.880	0.880	0.442	0.553401	0.000603	377
23	-0.660	0.650	0.331	-0.880	0.880	0.441	0.551400	0.002001	376
24	-0.660	0.650	0.331	-0.880	0.880	0.441	0.551400	0.000000	376

Table 4.1: Iterative statistical removal of erroneous match points from the map-image of Regent's Park area (units in pixels).

progressively in each iterative step from the first iteration to the last one. Each iteration removes bad match points and with the number of match points left, the X RMS error of that iterative process is calculated. The X RMS of all iterations is shown in Figure 4.10 (top-right). It shows that the accuracy of the X RMS increases with each iterative step. A similar effect is observed for the range of Y residuals and Y RMS, which is shown in figure 4.10 (middle-left) and Figure 4.10 (middle-right) respectively. Figure 4.10 (bottom-left) shows the relation between the iteration process and the match points. It can be seen that, with the increase in iterative steps, the number of match points decreases, filtering the match points with large errors. Figure 4.10 (bottom-right) shows the total RMS accuracy of each iteration. This figure infers that the use of the statistical model algorithm iteratively filters large errors which might be caused by the perspective distortion, shadows of objects and occlusions. It can also be seen, from the figure and from the table that after the 20th iteration, there is no significant removal of match points and change in the RMS values. This informs that the remaining match points after

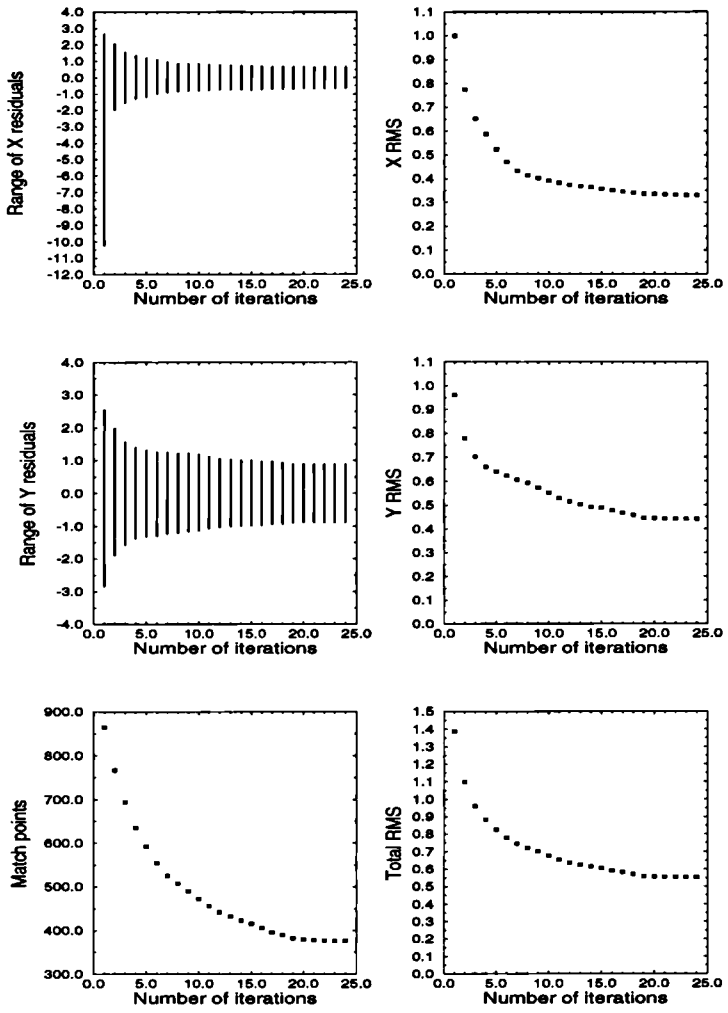


Figure 4.10: A graph representing statistical errors for matching map and image of the Regent's Park area.

this iteration belong to a single population.

- **Statistical analysis of iterative processes:** The above process used the statistical model algorithm to remove match points which were associated with large errors in the X or Y direction, in each iteration. Therein it was not mentioned how much error each match point contributed to the total RMS error. Error contribution of each match point can be calculated by computing the residual vector, which is shown in equation 4.4. Table 4.2 shows statistics for the residual vectors of each iteration. The effect of the residual vectors in terms of randomness or systematic errors of the match points can be shown by a plot, for each iteration. Three iterative processes

Iteration No.	Match points	Residual vector minimum	Residual vector mean	Residual vector maximum	T RMS of residual vectors	Standard deviation
1	865	0.020	1.152	10.383	1.389	0.775
2	767	0.022	0.961	2.730	1.098	0.531
3	694	0.010	0.856	1.902	0.960	0.435
4	635	0.010	0.788	1.789	0.882	0.397
5	593	0.032	0.740	1.677	0.827	0.369
6	555	0.036	0.698	1.564	0.780	0.348
7	525	0.040	0.667	1.496	0.745	0.332
8	508	0.030	0.648	1.426	0.722	0.319
9	490	0.014	0.630	1.440	0.701	0.307
10	472	0.032	0.610	1.379	0.676	0.291
11	456	0.063	0.592	1.273	0.653	0.276
12	442	0.054	0.577	1.202	0.636	0.267
13	432	0.051	0.567	1.162	0.624	0.261
14	423	0.040	0.557	1.170	0.613	0.256
15	416	0.050	0.550	1.187	0.605	0.253
16	406	0.045	0.538	1.124	0.592	0.246
17	396	0.054	0.528	1.140	0.581	0.242
18	390	0.050	0.520	1.160	0.572	0.238
19	382	0.045	0.509	1.056	0.559	0.232
20	379	0.050	0.506	1.054	0.556	0.231
21	378	0.036	0.504	1.062	0.554	0.229
22	377	0.036	0.504	1.062	0.553	0.229
23	376	0.036	0.502	1.054	0.552	0.229
24	376	0.036	0.502	1.054	0.552	0.229

Table 4.2: Statistics of residual vectors of map-image match points for the Regent's Park area (units in pixels).

were considered: first, intermediate and last, i.e. 1st, 7th and 24th, to display the effect of the residual vectors and their magnitudes. The plots are shown in Figure 4.11.

The residual vectors on the plots are magnified 50 times to show their effect. The first plot, [Figure 4.11 (top)] with the residual vectors of 865 match points, displays large errors in comparision to the two other plots shown in the Figure 4.11. The second plot, [Figure 4.11 (middle)] with 525 residual vectors, shows that many match points with large errors which showed up in the first plot have been removed. The last plot, [Figure 4.11 (bottom)] with residual vectors of 376 match points, show small errors in magnitude. The largest residual vector error in magnitude of this plot shows an error of 1.054 pixels and the total RMS error is 0.552 pixel of 376 match points, whereas the first plot shows the largest residual vector error of 10.383 pixels with a total RMS error of 1.389 pixels, and the second plot shows the largest residual vector of 1.496 pixel with a total RMS error of 0.745 pixel. In the

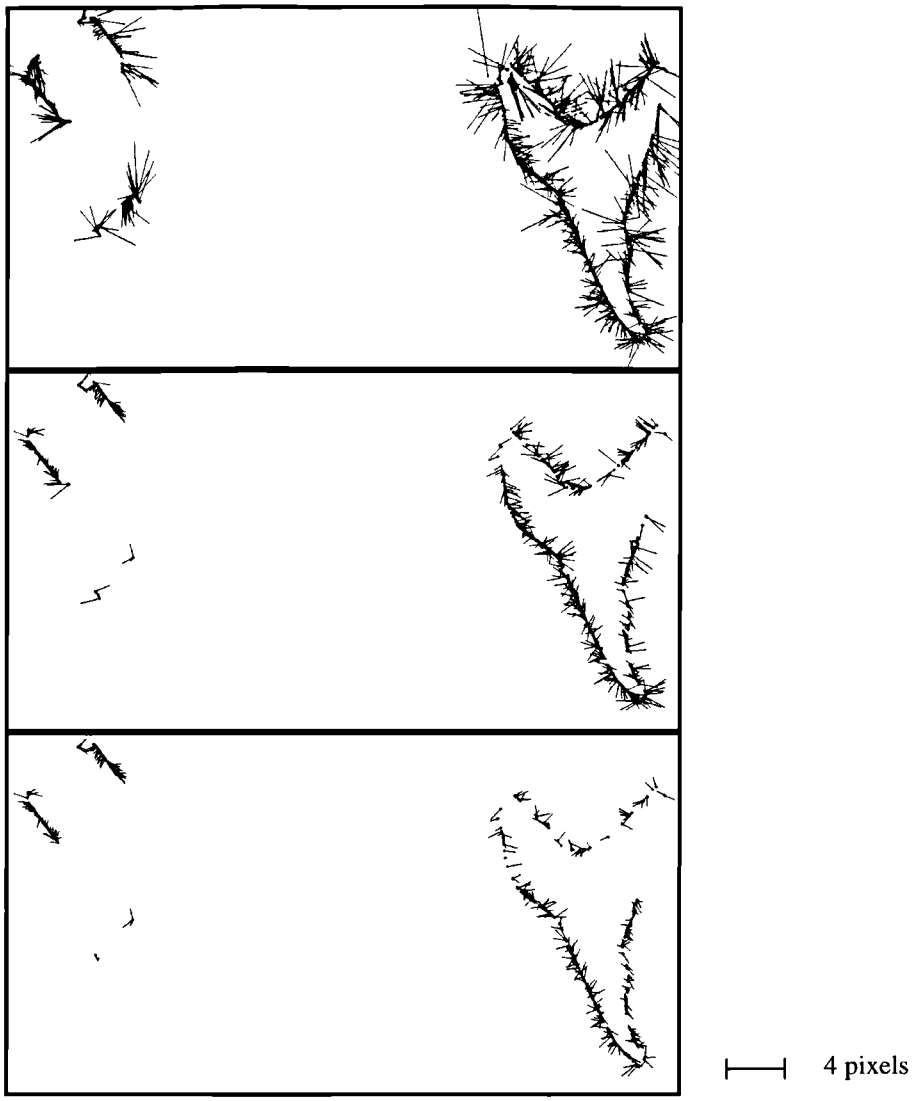


Figure 4.11: A plot of residual vectors magnified 50 times: 1st iteration (top), 7th iteration (middle) and 24th iteration (bottom) of the statistical model algorithm.

first plot, many residual vector errors are bunched into several groups showing patterns in different directions. These groups are caused by systematic errors, as explained before in section 4.3. For instance, upper righthand part of the lake boundary is completely covered by the trees, it shows systematic errors. The algorithm has removed all these errors which are caused by occlusion. Similarly, the Cricket Ground boundaries show systematic errors due to shadow effect of stadium, and the building due to its own height. They are removed by the algorithm.

The iterative process of the statistical model algorithm has shown that the processing filters out the large errors and selects match points of random errors. These match points were used to obtain transformation parameters for registration.

- **Validation of registration by a plot of planimetric points:** The transformation parameters was used first to transfer the image coordinates to the map coordinate system and then to resampled the image to register to the map. The registered image is shown in Figure 4.12. The area seen in the Figure is $\approx 4.4 km^2$ in extent and has a terrain variation of 15 meters (checked from the OS map). The map coordinates of the 376 planimetric match points were plotted on the registered image to show how accurately these points lie on the boundaries of the selected image polygons (on visual inspection), which were chosen for the registration.

  Figure 4.12 shows 376 planimetric points plotted as red dots (cross shaped) on the registered image. These points belong to the three features mentioned

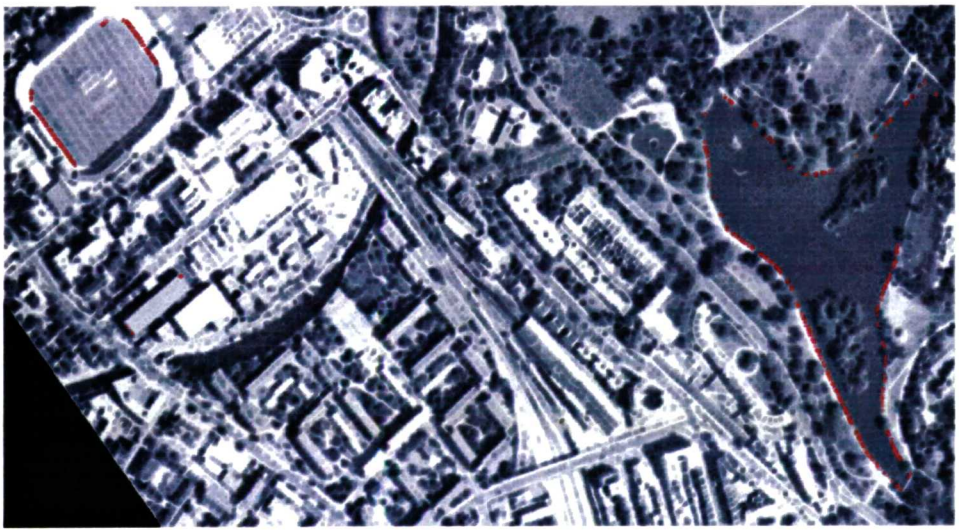


Figure 4.12: Plot of 376 planimetric match points on the registered London Regent's Park image.

above. It can be seen in this Figure that only a few red dots are shown on the building, and the cricket ground and the lake are plotted with many planimetric points. The statistical model algorithm has removed from the cricket

ground and the lake the match points which are associated with the occlusions and shadows. The building has a height component and the algorithm has removed many match points which are of large errors due to perspective distortion. Removal of large errors has registered the image to the map with a sub-pixel accurarcy.

Elchingen-Grosskuchen in Baden-Wuerttemberg (Germany)

Five features were extracted for matching from the map and the image of Elchingen-Grosskuchen. Matching these polygons produced 269 match points. The statistical model algorithm was applied to those match points which after 12 iterative steps selected only 137 match points. The statistics for each iterative step is shown in Table 4.3. The last iterative step shows the difference value between its RMS

Iteration No.	X MIN	X MAX	X RMS	Y MIN	Y MAX	Y RMS	T RMS	RMS_{i-1} - T RMS_i	Match points
1	-2.840	2.370	1.058	-2.950	2.310	0.918	1.40075	1.40075	269
2	-2.160	2.110	0.881	-1.740	1.890	0.662	1.10200	0.29875	234
3	-1.780	1.420	0.747	-1.240	1.310	0.531	0.916499	0.185501	207
4	-1.310	1.240	0.661	-0.940	0.980	0.381	0.762943	0.153556	178
5	-1.240	1.270	0.645	-0.760	0.800	0.255	0.693578	0.069365	159
6	-1.150	1.220	0.633	-0.510	0.360	0.171	0.65569	0.037888	148
7	-1.160	1.210	0.632	-0.310	0.340	0.165	0.653184	0.002506	146
8	-1.170	1.200	0.630	-0.310	0.340	0.161	0.650247	0.002937	144
9	-1.180	1.200	0.629	-0.290	0.350	0.150	0.646638	0.003609	139
10	-1.180	1.200	0.631	-0.280	0.300	0.148	0.648124	0.001486	138
11	-1.170	1.200	0.630	-0.280	0.270	0.146	0.646696	0.001428	137
12	-1.170	1.200	0.630	-0.280	0.270	0.146	0.646696	0.000000	137

Table 4.3: Iterative statistical removal of erroneous match points of map-image for the Elchingen-Grosskuchen area (units in pixels).

error and the preceding one as zero and there is no change in the number of match points. Further application of the algorithm on the last iterative steps shows no removal of match points. The match points of the final iterative step were used for obtaining the transformation parameters to register the image.

The graphs in Figure 4.13 show the content of table 4.3 more illustratively. With each iteration, the range of X residual values reduces progressively from the first iteration to the last one [Figure 4.13 (top-left)]. Each iteration removes match points with large errors which are greater than their $\pm 2\sigma$, and the match points within $\pm 2\sigma$ are used to calculate the X RMS error. The X RMS of all iterations is

shown in Figure 4.13 (top-right). This Figure shows that the accuracy of X RMS increases with each iterative step. The range of Y residuals and Y RMS also shows a similar effect, as shown in Figure 4.13 (middle-left) and figure 4.13 (middle-right) respectively.

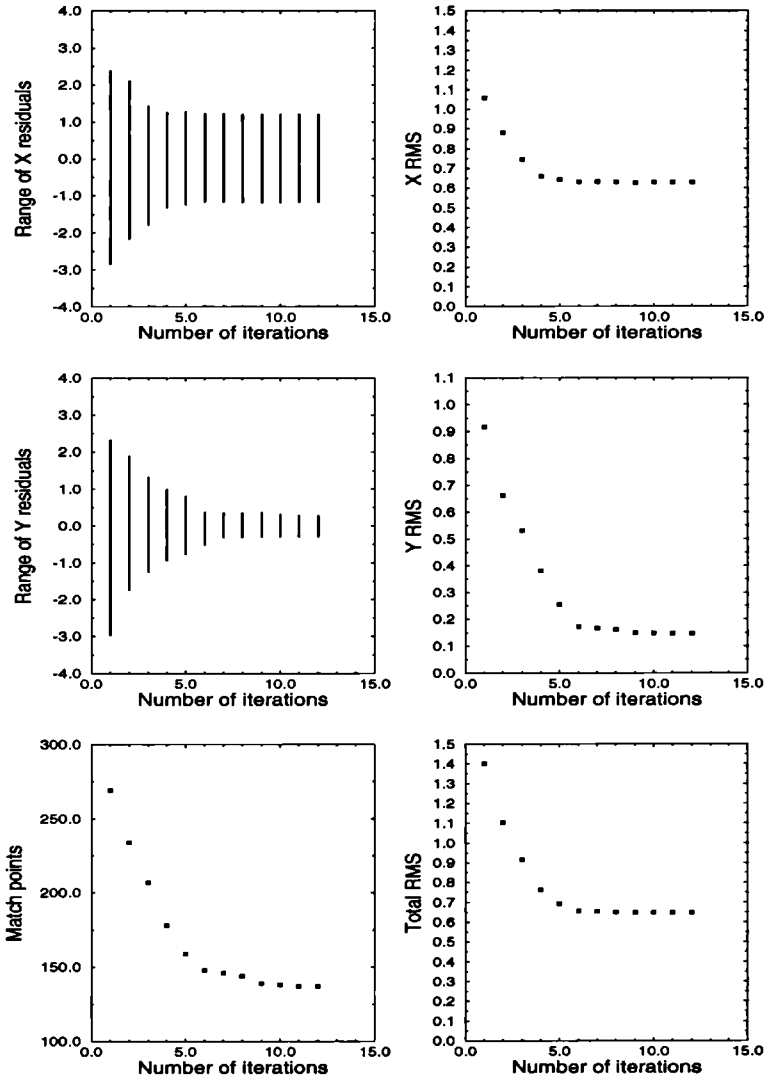


Figure 4.13: A graph representing statistical errors of matching map and image for the Elchingen-Grosskuchen area.

The relation between the iterative process and the match points is shown in Figure 4.13 (bottom-left), wherein it is shown that, with the increase in iterative steps, the number of match points decreases. Thus, it filters the match points with large errors and increases the RMS accuracy of each iteration [Figure 4.13

(bottom-right)].

- **Statistical analysis of iterative processes:** The residual vector of each match point plays a significant role in the measurement of the accuracy assessment of image registration. Table 4.4 shows the statistics of the residual vectors for each iteration. Therein it can be seen that with the increase

Iteration No.	Match points	Residual vector minimum	Residual vector mean	Residual vector maximum	T RMS of residual vectors	Standard deviation
1	269	0.080	1.166	4.095	1.400	0.777
2	234	0.051	0.938	2.591	1.102	0.579
3	207	0.000	0.786	2.064	0.916	0.472
4	178	0.045	0.658	1.536	0.763	0.388
5	159	0.020	0.589	1.444	0.693	0.366
6	148	0.064	0.550	1.232	0.656	0.358
7	146	0.063	0.546	1.222	0.653	0.359
8	144	0.063	0.542	1.213	0.650	0.360
9	139	0.072	0.538	1.222	0.646	0.360
10	138	0.072	0.538	1.222	0.648	0.362
11	137	0.072	0.536	1.222	0.647	0.364
12	137	0.072	0.536	1.222	0.647	0.364

Table 4.4: Statistics of residual vectors of map-image match points of the Elchingen-Grosskuchen area (units in pixels).

of each iterative step, the mean value of residual vectors decreases. This indicates that the accuracy of image registration increases with each iteration. The content of this Table can be emphasised by plotting the residual vectors for each iteration. Three plots are shown in figure 4.14 which shows the residual vector plot for all match points of iteration: 1st, 4th and 12th. The residual vectors on these plots are magnified 50 times to show their effect. The first plot [figure 4.14 (top)] with the residual vectors of 269 match points displays large errors with a RMS error of 1.4 pixels, whereas the two other plots show less error in magnitude. The second plot [figure 4.14 (middle)] shows residual vectors of 178 match points with 0.76 pixel RMS error and the third plot [figure 4.14 (bottom)] shows residual vectors of 137 match points with 0.65 pixel RMS error. The maximum residual vectors of these three plots show values of 4.095, 1.536 and 1.222 pixels respectively.

The statistical analysis of iterative processes has shown that it improves the accuracy of image registration with each iterative step.

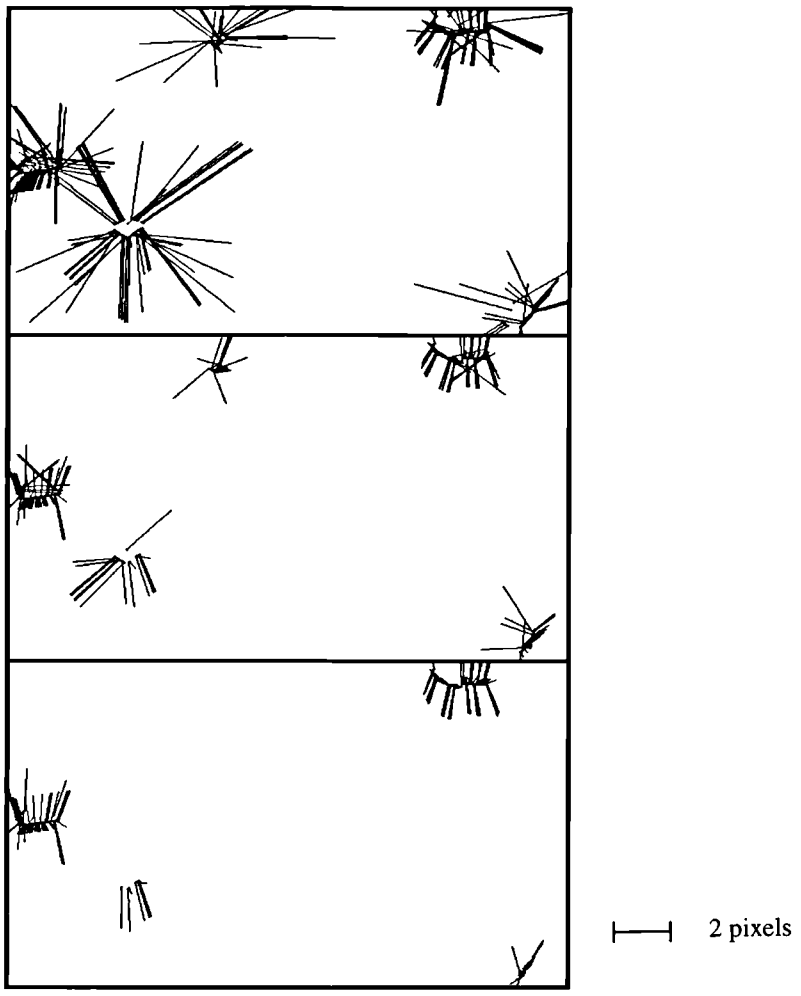


Figure 4.14: A plot of residual vectors magnified 50 times: 1st iteration (top), 4th iteration (middle) and 12th iteration (bottom) of the statistical model algorithm.

- **Validation of registration by a plot of planimetric points:** The transformation parameters were determined by using the match points (137) of the last iterative process of the statistical model algorithm. The image was registered to the map using these parameters. The map coordinates of the 137 match points (planimetric points) were plotted as red dots on the registered image to visually check the accuracy of the registration.

  Figure 4.15 shows 137 points as red dots on the registered image. The area seen in this figure is $\approx 26.5 km^2$ in extent and has a terrain variation of 40 meters (checked from ATKIS DEM).

  The features in this test were selected on the basis of their polygonal shapes

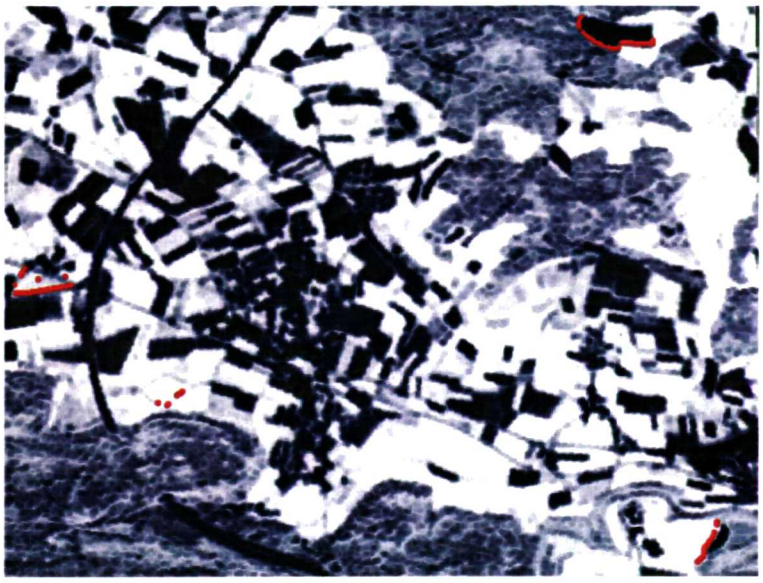


Figure 4.15: Plot of 137 planimetric match points on the registered Elchingen-Grosskuchen image.

(from a 10m resampled raster map created from an ATKIS 1:5,000 vector map), without knowing their feature type. The statistical model algorithm has removed the match points of one selected feature completely, as seen in Figure 4.15. This feature object is smallest in size and has a lower elevation (checked from ATKIS DEM) than the other selected objects.

By removing of the large errors, the image is registered to the map with a sub-pixel accuracy of 0.65 pixel. This test suggests that it is better not to select features on images which appear to be very small in size. There might be the possibility of the removal of the match points of such features by the algorithm if there were considerable height differences with the other selected objects.

London Lord's Cricket Ground (UK)

Seven features were extracted from the map and the image for matching around the Lord's Cricket Ground area. The features are a cricket ground, a boating pond and five buildings. Matching these polygons produced 1039 match points. In this test, more elevated objects were selected to see the performance of the algorithm

planimetric on the perspective distorted images (buildings). The algorithm was applied on 1039 match points and after 18 iterative steps it selected 552 match points. The statistics for each iterative step are shown in Table 4.5. Here, also,

Iteration No.	X MIN	X MAX	X RMS	Y MIN	Y MAX	Y RMS	T RMS	RMS_{i-1} - T RMS_i	Match points
1	-2.560	2.570	1.010	-2.820	2.510	1.080	1.47868	1.47868	1039
2	-2.100	1.970	0.803	-2.250	2.130	0.938	1.23477	0.24391	920
3	-1.630	1.530	0.728	-1.820	1.910	0.856	1.12371	0.11106	856
4	-1.500	1.450	0.683	-1.690	1.740	0.826	1.0718	0.05191	810
5	-1.370	1.380	0.637	-1.650	1.620	0.822	1.03993	0.03187	776
6	-1.320	1.270	0.591	-1.650	1.620	0.820	1.01078	0.02915	744
7	-1.230	1.190	0.547	-1.660	1.640	0.817	0.983208	0.027572	713
8	-1.100	1.090	0.504	-1.610	1.660	0.812	0.955699	0.027509	680
9	-0.990	1.020	0.471	-1.610	1.590	0.807	0.934393	0.021306	654
10	-0.940	0.980	0.456	-1.600	1.590	0.801	0.921703	0.01269	642
11	-0.890	0.950	0.438	-1.610	1.590	0.800	0.912055	0.009648	628
12	-0.850	0.910	0.417	-1.590	1.580	0.796	0.898613	0.013442	610
13	-0.820	0.850	0.390	-1.580	1.590	0.793	0.883713	0.014900	588
14	-0.760	0.790	0.372	-1.570	1.540	0.792	0.875013	0.008700	572
15	-0.740	0.740	0.359	-1.580	1.530	0.791	0.868655	0.006358	560
16	-0.720	0.710	0.354	-1.580	1.530	0.791	0.866601	0.002054	555
17	-0.700	0.700	0.351	-1.580	1.520	0.791	0.865380	0.001221	552
18	-0.700	0.700	0.351	-1.580	1.520	0.791	0.865380	0.000000	552

Table 4.5: Iterative statistical removal of erroneous match points for the map-image the Lord's Cricket Ground area (units in pixels).

the last iterative step shows the same result as a zero value between the difference of its RMS error and the preceding one. This indicates that further use of the algorithm on this step is of no use, i.e. there will be no removal of any match points. The match points of the last iterative step were used for obtaining the transformation parameters to register the image with the map.

The content of Table 4.5 is shown in the form of graphs in Figure 4.16 for better representation. The results in these graphs show that the algorithm filters out bad match points at each iterative step and increases accuracy in the X and Y directions. The total RMS of a sub-pixel accuracy is attained.

- **Statistical analysis of iterative processes:** The statistics of the residual vectors of the match points of each iteration is shown in Table 4.6. The behaviour of residual vectors can be shown by plotting them to see if they are either random or systematic. Three iterative steps were chosen to show this effect. The residual vector plots of iterations 1st, 6th and 18th are

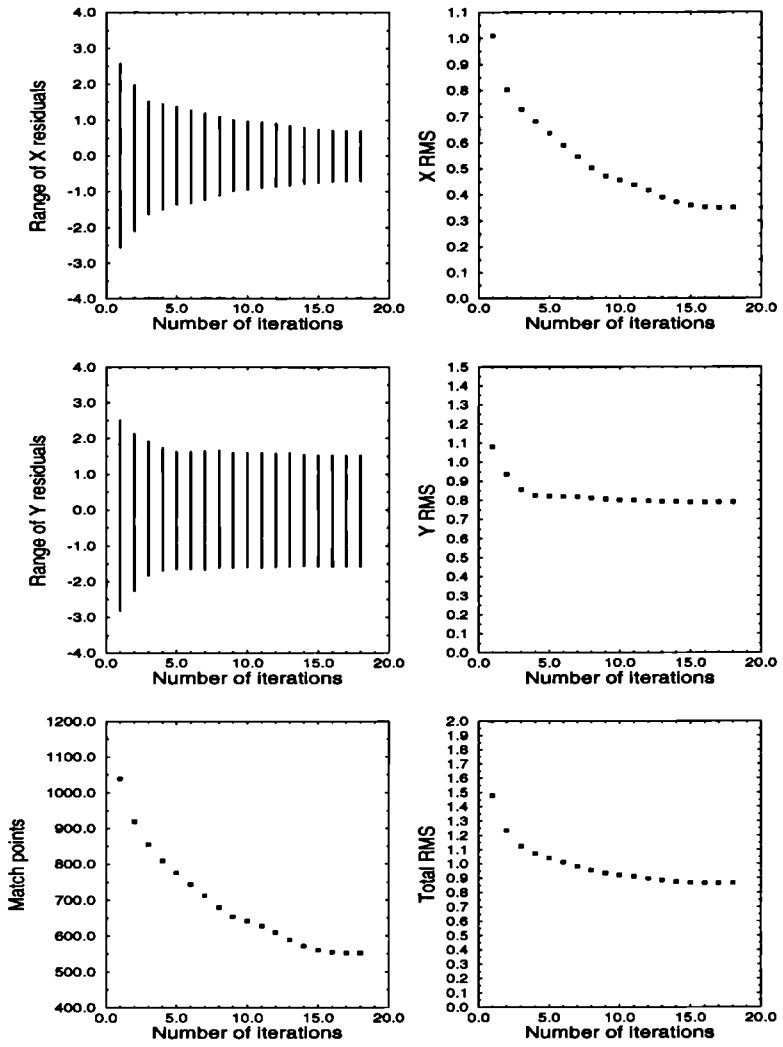


Figure 4.16: A graph representing the statistical errors of a matching map and image of Lord's Cricket Ground area.

shown in Figure 4.17. The residual vectors on these plots are magnified 50 times to show their effect.

The first plot, [Figure 4.17(top)] with the residual vectors of 1039 match points, display large errors in comparision to the two other plots shown in Figure 4.17. The second plot, [Figure 4.17 (middle)] with 744 residual vectors, shows that many match points with large errors which showed up in the first plot have been removed. The last plot, [figure 4.17(bottom)] with residual vectors of 552 match points, shows small errors in magnitude. The largest residual vector error in magnitude of this plot shows an error of 1.613 pixels

Iteration No.	Match points	Residual vector minimum	Residual vector mean	Residual vector maximum	T RMS of residual vectors	Standard deviation
1	1039	0.032	1.279	3.500	1.479	0.742
2	920	0.089	1.097	3.078	1.235	0.566
3	856	0.058	1.012	2.355	1.124	0.489
4	810	0.054	0.968	2.076	1.072	0.461
5	776	0.061	0.939	2.042	1.040	0.447
6	744	0.063	0.913	2.036	1.011	0.435
7	713	0.045	0.888	2.042	0.983	0.423
8	680	0.050	0.862	1.839	0.956	0.413
9	654	0.050	0.842	1.848	0.935	0.407
10	642	0.041	0.830	1.806	0.922	0.401
11	628	0.041	0.820	1.830	0.912	0.400
12	610	0.058	0.807	1.668	0.899	0.397
13	588	0.064	0.791	1.641	0.884	0.394
14	572	0.064	0.782	1.633	0.875	0.392
15	560	0.064	0.775	1.622	0.868	0.391
16	555	0.072	0.773	1.622	0.867	0.391
17	552	0.078	0.772	1.613	0.866	0.391
18	552	0.078	0.772	1.613	0.866	0.391

Table 4.6: Statistics of residual vectors of map-image match points for the Lord's Cricket Ground area (units in pixels).

and the total RMS error is 0.866 pixel whereas the first plot shows the largest residual vector error of 3.50 pixels with the total RMS error of 1.479 pixels, and the second plot shows the largest residual vector of 2.036 pixels with total RMS of 1.011 pixel.

It is noticed here, as well as in the last two tests, that in one particular direction (either X or Y), the RMS shows a considerably higher error than in the other. This might be due to the scanner producing systematic errors in the scanning direction. Nevertheless, the algorithm has shown a sub-pixel RMS accuracy for all three tests by the removal of large errors.

- **Validation of registration by a plot of planimetric points:** The transformation parameters were determined by using the match points (552) of the last iterative process of the statistical model algorithm. The image was registered to the map using these parameters. The map co-ordinates of 552 match points (planimetric points) were plotted as red dots on the registered image to show how accurately these points lie on the boundaries of the selected image polygons chosen for the registration.

  Figure 4.18 shows 552 planimetric points plotted as red dots on the registered

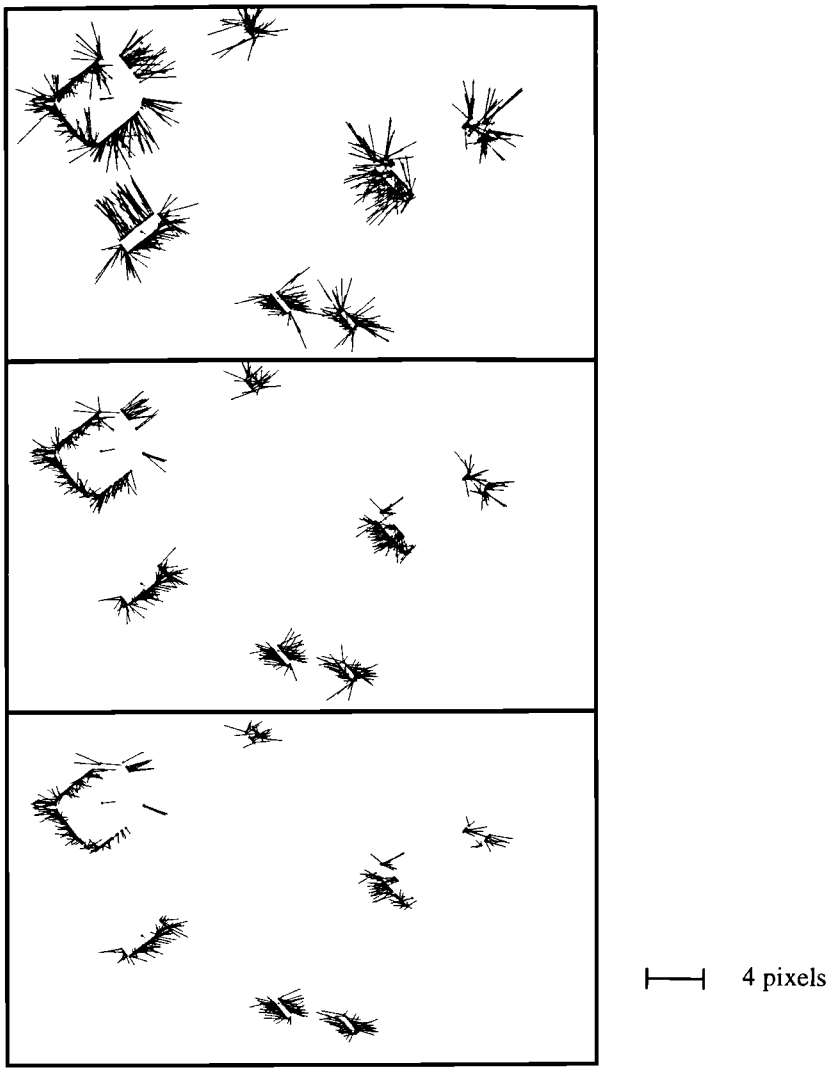


Figure 4.17: A plot of residual vectors magnified 50 times: 1st iteration (top), 6th iteration (middle) and 18th iteration (bottom) of the statistical model algorithm.

image. These points belong to the seven features mentioned above. The area seen in this figure is $\approx 1.4km^2$ in extent and has a terrain variation of 15m. Five buildings, shown marked with red dots, belong to a fairly dense urban-built area. On these buildings, the red dots are missing completely on the edges which are the roof line edges. The match points which belong to the roof line are distorted match points which are products of perspective distortion, and were removed by the algorithm. This was explained before in section 4.3 and section 4.5 in detail how to remove systematic errors. The look angle of a sensor and the orientation of elevated objects create roof lines

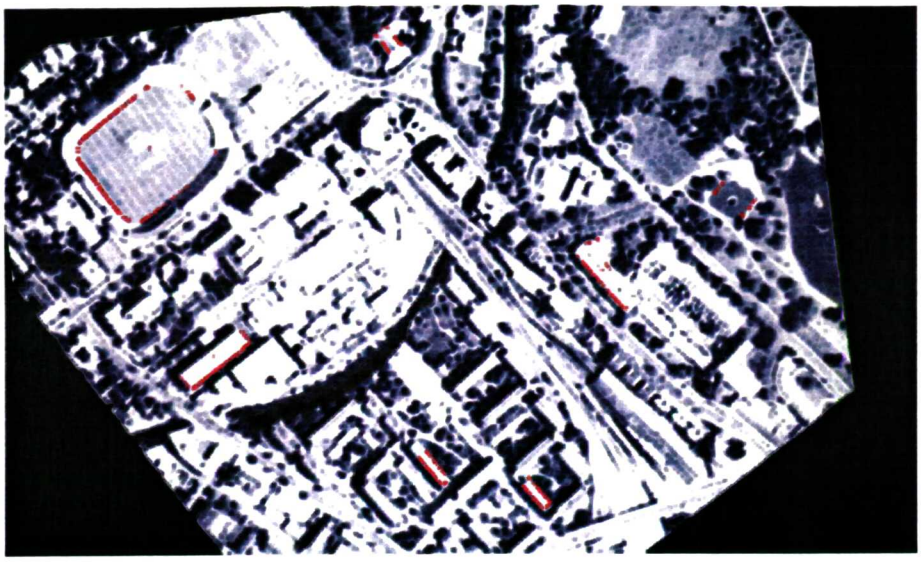


Figure 4.18: Plot of 552 planimatric match points on the registered Lord's Cricket Ground image.

in images whereas non-elevated objects in images do not show perspective distortion and so do not create roof lines. The match points (red dots) on the cricket ground and the boating pond were chosen by the algorithm, are well segmented in the image, and are free from the distortion caused by shadows or occlusions.

Performance of the algorithm: Different areal features of various sizes were chosen in the three tests to check the validity of the algorithm. In the second test, features were selected without their type being known. In the other two tests, ground lying features as well as elevated objects were considered to check the robustness of the algorithm. The statistical model algorithm showed a satisfactory result in the three tests. It is important to note here that the success of the algorithm for obtaining a precise image registration lies in the selection of well distributed areal features of any type, in images of flat or low relief terrain areas.

4.5.5 A comparision between the photogrammetric method and the statistical model for image registration

A digital photogrammetric workstation (DPW) can be used to register images for high relief areas to their map models. Dowman and Vohra (1999), and Bacher (1998) have reported on this using software package SOCET SET of HELAVA on the DPW770, first automatically generating a digital terrain model (DTM). They used two different strategies to compute surfaces, one for the ground surface and the other for the ground surface including ground features (trees, buildings, vegetation etc.). Bacher mentions the problem of automatic computation of DTM in densely built-up areas with strong vegetation coverage and on very steep terrain areas. He also notes that, for forest regions, the top of tree surfaces is recognised as part of the ground surface. Dowman and Vohra report on the high resolution digital aerial photographs of the St. Albans area, producing a stereo model. They recognise that interactive editing is required on the DTM for obtaining the correct foot prints of objects like buildings and vegetation. Using a DTM and a camera model, they produced an orthoimage with SOCET SET and registered it to Landline (database) data for validation, as shown in Figure 4.19.

The SOCET SET software uses a pyramid matching technique to generate numerous match points. Ground control points are required to initialise this technique, which needs the user's interaction. The last chapter described how the image to map registration can be obtained without a manual selection of ground control points. It was explained that patch matching can find the centroid of areal features to initialise the registration process, and the use of dynamic programming can generate numerous match points. If these techniques are used with the St. Alban's image for registration with its map model, and extract large areal features such as the cathedral, lake, stadium and field boundaries, the result should show a precise image registration, as with the three tests carried out and described in the previous and this chapter earlier.

Using the pyramidal matching technique, Dowman and Vohra (1999) produced a DTM of the St. Albans area, extracted buildings and tall trees from the DTM

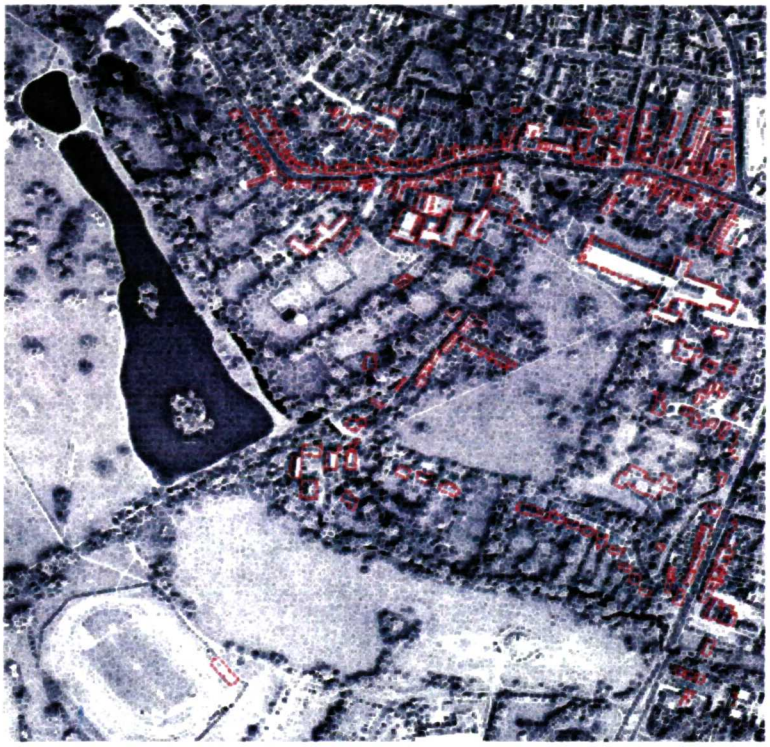


Figure 4.19: Orthoimage of St. Albans area registered to Landline data.

data and, combining it with the orthoimage, generated a perspective view of the area, as shown in figure 4.20.

Compared with the photogrammetric method discussed above, the method developed in this work shows satisfactory results in terms of automation. The method developed here does not require any manual selection of control points for flat or moderate terrain areas. The centroids of areal features extracted automatically are used for initialising the matching to generate numerous match points. Moreover, the statistical model developed in this work, applied to extracted match points, selects planimetric points automatically for registration from high resolution images of elevated and non-elevated objects.

The method developed in this study is economical benefical for flat and nearly flat terrains. The method removes height information in the images for precise map-image registration and it can be used for urban and non-urban areas. For high relief area this method could show results provided small areas in size are selected.

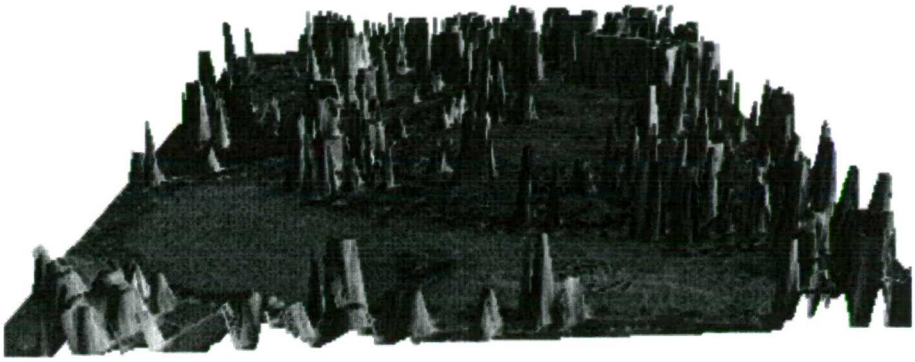


Figure 4.20: Perspective view of elevated objects (tall trees and buildings) in the St. Albans area.

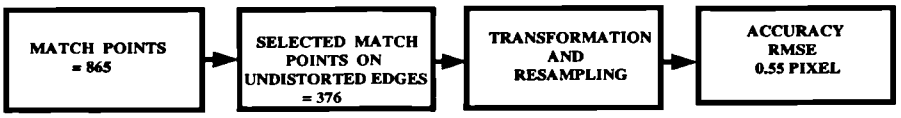
4.6 Conclusions

This chapter investigated a method of extracting planimetric points (control features) in images which coincide with the boundaries of real world objects. Researchers have been working for a long time on the problem of extracting control features in images and this has been comprehensively reviewed by Heipke (1997). Without *a priori* knowledge of the geometry of the 3D objects in images, the extraction of complete features in their true geometry and structure is not possible. However, it is possible to extract a partly correct shape and the position of the features in images (Marr, 1982).

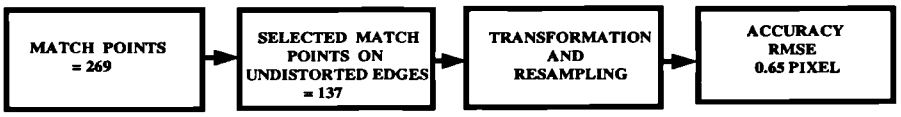
The previous chapter showed the result of matching parts of image features matched with their map model. This chapter extracted some planimetric points from the parts of the matched image features on the basis of a statistical model. An algorithm was developed, based on the model, to remove matched points with large errors due to perspective distortions, occlusions and shadows of elevated and non-elevated objects in images. The validity of the algorithm was tested on three data sets and the summary of results are shown in Figure 4.21. Each test has shown a precision of image registration to a sub-pixel accuracy.

The result has shown a satisfactory outcome. Flexibility of the image registration is achievable with selection of any type of areal feature in images of flat

Result of London Regent's Park:



Result of Elchingen-Grosskuchen:



Result of Lord's Cricket Ground:

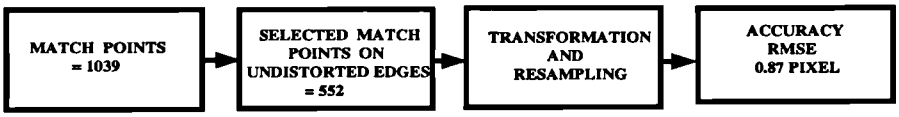


Figure 4.21: Precise registration of the images of the three test areas to their map model.

and low relief terrain areas. The registration system could also show attainability of accuracy for satellite images of high relief areas if the images are divided into subsets, which would minimise the terrain variation effect. In satellite images, the planimetric displacements due to elevation differences of the terrain are much smaller in comparision to aerial images; higher order polynomials would also be benificial for a higher accuracy of image registration.

The statistical model used in this work performs in an iterative-integrative process for removal of errors for precise registration. The statistical approach iteratively removes high error points and selects points which lie close to a plane. It may be said on the basis of the results that the statistical approach is not a trial-and-error technique for attaining precise registration. The iterative-integrated statistical technique has been shown to be a valid technique for obtaining control features in images.

Chapter 5

Automation of Map-Image Registration

5.1 Introduction

Automatic registration of images to maps is beneficial in many disciplines. Photogrammetrists and remote sensing scientists have for a long time been looking for a system which can update maps automatically. This will make their products of immense value to end users. On the other hand, the new availability of high resolution satellite data to civilians, for example the Russian DD5 and forthcoming US Ikonos (1m ground resolution) satellite data, has also increased the demand for the automation of map-image registration in various disciplines. Agricultural scientists, town-planners, map-makers, bureaucrats, decision-makers and many others studying disaster prone areas need a quick remedy solution to their problems. Moreover, the increasing volume of satellite data has increased the demand for data analysis. The number of experts in the field will be less and less able to cope with the enormous amounts of data becoming available. It seems that automation will be the only answer to solve such problems.

Many researchers have been engaged in finding the solution for automatic registration of images to maps, but no solution has yet been developed which can be used in a general and flexible way. The system developed in this study leads to

the automation of registration of images to maps. Automation of the registration process is based on methods of feature extraction, matching and elimination of errors as discussed in the previous chapters. The main operation is concerned with finding a set of common points, i.e. match points, on the map and on the image automatically. The features considered most suitable in this work are polygonal features because they are easily and uniquely identified on maps as well as on images. The arguments for the selection of areal features for map-image registration were discussed in detail in Chapter 2.

The registration system developed in this work is called the Automatic Registration of Map to Image Evaluation Software (ARMIES) system. This system is developed on the basis of the results obtained in the previous chapters. This chapter is concerned with the use of the system from the user's point of view. The main aim is that the operator has minimum interaction, with little or no expertise required, and that the system should be easy to operate. The system contains four modules which are described in the next section.

The issues discussed in this chapter are the strategy for the automation of map-image registration, and the result of the ARMIES system on a new data set. Finally concluding remarks on the ARMIES system and its limitations are made.

5.2 Strategy for the automatic registration of image to map

The main concern of the system developed here is the minimum interaction of the user with the system. Based on this, a strategy is built in such a way that the ARMIES system works as a black box. The system developed consists of four modules. The user's interaction is kept to a minimum and restricted to the first module only. The other three modules work automatically to provide the final result of map-image registration, without any intervention from the user. The work of each module is described below.

- **Preparation Module:** Extraction of selected areal features and their boundaries from a map and corresponding image.
- **Matching Module:** Matching the boundaries of extracted areal features of the map and image to generate a dense network of match points.
- **Error Elimination Module:** Elimination of such match points which are the result of perspective distortion, shadows and occlusions.
- **Registration Module:** Generation of transformation parameters using undistorted match points for a precise registration of image to map.

With these four modules, the system will only require the selection of areal features, which is performed manually. The user has to select a few desirable areal features (a minimum of three) from the displayed map and image on a computer. The computer will ask three questions. The questions are given below:

> Is there any rotational difference in the displayed map and image ?
> If yes, please insert approximate rotational difference in degrees :
> Module 1 will rotate the image and display it on the screen
>
> Is your Area Of Interest (AOI) in the displayed map and image ?
> If yes, please click the top-left and bottom-right hand corners of the
> selected area in the displayed map and image
> The module will extract and display the AOI
>
> For the map-image registration which polygons would you like to
> select ?
> Please select the polygons in the displayed map and image
> Select the map polygons by digitization and select the same polygons
> in the image by clicking on the selected polygons

The answers to these questions will execute the first module of the system. The line commands given in Chapter 3 in Section 3.5.3 are used for map digitization. Here, in this system, the commands, instead of being given by the user, are inserted

in a Arc Micro Language (AML) algorithm, which is part of Module 1. The user has only to digitize the selected map polygons semi-automatically on the Arc/Info system, as was explained in Section 3.5.3. No further interaction of the user is required in this system. It can be shown by a schematic diagram how the four modules of the system are interconnected. Figure 5.1 shows the four module boxes and the processes involved in each module. These processes are not discussed in this chapter. They were described in detail in Chapter 3 and Chapter 4.

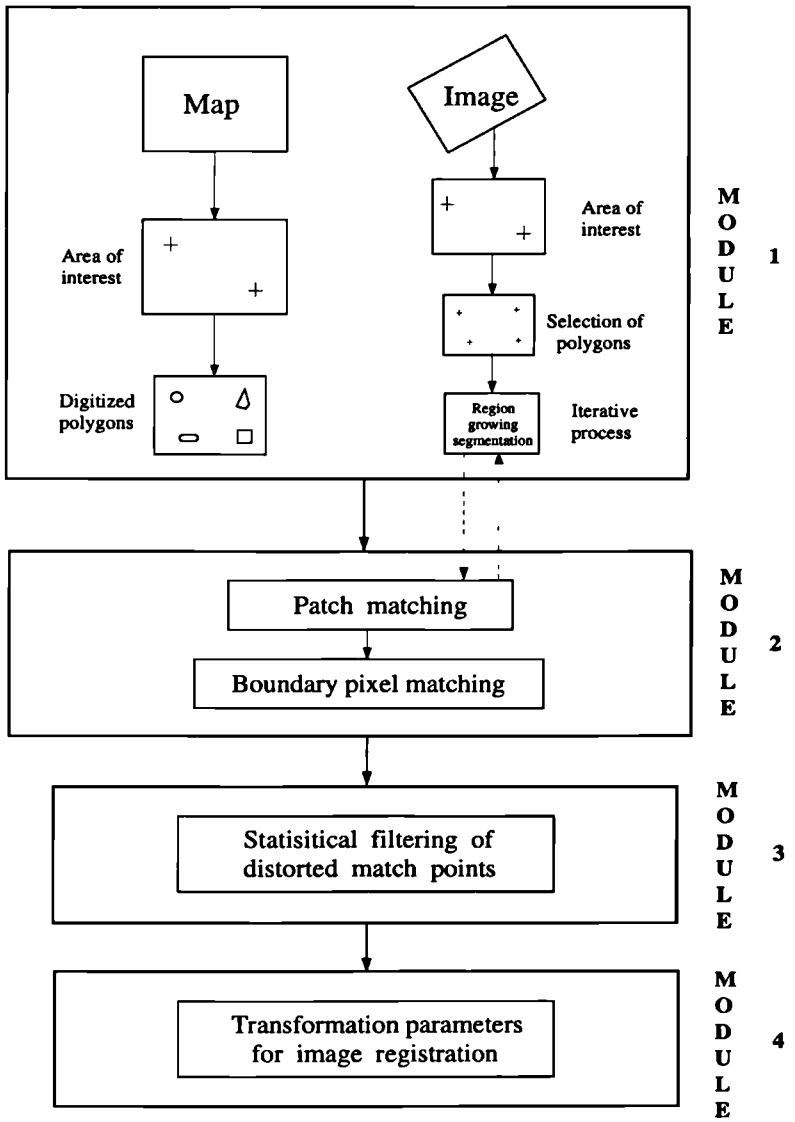


Figure 5.1: A schematic diagram of the Automatic Registration of Map to Image Evaluation Software (ARMIES) system.

5.3 Result of the ARMIES system

This section deals with the utility of the developed system. It comprises two subsections. They are the application of the system and the interaction of the user with the system.

5.3.1 Application of the system

This section describes the testing of the ARMIES system on a new data set for the user's point of view to acquire results which the user would like to see at the end of the processing to check its quality. The area chosen for testing was Farnborough (UK). A section of 1638 x 1556 pixels from a raster map (1:10,000) and a section of 430 x 360 pixels from the high resolution DD5 satellite image ($\approx 2.5m$ ground resolution) were extracted. They belong to the test area and are shown in Figure 5.2 (left) and Figure 5.3 (left) respectively. The selection of the sections from the map and image data were performed in a manner as shown in Module 1 in Figure 5.1.

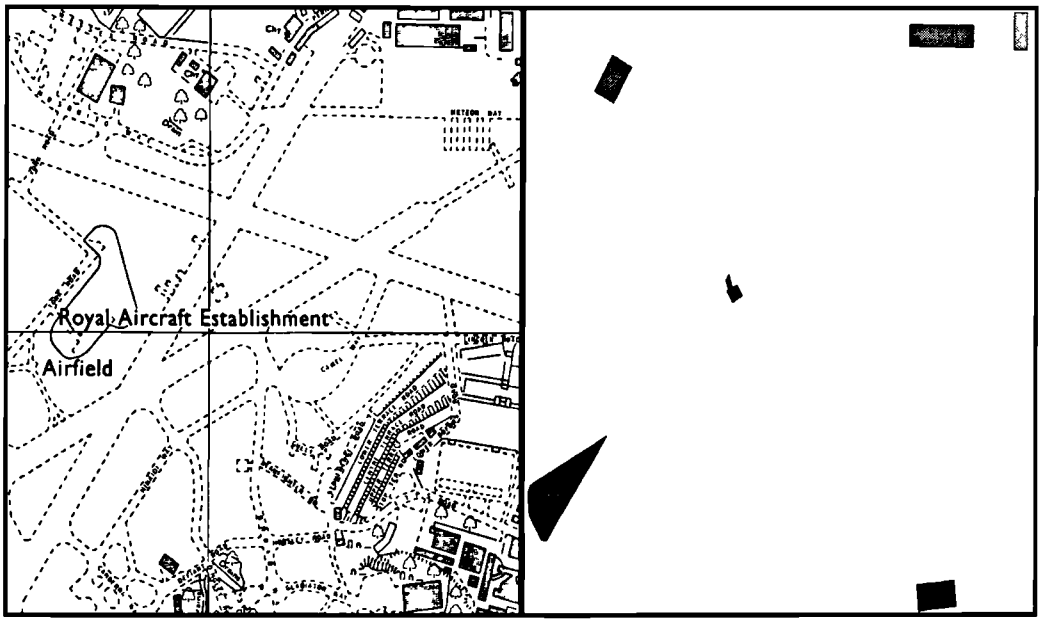


Figure 5.2: A section of the Farnborough map (left) and from the map six selected polygons for matching (right).

Six areal features were selected in the map. The digitization of these features was performed in less than a quarter of an hour using the digitization tool on

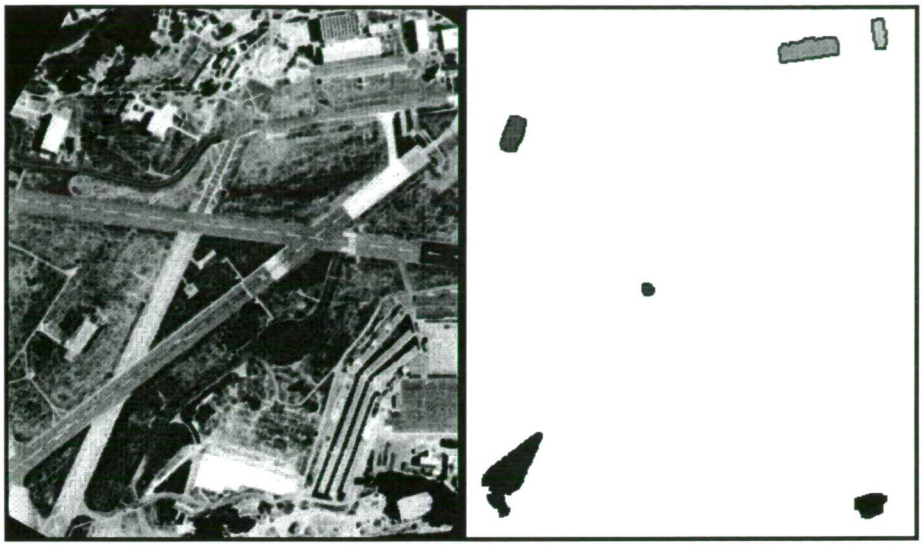


Figure 5.3: A section of a DD5 image of the Farnborough area (left) and the extracted six best shape image polygons to map model (right).

Arc/Info system. The rest of the map preparation for matching, as described in section 3.5.3, was done automatically by Module 1 of the ARMIES system and the result is shown in figure 5.2 (right).

The same areal features were selected in the displayed image. The selection was performed by a mouse click on the selected features. Further processes involved in Module 1 and in the other three modules were performed automatically to produce the final result of the map-image registration. Figure 5.3 (right) shows the output of the extracted best shape image polygons corresponding to the map polygons. These image polygons are the result of the iterative region growing process integrated with patch matching. In Chapter 3, this integrated process was explained in detail. It uses the shape attributes of map and image polygons along with the scale and translation difference of the map and image for patch matching to find the conjugate polygons. The shape attributes of the map and image polygons of this test are given in Appendix A. The scale and translation difference between the map and image are shown in Table 5.1. To identify the conjugate map-image polygons, all attributes of the image and map polygons were multiplied by scale ratio (4.10 : 1) and the translation difference was added to centroids of map polygons in patch matching. The detail of the patch matching technique and its components in matching were

Scale ratio between image and map	Translation difference between map and image co-ordinates	
	in X direction (in pixels)	in Y direction (in pixels)
4.10 : 1	-71.00	126.5

Table 5.1: Scale and translation difference between map and image in the Farnborough test area.

discussed in Chapter 3 with examples.

The identified conjugate map and image polygons, [Figure 5.2 (right) and Figure 5.3 (right)] by module 2, were further treated in the same module to achieve corresponding map-image match points, as shown in Figure 5.4. This Figure shows

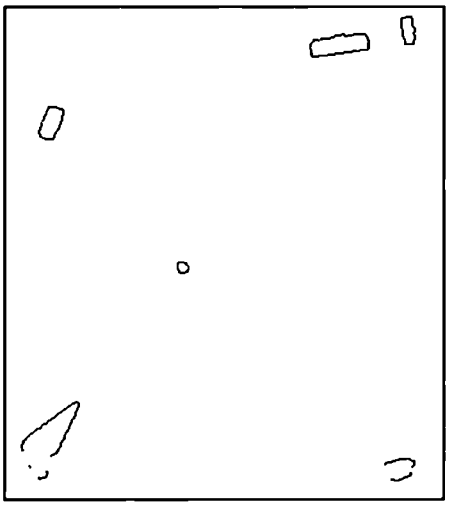


Figure 5.4: Match points (857) of map and image in the Farnborough area.

a large number of match points. This indicates that the integration of the iterative region growing process and patch matching have extracted well segmented image patches which closely resemble the map model. These match points, 857 in numbers, are not all planimetric points. A number of match points out of these are affected by distortion (caused by perspective view, occlusion or shadows). They are bad match points for map-image registration. Module 3 of the ARMIES system was performed on these match points for filtering mismatch or distorted points and to collect the match points which belong to real world objects. The statistical algorithm in this module carried out 22 iterations to remove the bad match points and to collect good match points (354 planimetric points). The principle and processing technique of this algorithm was explained in detail in the last chapter. Table 5.2

below gives statistics for the two iterations, first and last, to show the map-image registration accuracy using all match points (857), including distorted and good match points (354) after filtering the large error match points by the algorithm.

Iteration No.	X MIN	X MAX	X RMS	Y MIN	Y MAX	Y RMS	T RMS	Match points
First	-2.570	2.440	0.979	-2.470	2.680	0.992	1.39374	857
last (22)	-0.800	0.810	0.407	-0.830	0.830	0.420	0.58485	354

Table 5.2: Result of statistical algorithm removing erroneous match points of the map-image of the Farnborough area (units in pixels).

The Table clearly shows the use of the statistical algorithm. The algorithm greatly improves map-image registration ability. Module 4 used these good match points to determine a transformation parameters by applying an affine transformation. The transformation parameters were used to register the image to the map with a sub-pixel accuracy. To visually show the registration accuracy, the registered Farnborough image was plotted with 354 planimetric points (red dots), as shown in Figure 5.5.

Figure 5.5: Planimetric match points (354) plotted on the registered Farnborough image.

At the end of the process, the ARMIES system will give output in the form of statistics and figures, as shown above, for quality assessment of the result.

5.3.2 Interaction of the user

The performance of the system was checked on a data set as shown above. The interaction of the user with the ARMIES system is minimal for the registration of the image to map and without any expertise needed. Moreover, the amount of interaction involved consists only in starting the system, i.e. only to select desirable areal features from the map and image for registration.

The quality of the registration was checked visually on the computer by flickering the map and registered image. It was noticed that each part of the image was registered well except for the bottom right hand corner of the scene. A building was selected for registration from this part of the image, as shown in Figure 5.5. It can be seen in this figure that the building is surrounded by trees. Most of the boundary parts of the building are blocked (occluded). This creates a problem in the extraction of the correct boundaries of the building in the segmentation process [see Figure 5.3 (right)] which has led to mis-registration of a pixel or two in this part of the scene. This information suggests that the user should be careful in selecting features for automatic map-image registration as well as for the achievement of a higher order of accuracy. The user should follow two steps for a precise registration:

- selection of features should be well distributed in the image, and
- Such features whose maximum boundaries are covered by other features should not be selected.

To follow these two instructions, the user will not require any expertise and the system will attain automation and precise map-image registration.

5.4 Conclusions

The automatic map-image registration system developed in this study has shown a satisfactory outcome. The main concern was the no or minimum interaction of the user with the system. This was achieved with the minimum interaction of the

user, as described in this chapter. The system developed has also shown a sub-pixel registration accuracy besides automation.

The developed system has shown few limitations. The system needs at least three areal features for image registration. Moreover, the selected areal features should be well separable, i.e. not completely occluded by surrounding features, and spread over the image scene. Otherwise, the automation may show a poor accuracy of registration.

Chapter 6

Conclusions

This thesis has described the work carried out on automated map-image registration. The work can be divided into two parts: the extraction of areal feature boundaries and the elimination of false extracted boundaries. False boundaries are caused by object heights, shadows and occlusions. Both parts were discussed in Chapter 3 and Chapter 4. Chapter 3 in particular described the development of the feature extraction system and Chapter 4 established the development of a statistical model system for the removal of the defective boundary parts of the extracted features. In Chapter 5, both systems were combined to show automation of map-image registration. The system described in this thesis achieved new functionality that can be enumerated as follows:

- automatic extraction of areal control features obtained by integration of iterative region growing with patch matching,
- new advancement in patch matching to match correct patches of map and image of different scales,
- correct identification of objects of the same size, shape and orientation in two images by patch matching, and
- precise map-image registration obtained by filtering erroneous match points using statistical model.

6.1 Ideas and contributions

Automation of registration is not possible unless correct image and map features are identified automatically. To solve this, maps and images used in this work were prepared to extract object boundaries from the maps and images. The preparation was needed to match their boundaries. Preparation of the map was done by selecting a few areal features (polygons) in the Arc/Info system and digitizing them. The same features were selected in the displayed image. The selection was performed by a mouse click on the selected features. The rest of processes to prepare the image for matching with its corresponding map was done automatically by the system described in this thesis, and resulted in the form of the boundaries of selected areal features. The system includes processes of smoothing, iterative region growing to extract selected areal features in each segmented image, and patch matching. The extracted image features present different sizes and shapes in different segmented images. Such differences are due to variations in the thresholds that were used in region growing. One of these extracted image features has a closest resemblence to its corresponding map areal feature. The closest resemblence was determined, in this research, by finding a minimum cost function with the help of patch matching. The integration of patch matching with region growing process allowed the automatic extraction of correct areal image features (areal control features) to their corresponding map ones.

In this study, a polygon attribute was looked for which alone can solve the problem of identifying the conjugate polygons in the map and image. The answer to this was found by understanding the sensitivity of each polygon attribute in patch matching. The sensitivity of the attributes are dependant on three factors: rotation, scale, and the translation difference between map and image. Table 6.1 gives the information about the attributes which are variant or invariant to these factors.

The centroid of the polygon is the only attribute which is dependant on the three factors as shown in the Table 6.1. This attribute was exploited in this

Factors	Attributes of polygon variant to factors in patch matching					
	Code	Differences	Area	Perimeter	Bounding Rectangle	Centroids
Orientation	yes	no	no	no	yes	yes
Scale	yes	yes	yes	yes	yes	yes
Translation	no	no	no	no	no	yes

Table 6.1: Attributes and factors controlling patch matching.

work and was used for identifying the position of conjugate polygons in the map and the image. Besides this, the centroid was also used for two other purposes:

1. for determining the scale ratio between map and image, and
2. as control points for the coarse registration between map and image.

For the fine registration between map and image an iterative statistical approach developed in this research was used for the precise map-image registration.

There are some limitations to the developed auto-map-image registration system. A small amount of user interaction is required in initialising the system. To account for the amount of user interaction, the system can be divided into four blocks, as shown in Figure 6.1.

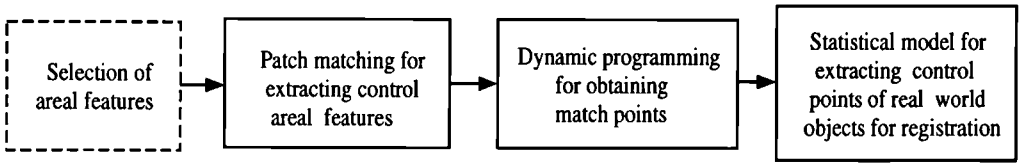


Figure 6.1: User interaction with the registration system.

The first block in the Figure is the block where user interaction is involved with the system and the remaining three blocks are automated to provide registration with high accuracy. This Figure clearly indicates that the developed system is semi-automatic. It will take a considerable amount of time to reach the full-automation stage. The results in this work have shown a satisfactory outcome with a small amount of interaction from the user. It is probable that an automation system without user interaction would lead to

poor results. The Archangel Project (Dowman, 1998) has also worked along similar lines and has shown that user interaction with the system gives better results.

The major contribution of the work in this thesis is the development of techniques which can be used for automated urban and non-urban mapping.

6.2 Future work

The last section of this thesis makes some comments and suggestions for future work on the current topic.

There are two main aspects to be considered for future work:

- testing the efficiency of the system for different types of images, and
- testing for full automation of the system.

Efficiency is related to the accuracy of the system, and full automation to the absence of the interaction of the user with the system. For no indulgence of the user with the system, the system should have an internal in-built part of the system to select areal features. The quality of the registration accuracy should be checked for higher resolution images than the images considered in this thesis. This thesis tested the efficiency of the system up to 2m ground resolution images. The system should be tested further on 0.2m images or higher for checking its robustness and recognising its weaknesses.

There is further scope for developing an automated change detection system attached to the present system for map revision.

The author feels that research explores the path of truth. To follow the road of truth is hard and sometimes painful; but it provides insights of purity and perfection. Research with truth advances the state-of-the-art for the betterment of human-kind.

Appendix A

Shape attributes of polygons

Polygon Number	Code								Differences							
	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
1	39	1	101	0	39	2	99	1	270	4	4	0	0	0	0	4
2	190	2	57	3	190	2	57	3	481	11	3	0	0	0	0	9
3	25	55	44	31	27	55	42	33	38	138	2	0	0	0	0	134
4	10	19	36	27	11	18	37	27	43	65	4	2	0	0	0	71
5	100	141	94	47	14	181	100	1	191	243	3	0	0	0	0	241
6	101	13	66	6	104	12	65	8	294	40	3	0	0	0	0	38

Table A.1: Chain code and first derivative frequencies of map polygons of Farnborough area (units in pixels).

Polygon Number	Perimeter (pixels)	$\sqrt{Area}$ (pixels)	Bounding Rectangle (pixels)		Centre of Gravity (pixels)	
			width	height	X	Y
1	283	62.50	41	102	1506.66	57.86
2	505	106.00	195	62	1269.95	69.78
3	313	85.95	113	130	264.39	186.77
4	186	34.65	51	76	631.22	752.04
5	679	156.59	242	282	117.24	1288.79
6	376	90.89	122	85	1258.36	1586.64

Table A.2: Attributes of map polygons of Farnborough area.

Polygon Number	Code								Differences							
	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
1	7	1	20	3	7	2	18	4	40	9	4	0	0	0	0	9
2	33	10	11	2	39	7	11	5	73	22	3	0	0	0	0	20
3	9	8	17	3	9	8	17	3	29	22	3	0	0	0	0	20
4	5	1	6	2	5	2	4	3	12	9	2	0	0	0	0	5
5	18	34	23	11	19	27	36	5	65	48	7	0	0	0	1	52
6	18	4	11	4	12	8	9	2	36	14	4	0	0	0	0	14

Table A.3: Chain code and first derivative frequencies of image polygons of Farnborough area (units in pixels).

Polygon Number	Perimeter (pixels)	$\sqrt{Area}$ (pixels)	Bounding Rectangle (pixels)		Centre of Gravity (pixels)	
			width	height	X	Y
1	63	13.45	11	24	325.25	19.73
2	119	25.16	48	20	271.65	32.39
3	75	18.55	20	28	36.00	99.68
4	29	7.28	9	9	140.10	219.65
5	174	33.48	47	67	34.71	377.26
6	69	17.17	24	19	318.90	397.72

Table A.4: Attributes of image polygons of Farnborough area.

References

[**Abbasi-Dezfouli et al., 1994**] M. Abbasi-Dazfouli and T. Freeman, 1994. Patch matching in stereo-images based on shape. *International Archives of Photogrammetry and Remote Sensing*, 30(3): 1-8.

[**Ackermann, 1984**] F. Ackermann, 1984. Digital image correlation: performance and potential application in photogrammetry. *Photogrammetric Record*, 64(11): 429-439.

[**Ackermann, 1996**] F. Ackermann, 1996. Some considerations about automatic digital aerial triangulation. In: O. Kolbl (Ed.), *OEEPE - Workshop on Application of Digital Photogrammetric Workstations*, OEEPE Offical Publications No. 33, pp. 157-164.

[**Attneave, 1954**] F. Attneave, 1954. Some information aspects of visual perception. *Psychological Review*, Vol. 61(3).

[**Bacher, 1998**] U. Bacher, 1998. Experimental studies into automated DTM generation on the DPW770. *International Archives of Photogrammetry and Remote Sensing*, 32(4): 35-41.

[**Ballard and Brown, 1982**] D. H. Ballard and C. M. Brown, 1982. *Computer Vision*. Prentice Hall, London, 523 pages.

[**Barnard and Thompson, 1980**] S. T. Barnard and W. B. Thompson, 1980.

Disparity analysis of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI 2(4): 333-340.

[Bellman and Dreyfus, 1962] R. Bellman and S. Dreyfus, 1962. *Applied dynamic programming.* Princeton University Press, Princeton, New Jersey.

[Billingsley, 1983] F. C. Billingsley, 1983. Data processing and reprocessing. *The Manual of Remote Sensing: Chapter 17*, R. N. Colwell, Ed Falls Church, Va: American Society of Photogrammetry and Remote Sensing, Vol. 1: 1039-1040.

[Braun et al., 1996] J. Braun, L. Tang, R. Debitsch, 1996. PHODIS AT - An automated system for aerial triangulation. *International Archives of Photogrammetry and Remote Sensing*, 31(B2): 32-37.

[Brooks, 1978] M. J. Brooks, 1978. Rationalizing edge detectors. *CGIP*, 8: 277-285.

[Canny, 1986] J. Canny, 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI, 8(6): 679-698.

[Caspary, 1987] W. F. Caspary, 1987. *Concepts of network and deformation analysis.* School of Surveying, The University of New South Wales, Monograph 11, Kensington N. S. W. 183 pages.

[Cooper and Robson, 1996] M. A. R. Cooper and S. Robson, 1996. Theory of close range photogrammetry. In *Close Range Photogrammetry and Machine Vision*, Whittles Publishing, Scotland, edited by K. B. Atkinson, pp. 9-51.

[Day and Muller, 1989] T. Day and J-P. Muller, 1989. Digital elevation model production by stereomatching SPOT image-pairs: a comparison of algorithms. *Image and Vision Computing*, 7(2): 95-101.

[Deriche and Giraudon, 1993] R. Deriche and G. Giraudon, 1993. A computational approach for corner and vertex detection. *International J. Comput.*

Vision, 10(2): 101-124.

[**Dowman, 1998**] I. J. Dowman, 1998. Automated procedures for integration of satellite images and map data for change detection: The Archangel Project. *International Archives of Photogrammetry and Remote Sensing*, 32(4): 162-169.

[**Dowman et al., 1996**] I. J. Dowman, A. Morgado and V. Vohra, 1996. Automatic registration of images with maps using polygonal features. *International Archives of Photogrammetry and Remote Sensing*, 31(3): 139-145.

[**Dowman and Ruskone, 1997**] I. Dowman and R. Ruskone, 1997. Extraction of polygonal features from satellite images for automatic registration: The Archangel Project. *Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, ETH Zurich, Ascona, Switzerland, pp. 343-354.

[**Dowman and Vohra, 1999**] I. Dowman and V. Vohra, 1999. *A study on feature extraction and change detection from high resoution satellite images (part 3): DERA Project.* A final report by University College London. 20 pages.

[**Doyle, 1982**] F. J. Doyle, 1982. Satellite systems for cartography. *International Archives of Photogrammetry and Remote Sensing*, 24(1): 180-185.

[**Fischler and Bolles, 1985**] M. Fischler and R. Bolles, 1985. Perceptual organization and curve partitioning. *Image Partitioning and Perceptual Organization, IEEE*: 210-215.

[**Fonseca and Munjunath, 1996**] L. M. G. Fonseca and B. S. Manjunath, 1996. Registration techniques for multisensor remotely sensed imagery. *Photogrammetric Engineering & Remote Sensing*, 62(9): 1049-1056.

[**Forshaw et al., 1983**] M. R. B. Forshaw, A. Haskell, P. F. Miller, D. J. Stanley and J. R. G. Townshend, 1983. Spatial resolution of remotely sensed imagery: a review paper. *International Journal of Remote Sensing*, 4(3): 497-520.

[**Forstner, 1986**] W. Forstner, 1986. A feature based correspondence algorithm for image matching. *International Archives of Photogrammetry and Remote Sensing*, Rovaniemi, vol. 26-3/3: 150-166.

[**Forstner, 1995**] W. Forstner, 1995. Object and image models for image interpretation. *Second Course in Digital Photogrammetry*, Bonn.

[**Forstner and Gulch, 1987**] W. Forstner and E. Gulch 1987. A fast operator for detection and precise location of distinct points, corners and centres of circular features. *Proceedings of ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data*, Interlaken: 281-305.

[**Freeman, 1961**] H. Freeman, 1961. Boundary encoding and processing. *Picture processing and psychopictorics*, Ed. B. S. Lipkin and A. Rosenfeld, Academic Press, New York: 241-266.

[**Freeman and Abbasi-Dezfouli, 1993**] T. G. Freeman and M. Abbasi-Dezfouli, 1993. Algorithms for determining terrain height from SPOT satellite stereo-images. *Proceedings of Advanced Remote Sensing Conference*, The University of New South Wales, Sydney, 133-145.

[**Fritz, 1996**] L. W. Fritz, 1996. The era of commerical earth observation satellite. *Photogrammetic Engineering & Remote Sensing*, 62(1): 173-184.

[**Gahegan and Flack, 1996**] M. Gahegan and J. Flack, 1996. A model to support the integration of image understanding techniques with GIS. *Photogrammetric Engineering & Remote Sensing*, 62(5): 483-490.

[**Ghosh, 1988**] S. K. Ghosh, 1988. *Analytical photogrammetry.* New York: Pergamon. Second edition, 308 pages.

[**Gonzalez and Wintz, 1987**] R. C. Gonzalez and P. Wintz, 1987. *Digital image processing.* Addison-Wesley publishing company. Second edition, 503 pages.

[**Greenfeld and Schenk, 1989**] J. Greenfeld and T. Schenk, 1989. Experiments with edge-based stereo matching. *Photogrammetric Engineering & Remote*

Sensing, 55(12): 1771-1777.

[Gruen, 1985] A. W. Gruen, 1985. Adaptive least squares correlation: a powerful image matching technique. *SAfr. J. of Photogrammetry, Remote Sensing and Cartography*, 14(3): 175-187.

[Gugan and Dowman, 1988] D. J. Gugan and I. J. Dowman, 1988. Accuracy and completeness of topographic mapping from SPOT imagery. *Photogrammetric Record*, 12(72): 787-796.

[Gulch, 1991] E. Gulch, 1991. Automatic extraction of geometric features from digital imagery. *Digital Photogrammetry Systems*, H. Ebner, D. Fritsch, C. Heipke (Eds.): 311-312.

[Haala et al., 1994] B. Haala, M. Hahn, D. Schmidt, 1994. Quality and performance analysis of automatic relative orientation. *SPIE volume 1994*: 140-150.

[Haala and Vosselman, 1992] N. Haala and G. Vosselman, 1992. Recognition of road and river patterns by relational matching. *International Archives of Photogrammetry and Remote Sensing*, 29(B3): 969-975.

[Hahn and Forstner, 1988] M. Hahn and W. Forstner, 1988. The applicability of a feature based and a least squares matching algorithm for DEM-Aquisition. *International Archives of Photogrammetry and Remote Sensing*, 27(3): 137-150.

[Hanaizumi et al., 1990] H. Hanaizumi, T. Hibino and S. Fujimura, 1990. An automatic method for terrain height estimation from SPOT stereo pair images using correlation between epipolar lines. *Proceedings IGARSS '90*: 1959-1962.

[Hannah, 1980] M. J. Hannah, 1980. Bootstrap stereo. *Proceedings Image Understanding Workshop*, pp. 201-208.

[**Haralick and Shapiro, 1985**] R. M. Haralick and L. G. Shapiro, 1985. Image segmentation techniques. *Computer Vision, Graphics and Image Processing*, 29:100-132.

[**Haralick and Watson, 1981**] R. M. Haralick and L. Watson, 1981. A facet model for image data. *CGIP*, 15: 113-129.

[**Harshbarger, 1971**] T. R. Harshbarger, 1971. *Introductory statistics: a decision map.* Macmillan, New York, 558 pages.

[**Hartley, 1988**] W.S. Hartley, 1988. Topographic mapping with SPOT 1 data: a practical approach by the Ordnance Survey. *Photogrammetric Record*, 12(72): 833-846.

[**Hartley, 1991**] W.S. Hartley, 1991. Topographic mapping and satellite remote sensing: is there an economic link? *International Journal of Remote Sensing*, 12(9): 1799-1810.

[**Heipke, 1997**] C. Heipke, 1997. Automation of interior, relative and absolute orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(1997): 57-73.

[**Hellwich et al., 1994**] O. Hellwich, C. Heipke, L. Tang, H. Ebner, W. Mayr, 1994. Experiences with automatic relative orientation. *International Archives of Photogrammetry and Remote Sensing*, 30(3): 370-378.

[**Hildreth, 1983**] E. C. Hildreth, 1983. The detection of intensity changes by computer and biological vision systems. *CVGIP*, 22: 1-27.

[**Hoffmann, 1992**] C. G. Hoffmann, 1992. Satellite image data and their potential use of up-dating and creating large scale topographic maps. *International Archives of Photogrammetry and Remote Sensing*, 29(B4): 496-499.

[**Horn, 1986**] B. K. P. Horn, 1986. *Robot vision.* MIT press, 1986.

[**Kai and Muller, 1991**] Li Kai and J-P. Muller, 1991. Segmentation satellite imagery: a region growing scheme. *Proceedings of IGARSS'91*, vol. II, pp 1075-1078, Helsinki, June 1991.

[**Konecny, 1992**] G. Konecny, 1992. Status of world topographic and cadastral mapping. *Report of the Department of Technical Co-operation for Development of the United Nation*: 31 pages.

[**Kasturi and Jain, 1991**] R. Kasturi and R. C. Jain, 1991. Computer vision: principles. IEEE Computer Society Press, Los Alamitos, California.

[**La Moigne and Tilton, 1995**] J. La Moigne and J. C. Tilton, 1995. Refining image segmentation by integration of edge and region data. *IEEE Transactions on Geoscience and Remote Sensing*, 33(3): 605-615.

[**Lee et al., 1993**] A. J. Lee, N. H. Carender, D. J. Knowlton, D. M. Bell, J. K. Bryan, 1993. Fast autonomous registration of Landsat, SPOT, and digital imagery. *SPIE vol. 1994*: 68-79.

[**Lemmens, 1996**] J. P. M. Lemmens, 1996. A survey on boundary delineation methods. *International Archives of Photogrammetry and Remote Sensing*, 31(3):435-441.

[**Lenzen, 1992**] T. W. Lenzen, 1992. The significance of SPOT imagery to GIS applications: a comparison of databases, their accuracies and uses. *International Archives of Photogrammetry and Remote Sensing*, 29(B4): 683-687.

[**Low, 1991**] A. Low, 1991. *Introductory computer vision and image processing.* McGraw-Hill, London.

[**Lue, 1996**] Y. Lue, 1996. Towards a higher level of automation for SoftPlotter. *International Archives of Photogrammetry and Remote Sensing*, 31(B3): 478-483.

[**Lyvers and Mitchell, 1988**] E. P. Lyvers and O. R. Mitchell, 1988. Precision edge contrast and orientation estimation. *PAMI*, 10(6): 927-937.

[**Madani, 1996**] M. Madani, 1996. Digital aerial triangulation - the operational performance. *International Archives of Photogrammetry and Remote Sensing*, 31(B3): 490-496.

[**Maitre et al., 1989**] H. Maitre and Y. Wu, 1989. A dymanic programming algorithm for elastic registration of distorted pictures based on autoregressive model. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(2).

[**Marr, 1975**] D. Marr, 1975. Analyzing natural images; a computational theory of texture vision. *Technical Report 334, AI Lab, MIT*, June 1975.

[**Marr, 1982**] D. Marr, 1982. *Vision - a computational investigation into the human representation and processing of visual information.* W. H. Freeman and Co., San Francisco.

[**Marr and Hildreth, 1980**] D. Marr and E. Hildreth, 1980. Theory of edge detection. *Proc. R. Soc. Lond.*, B(207):187-217.

[**Marsik, 1996**] Z. Marsik, 1996. Equivalent maps for areal elements of environment ? *International Archives of Photogrammetry and Remote Sensing*, 31(3): 503-510.

[**Mather, 1987**] P. M. Mather, 1987. *Computer processing of remotely sensed images.* John Wiley & Sons, New York.

[**Matsuyama et al., 1980**] T. Matsuyama and M. Nagao, 1980. *A structural analysis of complex aerial photographs.* Plenum Press, New York.

[**Medioni and Nevatia, 1984**] G. Medioni and R. Nevatia, 1984. Matching images using linear features. *IEEE Trans. Pattern Anal. Machine Intell.*, 6: 675-685.

[**Moffitt and Mikhail, 1980**] F. H. Moffitt and E. M. Mikhail, 1980. *Photogrammetry.* Harper & Row, Publishers, New York. Third edition, 648 pages.

[**Moravec, 1977**] H. P. Moravec, 1977. Towards automatic visual obstacle avoidance. *Proc. of 5th Joint Conference on Artifical Intelligence*, Cambridge.

[**Moravec, 1980**] H. P. Moravec, 1980. *Obstacle avoidance and navigation in the real world by a sensing robot rover.* PhD Thesis, Stanford University.

[**Morgado, 1996**] A. Morgado, 1996. *Automated procedures for orientation of digital images.* PhD Thesis, University of London. 252 pages.

[**Morgado and Dowman, 1997**] A. Morgado and I. Dowman, 1997. A procedure for automatic absolute orientation using aerial photographs and a map. *ISPRS Journal of Photogrammetry & Remote Sensing*, 52: 169-182.

[**Morris and Farrow, 1988**] A. C. Morris, A. Stevens, J.-P. Muller, 1988. Ground control determination for registration of satellite imagery using digital map data. *Photogramm. Record*, 12(72), 809-822.

[**Murray et al., 1988**] K. J. Murray and J. E. Farrow, 1988. Experiences producing small scale line mapping from SPOT imagery. *International Archives of Photogrammetry and Remote Sensing*, 27(2): 407-421.

[**Murray and Newby, 1990**] K. J. Murray and P. R. T. Newby, 1990. Mapping from SPOT imagery at the Ordnance Survey. *International Archives of Photogrammetry and Remote Sensing*, 28(4): 430-438.

[**Nalwa, 1993**] V. S. Nalwa, 1993. *A guide tour of computer vision.* Addison Wesley, New York, 1993.

[**Nalwa and Binford, 1986**] V. S. Nalwa and T. O. Binford, 1986. On detecting edges. *PAMI*, 8(6): 699-714.

[**Newton, 1993**] W. Newton, 1993. *An approach to the identification of forest in Thematic Mapper imagery within the context of a change detection system.* MSc in Remote Sensing project, UCL, London: 94 pages.

[**Newton et al., 1994**] W. Newton, C. Gurney, D. Sloggett, I. Dowman, 1994. An approach to the automated identification of forests and forest change

in remotely sensed images. *International Archives of Photogrammetry and Remote Sensing*, 30(3): 607-614.

[**Paderes and Mikhail, 1983**] F. C. Paderes and E. M. Mikhail, 1983. Photogrammetry aspects of satellite imageries. *Proceedings of ACSM-ASP Convention*, Salt Lake City, Utah: 626-635.

[**Paderes et al., 1984**] F. C. Paderes, E. M. Mikhail and W. Forstner, 1984. Rectification of single and multiple frames of satellite scanner imagery using points and edges as control. *NASA Symposium on Mathematical Pattern Recognition and Image Analysis*, 1984.

[**Pal and Pal, 1993**] N. R. Pal and S. K. Pal, 1993. A review of image segmentation techniques. *Pattern Recognition*, 26(9): 1277-1294.

[**Pohl and Genderen, 1998**] C. Pohl and J. L. Van Genderen, 1998. Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, 19(5): 823-854.

[**Rees, 1989**] D. G. Rees, 1989. *Essential statistics.* Chapman & Hall. Second edition, London, 258 pages.

[**Ridley et al., 1997**] H. M. Ridley, P. M. Atkinson, P. Aplin, J. P. Muller and I. Dowman, 1997. Evaluating the potential of the forthcoming commercial U.S. high-resolution satellite sensor imagery at the Ordnance Survey. *Photogrammetic Engineering & Remote Sensing*, 63(8): 997-1005.

[**Rosenfeld and Kak, 1982**] A. Rosenfeld and A. C. Kak, 1982. *Digital image processing*, Vol. 2. Academic Press, 1982.

[**Rowntree, 1981**] D. Rowntree, 1981. *Statistics without tears: a primer for non-mathematicians.* Penguin Books Ltd, London, 199 pages.

[**Russ, 1995**] J. C. Russ, 1995. *The image processing handbook.* CRC Press, Inc., Second edition, Boca Raton. 674 pages.

[**Schenk et al., 1991**] T. Schenk, J. C. Li and C. Toth, 1991. Towards an autonomous systems for orienting digital stereopairs. *Photogrammetric Engineering & Remote Sensing*, 57(8): 1057-1064.

[**Schickler, 1992**] W. Schickler, 1992. Feature matching for outer orientation of single images using 3D wireframe control points. *International Archives of Photogrammetry and Remote Sensing*, 29(B3): 591-598.

[**Shahin and Novak, 1994**] F. Shahin and K. Novak, 1994. The exterior orientation of digital images by road matching. *International Archives of Photogrammetry and Remote Sensing*, 30(2): 174-181.

[**Stevens et al., 1988**] A. Stevens, A. C. Morris, T. J. Ibbs, J. P. Muller, 1988. Automatic Generation of Image Ground Control Features From a Digital Map Database. *IAPRS*, 27(2/3): 402-413.

[**Torre and Poggio, 1986**] V. Torre and T. A. Poggio, 1986. On edge detection. *PAMI*, 8(2): 147-163.

[**Toutin, 1995**] T. Toutin, 1995. Multisource data integration: comparison of geometric and radiometric methods. *International Journal of Remote Sensing*, 16(15): 2795-2811.

[**Triola, 1989**] M. F. Triola, 1989. *Elementary statistics.* The Benjamin Cummings Publishing Company, Inc. Fourth edition, California, 784 pages.

[**Tseng et al., 1997**] Y-H. Tseng, J-J. Tzen, K-P. Tang and S-H. Lin, 1997. Image-to-image registration by matching area features using fourier descriptors and neural networks. *Photogrammetric Engineering & Remote Sensing*, 63(8): 975-983.

[**Veillet, 1992**] I. Veillet, 1992. Accuracy of SPOT triangulation with very few or no ground control points. *International Archives of Photogrammetry and Remote Sensing*, 29(B4): 448-456.

[**Vohra and Dowman, 1996**] V. K. Vohra and I. J. Dowman, 1996. Automatic extraction of large buildings from high resolution satellite images for registration with a map. *International Archives of Photogrammetry and Remote Sensing*, 31(3): 903-908.

[**Vohra and Dowman, 1997**] Vohra, V. and Dowman I., 1997. Automatic extraction of significant features from high resolution satellite images for registration with a map. *Developing Space '97: Proceedings of the 1997 Remote Sensing Society Annual Student Meeting*, 23-29.

[**Weidner, 1995**] U. Weidner, 1995. Image processing. *Second Course in Digital Photogrammetry*, Bonn.

[**Wolf, 1983**] P. R. Wolf, 1983. *Elements of photogrammetry.* McGraw-Hill International, Second edition, New York, 628 pages.