

THE UNIVERSITY OF CALGARY

Automating Procedures On A Photogrammetric Softcopy System

by

Yubin Xin

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF GEOMATICS ENGINEERING

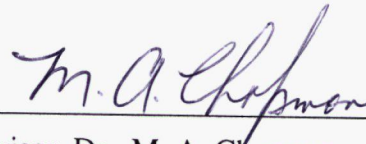
CALGARY, ALBERTA

SEPTEMBER, 1995

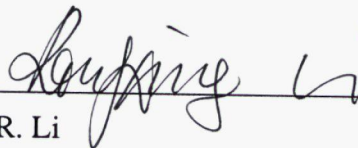
© Yubin Xin 1995

THE UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Automating Procedures on a Photogrammetric Softcopy System" submitted by Yubin Xin in partial fulfillment of the requirements for the degree of Master of Science.



Supervisor, Dr. M. A. Chapman
Department of Geomatics Engineering



Dr. R. Li
Department of Geomatics Engineering



Dr. A. K. Mehrotra
Department of Chemical & Petroleum Engineering

Date: September 26, 1995

ABSTRACT

Geo-referenced digital images taken by CCD cameras and integrated with GPS/INS data provide an efficient way for monitoring special objects with high resolution and visual capability in a highway/road inventory system. This research represents some techniques for extraction of spatial information and generation of GIS database from geo-referenced CCD.

Unlike conventional photogrammetric systems that use optical instruments and perform measurement on a very limited number of high resolution photos, a softcopy photogrammetric system uses computer and large number of digital images. This fast, accurate and convenient system consists of GUI, geometric functionalities and real time results displaying in a graphics environment. It can process images to improve the quality in terms of gray level enhancement, edge sharpening and noise removal, perform the measurement and the triangulation to get 3D object coordinates. A more important part in this softcopy package is the automatic/semi automatic object measurement and positioning. To this end, cross correlation matching, least squares matching, feature based road central line and road edge matching are studied and implemented. Results indicate that the semi-automatic highway/road feature extraction is an efficient way for map production and GIS database generation.

ACKNOWLEDGMENTS

I wish to express my deep gratitude to my supervisor, Dr. M.A. Chapman, for his consistent guidance, assistance, and encouragement throughout my graduate studies. I am grateful to Dr. R. Li for the valuable suggestion and discussion during my research work. Appreciation should also go to my fellow graduate students, Mr. L. Qian and Mr. H. Li during my thesis research. The research is supported by the University of Calgary and the Geofit Inc., Laval, Quebec.

TABLE OF CONTENTS

APPROVAL PAGE	ii
ABSTRACT	iii
ACKNOWLEDGMENTS	iv
TABLE OF CONTENTS	v
LIST OF FIGURES	viii

CHAPTER	Page
1 INTRODUCTION	1
1.1 Background	3
1.2 CCD cameras and geo-referenced images	5
1.3 VISAT -- A mobile highway survey and data extraction system	7
1.4 Research objectives	9
2 DESIGN CONCEPT OF GEOSTATION	11
2.1 Geo-referenced CCD image data transfer	11
2.2 Multiview subsystem	12
2.3 Overview subsystem	15
2.4 System configuration and GUI development	17
3 MAIN FUNCTIONS OF GEOSTATION	20
3.1 Photogrammetric principles	20
3.2 Epipolar geometry	23
3.3 The space intersection and bundle adjustment	27
3.4 Database generation and GIS link	29

3.4.1	Database generation.....	29
3.4.2	GIS link.....	33
4	IMAGE PROCESSING	38
4.1	Grey level enhancement.....	38
4.1.1	Histogram equalization	38
4.1.2	Linear stretch.....	40
4.2	Image sharpening	44
5	SINGLE POINT IMAGE MATCHING.....	50
5.1	Area-based single point matching	50
5.1.1	Correlation matching	51
5.1.2	Least squares matching.....	55
5.2	The reduction of searching range.....	58
5.3	The optimal matching window size and shape.....	61
5.4	The application of triangulation from multi images.....	64
5.5	Feature supported area-based single point matching.....	67
6	ROAD FEATURE EDGE EXTRACTION IN 3D	71
6.1	Edge detection and noise removal.....	72
6.2	Edge binarization and skeletonization	77
6.3	Edge following.....	81
6.3.1	Direction matrix line following.....	81
6.3.2	Least squares line following	85
6.3.3	Polynomial edge fitting	88
6.4	Edge vectorization	92
6.5	Automatic extraction of 3D road feature edges.....	94
6.5.1	Matching the starting point	94

6.5.2	Matching the vertices.....	97
6.5.3	Dynamic link of matching edge from sequential image pairs	99
7	CONCLUSIONS AND RECOMMENDATIONS.....	102
7.1	Conclusions.....	102
7.2	Recommendation.....	105
	REFERENCES.....	108

LIST OF FIGURES

No.		Page
2.1	Hierarchical structure of the whole system	11
2.2	3D determination through space intersection.....	13
2.3	Graphic User Interface (GUI) of Multiview.....	15
2.4	Graphic User Interface (GUI) of Overview	16
2.5	Geostation system architecture.....	19
3.1	Spatial relationships in the system	21
3.2a	Transformation from digital image coord to photo coord	24
3.2b	Transformation from photo coord to object coord.....	24
3.2c	Transformation from space coord to baseline coord	26
3.3	Main structure of the GIS system.....	34
3.4	Display of measurement results in Arc/Info	36
4.1	Original and equalized images	40
4.2	Two-segment linear stretch and processed result.....	42
4.3	Three-segment linear stretch and processed result	43
4.4	Change of histogram before and after processing.....	44
4.5a	3x3 neighborhood for Sobel filters	45
4.5b	Expanded neighborhood	46
4.6	Result of Sobel “addback”	47
4.7	Grey level difference mask	48
4.8	Processed image using grey level “addback”	49
5.1	Target and searching masks	51
5.2	Weight distribution	54

5.3	Matching result by using 3x15 window	64
5.4	Edges in vector form.....	68
5.5	Mismatch by large geometric disformation	69
5.6	Correct match by combined conditions.....	70
6.1	Edge detection in local areas	73
6.2	Template matrices.....	78
6.3	An example of edge filtering	79
6.4	Description of the detailed skeletonization procedure.....	80
6.5	Ordering of neighborhood pixels	81
6.6	Directional matrices	83
6.7	Large directional matrices	84
6.8	Conversion from raster to vector.....	93
6.9	An example of edge matching	101
6.10	Projection on Overview	101

Chapter 1

Introduction

During the last few years, the interest in the use of digital images has been increasing. This has resulted in the development of digital photogrammetric systems for the processing of those image data. These developments date back to the early 1980s (Sarjakoski, 1981; Case, 1982, Dowman, 1992). Since 1988, this field has been growing rapidly due to the increasing availability of digital imagery and the revolutionary hardware/software progress in computer science.

Digital photogrammetry is the technology used to acquire spatial/non-spatial information about objects in the three-dimensional (3D) space from 2D digital images of these objects. Digital photogrammetry has the potential for automating large parts of the photogrammetric processing chain and thus for the implementing of rapid information acquisition. And, the goal of using inexpensive equipment and less qualified labor can also be reached in the development of digital photogrammetry.

Softcopy photogrammetric systems are based on digital images and their orientation details. A softcopy photogrammetric system consists of hardware and software to carry out photogrammetric tasks in an interactive or automated way using digital image data as the original data. The three dimensional coordinates, geometric

object surface descriptions, structured vector data, and attribute information about objects can be extracted as final results. They are near real time photogrammetric system.

Many factors have influenced the developments of softcopy photogrammetric systems, e.g. the availability of accurate and affordable CCD cameras, low-cost computer hardware which provides the required speed, storage, and display facilities (Sizlet, 1992). “So that the system would solve the standard photogrammetric problems in a user-friendly and efficient way and even perform tasks... beyond the possibilities of conventional photogrammetric instrumentation”(Gagnon et al. , 1990).

While the design issues discussed above are important, the heart of a softcopy photogrammetric system is in the implemented software. Analytical plotters are already equipped with programs related mainly to the positioning of the measuring mark at given locations in images in order to manually perform data acquisition. However, the use of digital imagery has the potential for the automating of substantial parts of the processing chain of digital photogrammetry. Automation in digital photogrammetry can be divided into two forms: completely automatic processing and semi-automatic interactive processing. Here image matching algorithms play the essential role in obtaining solutions.

Although the ideas of completely automatic systems are discussed in many papers, practical, usable systems of this kind are out of reach in the foreseeable future. Semi-automatic processes assist in the measurement of the image coordinates of a point with given coordinates in one image. This procedure can be applied when the operator specifies a measurement target by giving one point in an image, and then the image coordinates of the point in the other image(s) are found automatically. This procedure is

also used in feature extraction. The operator can provide attributes for the relevant feature along with an initial position or shape; the system then finds the precise location of the feature.

We can observe that the automation of parts of the photogrammetric processing chain has made substantial progress in the last few years. However, the algorithms need to be implemented in an interactive environment, such as a softcopy system. The system is equipped with the necessary photogrammetric software together with a comfortable user interface. The human operator provides initial values for the algorithms, verifies and corrects the automatically obtained results, and performs additional measurements manually.

1.1 Background

Unlike conventional photogrammetric systems that use optical instruments and perform measurement on a very limited number of high resolution film-based photos, softcopy photogrammetric systems use computers and a large number of digital images. Although the resolution of digital images acquired from CCD cameras is not as high as that of conventional photos, those digital images employed are much less expensive and can be easily accessed, stored and analyzed. Advantages can be taken of the flexibility of computer systems in terms of data storage, management and computing capability. For the purpose of image processing and analyzing, the current softcopy photogrammetric systems are mostly implemented on workstations. The requirement for this kind of system is not only a measuring and data acquisition system, but also a comprehensive geo-information extraction, analysis and transformation system. Another important factor has

become more recognized recently -- the need for integration of information from different sources. The development of Geographic Information Systems (GIS) and Computer Aided Design (CAD) systems for handling spatial data offers this information to many different users. GIS and CAD progress has largely contributed to the development of softcopy photogrammetry, which acts as the data acquisition component and provides an active interface for the manipulation of the database, and gives access to a large toolbox of photogrammetric, data management, and visualization techniques. For all applications, the accuracy of the product depends on the image resolutions and the quality of geo-reference data.

In this thesis, a prototype for a digital photogrammetric system is described. This is a fast, accurate and convenient system for measurement, which consists of two parts: hardware/software for data acquisition and a workstation for measurement and database generation. Whenever an area needs to be surveyed, a vehicle equipped with instruments travels through that area and takes a series of images of the objects to be surveyed. Upon its return, a mass storage module is dismounted from the computer located in the vehicle and taken to the office workstation. The data are processed and displayed on the workstation screen. Through interactions, the operator can measure all of the objects that appear in these images and, subsequently, store the 3D data into GIS databases. At the same time, a 3D projection system is used to give the operator a realistic sense of the object position as well as the resulting map in a typical vector format.

In this package, automatic/semi-automatic object measurement is possible. Image matching is the most important method used to automate the measuring procedures. In

this measuring procedure, either for the interactive or for the automatic/semi-automatic module, image quality is a critical factor. As a result, various pre-processing approaches to improve image quality will be discussed and compared. Correlation area-based single point matching will be firstly discussed and implemented. Since all the images are geo-referenced, epipolar geometry will be applied for the purpose of improving the efficiency and the reliability of the matching. In addition, least squares single point matching (LSM) will be used to refine the matching results. Because of large geometric distortion in close range photogrammetry, optimal window size and shape for reliability and accuracy will be investigated. Furthermore, feature-based techniques will be employed to improve the reliability of the result of single point matching. A combined matching method will be developed and tested. For the improvement of accuracy, multi image single point matching will be developed and implemented. Finally, a feature-based matching method will be investigated and studied. Automatic/semi-automatic information extraction from the stereo image data has become the most important research topic in our system. The main features to be extracted from highway image pairs are edges and center lines of roads. This research will concentrate on a semi-automatic edge extraction and positioning procedure, which includes edge detection, binarization, skeletonization, edge following, vectorization, and edge-based stereo matching.

1.2 CCD cameras and geo-referenced images

In the field of digital photogrammetry, CCD (Charge-Coupled Device) cameras play a very important role. The mechanical and electronic design of CCD cameras is not usually oriented for photogrammetric purposes. In digital image acquisition systems,

frame grabbers are critical components. When an object is focused by a lens system on the sensor array, each element cell detects the intensity of the light incident on it and transfers the intensity as an electrical charge (Beyer, 1992). Then the analog transmitted imagery is converted into a matrix of numbers---the digital image, by the frame grabber.

For photogrammetric applications, digital images are becoming a replacement of conventional photographs. Compared with film-based cameras, CCD cameras have many advantages, which include:

- no film distortion;
- no emulsion problems;
- geometric stability of the image element array;
- the need for fiducial marks to define the interior geometry of the cameras is eliminated. (Curry et al., 1987)

Furthermore, another major advantage of the CCD video images is the high computer compatibility. Therefore, with fast evolution of computer hardware and software, low cost digital images can be easily stored, transferred, processed and accessed. These factors will greatly contribute to practical photogrammetric systems. However, CCD video images have one major disadvantage, i.e., they generally have a low spatial resolution. It can be improved either by some image processing technology or by the ever improving manufacturing techniques.

Because regular CCD cameras are non-metric cameras, the calibration procedure is a critical step in the whole system. However, CCD images without geometric control

cannot supply three dimensional information about objects which are to be acquired for GIS database generation. In the field of photogrammetry, geo-reference is an expensive, time consuming process. Usually, a large number of ground control points are required to obtain whole geometric control for the photos or images. Nowadays, high precision positioning can be achieved by applying differential Global Positioning System (GPS), with an additional receiver located at a control point with known coordinates. As for the cameras attitude information, an Inertial Navigation System (INS) can be used to get angular orientation parameters. The integration of GPS and INS technologies is especially advantageous in situations where GPS signals cannot be received in good condition. In this way, expensive ground points and most field work can be eliminated, and complete geometric control of the images can be achieved. Therefore, this kind of images can be defined as geo-referenced images.

1.3 VISAT -- A mobil highway survey and data extraction system

In this thesis, a prototype digital photogrammetric system, VISAT, is described and discussed. This is a fast, accurate and convenient surveying system for highway information acquisition (Li, et al, 1994). It consists of two parts: a data acquisition part with a mobile platform which can record surveyed objects within short range and a workstation for measurement and database generation.

In the VISAT system, the geometrical control of object measurements is a hierarchical procedure. The location of the survey van at any time is determined by GPS through a receiver mounted on the top of the van. High precision positioning is achieved by applying differential GPS with an additional receiver located at a control point with

known coordinates. An INS installed inside the van supplies angular orientation parameters of the van in the three dimensional coordinate system as well as interpolated positions during GPS outages. CCD cameras (two to six) installed on the top of the van record images of objects along the road within a corridor of about 50 meters, while the INS/GPS system simultaneously records the time, positions and orientations of the cameras. The computer on board the van records the image data and the navigation data onto a mass storage unit. This unit can consist of a hard disk array or magnetic tapes depending on the selected system configuration. For each exposure station, there is one record of data corresponding to an information header and the image data.

The information header includes the navigation data for the CCD cameras. Corresponding to each camera, there are two parts of data describing position and attitude. A set of 3D coordinates in a global reference system (e.g. 3TM) defines the perspective centres of the cameras and a rotation matrix defines the orientation of each camera. In addition, some camera calibration parameters are included. Consequently, for one frame of data, all of the geo-reference data are available for the corresponding exposure station. These data are the basis of the accurate measurements which can be made on the objects appearing in multiple images.

The Geostation consists of a SUN SPARC IPX with two monitors. One monitor is used for the so-called Multiview subsystem, where stereo images and graphic user interface elements are displayed. Most of the interactive and automatic procedures are conducted on this screen. The other monitor is used for the overview application with the base map and P-lines (i.e. routes of the van) displayed. Whenever an object is measured,

its projection on the base map is shown on the Overview screen. The main purpose for Overview is to give the user a reference for the measurement data using a base map. For high end applications, we can have more than one monitor for Multiview, or even several workstations networked together in case of time critical surveying in which several operators can simultaneously carry out measurements of data from the same data set. For low end applications, Overview and Multiview can share a single screen or Overview may be disabled.

1.4 Research objectives

In this study, a softcopy photogrammetric system, Geostation, will be described, and the system designing method and some functionality models will be studied. The primary objectives are:

- to establish a measurement procedure for extracting 3D information of objects;
- to develop a Graphical User Interface(GUI).
- to generate a database for storing spatial/non-spatial results;
- to transfer the measurement results into GIS package, ArcInfo, for further GIS applications;
- to set up a preprocessing system to enhance the quality of images.
- to develop and implement automatic/semi-automatic measurement methods through image matching;

There are seven chapters in this thesis. The main content of each chapter is listed as follows:

In Chapter 2, a brief description of Geostation designing concept is given. The Multiview and Overview subsystems are discussed together with the GUI development.

Chapter 3 deals with the main functionalities in this system, include basic digital photogrammetric principles and implementation as well as a GIS link.

In Chapter 4, preprocessing of digital images is discussed. Here, several image enhancement techniques are studied or improved. Spatial and spectral filtering theory is also presented. Some of the methods have been implemented.

Chapter 5 focuses on variations of the single point image matching approaches. The ability of each of these methods is discussed and examined. Some matching functions are successfully integrated in our system.

In Chapter 6, feature-based edge matching methods are investigated and presented. Results of different feature matching techniques applied to real stereo images are presented. 3D information of road center line and road edge are extracted from these images.

Conclusions and recommendations for further research are given in Chapter 7.

Chapter 2

Design Concept of Geostation

2.1 Geo-referenced CCD image data transfer

The structure of the whole system is hierarchical (Figure 2.1 Qian, et al, 1994). Through a calibration procedure, the spatial relationships between INS and GPS centres, camera centres relative to the INS/GPS, the camera's inner and relative orientation parameters can be determined (El-Sheimy and Schwarz, 1993). In the survey van, GPS time controls the synchronization of all the system components. The prototype system is equipped with three CCD cameras with a basic format of 512x480 pixels with 256 grey levels. A frame grabber grabs three images every 0.7 second. Acquired digital images are stored on the PC hard disks and, subsequently, downloaded to 8mm tapes or directly transferred to a workstation through networks.

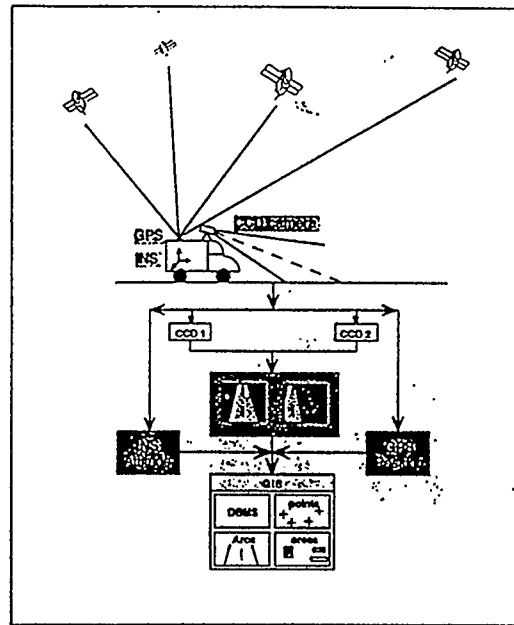


Figure 2.1 Hierarchical structure of the whole system

Through the postprocessing procedure, which integrates and processes GPS and INS data, accurate and complete camera orientation and positioning information is obtained and added to each record of image data, so that each record includes three geo-referenced images at one exposure station, and each P-line file includes hundreds of records. This kind of P-line file is consequently transferred to hard disk drive of the workstation - the Geostation.

To display multiple image frames in a window environment where all kinds of spatial objects can be measured and the results of the measurements in required format can be obtained, displayed and transferred. The corresponding Graphic User Interface (GUI) has two parts: multiview and overview. Here, one workstation with two monitors is our standard configuration.

2.2 Multiview subsystem

After the P-line data set is loaded onto the Geostation, a user can perform measurements of objects appearing in the overlapped area of the stereo images displayed on the Multiview screen. On the Multiview screen, three images (normally belonging to one record) are displayed simultaneously. Because all of the positioning and orientation parameters for each image are known, the 3D ground coordinates of desired points from at least two images can be measured (see Figure2.2). In the current version, we can select any image with the object on it within the image data set to do the intersection. This function is implemented by directly observing conjugate image points on the screen instead of using conventional photogrammetric instruments that are costly and require trained

operators. In this system, photogrammetric functions and GIS structure are integrated, so that it can be used in digital mapping, photo triangulation and some GIS applications.

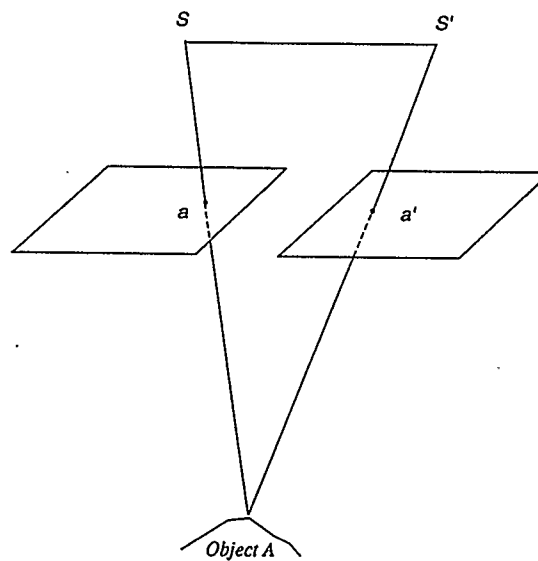


Figure 2.2 3D Determination through space intersection

One of the primary benefits of the digital approach is the efficiency to access images. Through the image navigational functions in Multiview, such as forward/backward, fast forward/backward or directly accessing records, we can easily select and display images along P-lines. Corresponding points of a desired object in different images can be zoomed in order to give a clearer view. In this way, both measurement efficiency and accuracy are improved. In GIS terms, geometric objects can be divided into three types: point, arc, and polygon. Also, each object may have its own non-spatial attributes, such as a telephone pole, a traffic sign, etc. We define all objects as GIS elements. When any measurement is performed, non-spatial attributes of the object

can be specified and be attached to the result - an integrated structural entity. Then the resulting GIS element is displayed on the Overview basemap in real time.

The Multiview module includes the master interface of Geostation (Qian, 1995). It consists of a menu bar, a tool window, several image display windows, and a status window displaying operating status, a message window which basically reports the message from intersection programs, and an option window in which users can change zoom factors (see Figure2.3). Since duplicated menu items and icons are available at the same level, functions can be accessed either through menu items or through icons. The main functions in the Multiview part are as follows:

- specify user configuration
- open/select a dataset
- navigate/display image records
- display camera geo-referencing data
- add/delete items in object attributes list
- define geometric type/object attribute
- choose observation points
- calculate 3D coordinates
- calculate basic geometric elements
- store measurement results in files
- control the communication with Overview
- automate measurement procedures

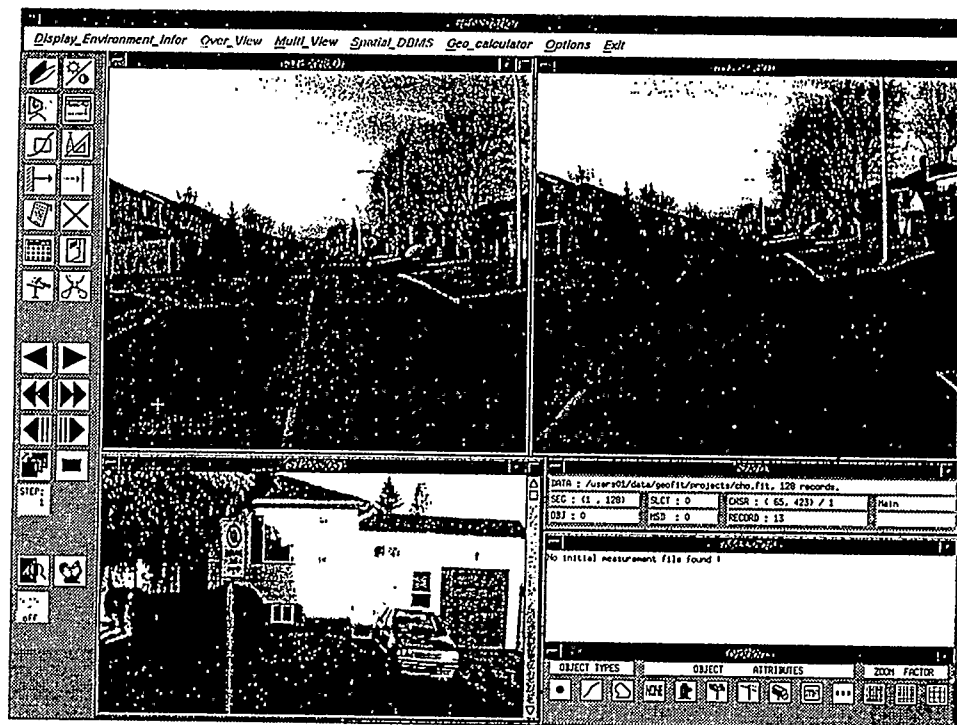


Figure 2.3 Graphic User Interface (GUI) of Multiview

2.3 Overview subsystem

The Overview module firstly gives the user a general idea of where the current image record is in the whole P-line. We can either use an existing softcopy map as the basemap or not use any basemap at all. The 3D results of the photogrammetric measurements of the digital images in Multiview are transferred to the Overview subsystem and are projected on the Overview screen in real time. Areas to be measured can be selected and zoomed. The current position of the images record and the camera orientations along the P-lines are also displayed (see Figure 2.4). When the image navigation function is used in Multiview, the corresponding image record position is also updated in Overview and this position is usually put at the center of the window in the

chosen displaying scale to supply the user a better viewing. Alternately, a specific image record at one exposure station can also be selected on the Overview screen. The image display on the Multiview screen would be updated at the same time.

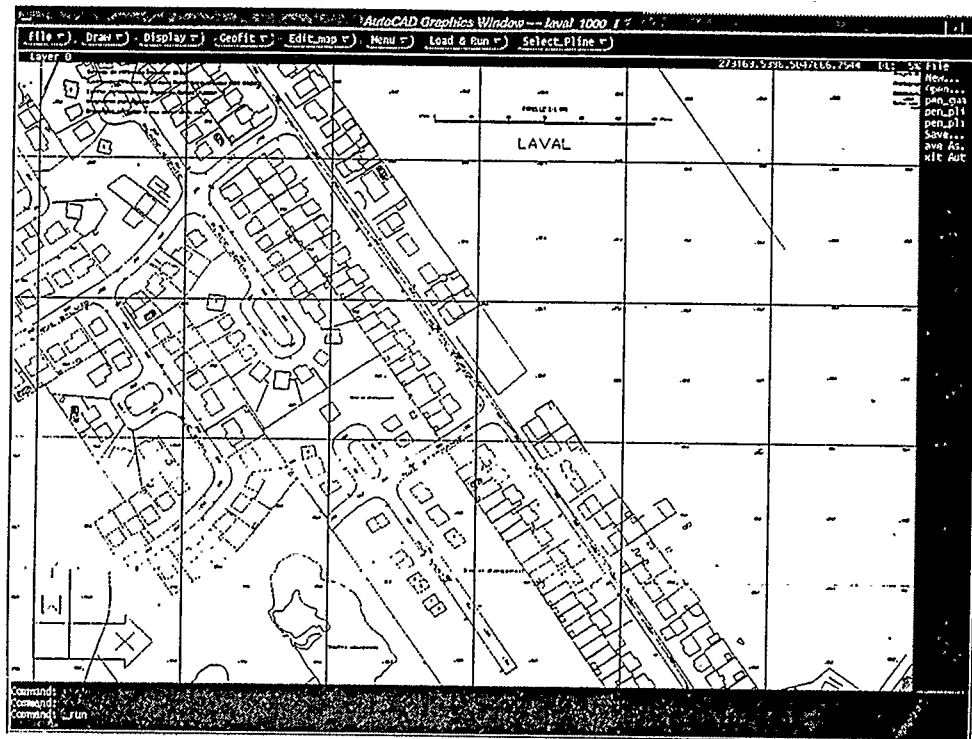


Figure 2.4 Graphic User Interface (GUI) of Overview

Major functions in this module include:

- loading a basemap in either vector or raster format;
- zooming in/out, pan display
- displaying P-line, projection of 3D measurement results
- tracking current exposure station and orientation
- selecting vertices for functions such as extending arc/polygon
- loading additional LISP functions
- selecting P-line data set by project name, period, P-line name or owner

- editing the results of photogrammetric measurements
- converting the results to DXF file for further applications

Overview uses AutoCAD R12 environment as a basis. It is also used as an editing tool for the objects in Overview. They can be deleted or modified on-line (Rogers, 1988). All of the displayed elements are in color. We keep the basic resulting element type such as point, arc and polygon. When a resulting element is transferred to Overview, it attaches its GIS attributes, so that different symbols are used to represent different attributes. The database is updated whenever the data field of any object has changed. At any time, all of the data in Multiview, Overview and database are kept consistent with each other.

The communication between Multiview and Overview is very important in the system. In the current stage, some specific files are used as communication buffer between the two subsystems. In the future, socket techniques will be used to replace these files in communication, in which case, it becomes more reliable, efficient and expandable (Stevens, 1990).

2.4 System configuration and GUI development

Geostation is runs under X-Windows 11 R5 and MOTIF 1.2 using a UNIX operating system. For the implementation of Graphic User Interface (GUI), a high level object-oriented development package WNDX was used as the toolkit to design the GUI of Multiview. This toolkit is portable from different workstations, such as IBM-PC and Machintosh desktop computers. It provides a single Application Program Interface (API)

for the developer that is supported in several GUI environments. The Overview module is interfaced with AUTOCAD R12 with CADOVERLAY V4.0. By using this package, we can easily access, display, and edit vector and raster basemap, and drawing files of measurement results. Also, we can make it easily interface to Multiview, GIS packages and some surveying and mapping systems.

One of the most significant improvements from conventional to softcopy photogrammetry is in the human-machine interface, which is commonly known as GUI. In Geostation system, raw image data are used and displayed on the screen in the form of bitmaps. GUI was developed, such that it supplies users a completely transparent access to photogrammetric software and other functions. The most important operation is the measurement function. In conventional and some other softcopy photogrammetric systems, an operator performs measurements through stereo viewing on pairs of photos or images. Although the operator can view the 3D stereo model, there are still several disadvantages: special displaying/viewing devices are needed; well trained operators are required; and, the number of images to form stereo pairs are limited. On the other hand, GeoStation displays the desired stereo images in separate viewing fields on the screen and the image pairs can be formed of any image wherever the chosen object appears. Beside these, the graphic utilities, such as on-line zooming and image enhancement, are easily implemented. When the operator locates the cursor on the object on the screen, the image coordinates can be obtained, the 3D ground coordinates can be calculated from at least two conjugate image points.

GUI supplies a user-friendly way to access all the functions and to change the options conventionally. The main GUI tools in Multiview are menus and icons. On a menu, related functions or commands can be simply grouped as sub-menus and users can quickly find and identify the desired function(s). This is an efficient way to arrange structured and logically related functions. For direct and quick access, icons are a good choice. By using this kind of on-line tool, the efficiency of the softcopy photogrammetric system is greatly improved. For the Overview part, all the functions are interfaced to AUTOCAD and CADOVERLAY environment. Through the development interface using AutoLisp, a GUI tool system was established which allows users to access or edit raster and vector maps, to run the user defined functionalities or to update current viewing. This tool system consists of pulldown/sidebar menus and dialogs. Here, both AUTOCAD/CADOVERLAY standard functions and user-developed functions work cooperatively with the user and communicate with Multiview. Through AUTOCAD output facilities, a GIS package can also be interfaced to GeoStation (see Figure 2.5).

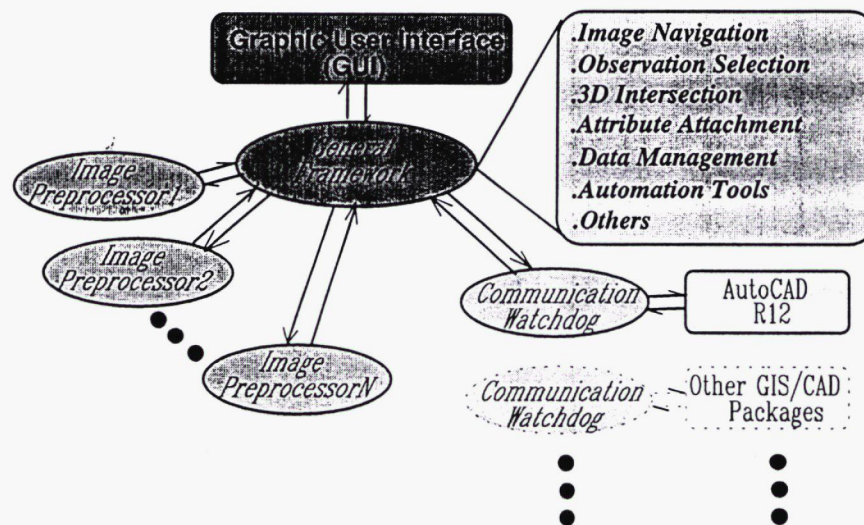


Figure 2.5 Geostation system architecture

Chapter 3

Main Functions of Geostation

3.1 Photogrammetric principles

The VISAT system has three functional modules:

- a) system calibration;
- b) data acquisition; and
- c) information extraction.

The purpose of calibration is to obtain the inner and relative orientation parameters between the components of the system, including three CCD cameras, GPS antenna and INS reference systems. Firstly, the relative angle orientation parameters between the INS coordinate system and the CCD cameras are determined. Also, the spatial offsets between GPS antenna, INS center and the perspective centres of the three cameras are modeled. At this time, a control ground site is used to perform this calibration. Because all of the unknown parameters are modeled in the form of observation equations, space resection or bundle adjustment algorithms can be applied to solve this mathematical model and the unknown parameters can be obtained. Since there is high correlation between different parameters, a bundle adjustment algorithm is preferred. Under this process, the whole system is united to a unique coordinate frame (Li, 1994). The spatial relations in the system are showed in Figure 3.1 (El-Shiemy et al, 1993).

Data acquisition includes the determination of positioning and orientation information at exposure time and the acquisition of image grey level data. At any time,

the status of the van can be described by six parameters. These parameters are derived from GPS and INS data. The INS provides the basic position and orientation information for the CCD cameras, and the GPS provides position and velocity information for the INS and the CCD cameras (Schwarz et al, 1993). The combination of GPS and INS technologies especially offers an advantage when there are some absences of either GPS or INS signals. For example, in downtown areas the GPS signals are obstructed by high buildings. At such time, the loss of positioning information is compensated by the INS data. In this way, the position errors are in the order of 10 cm with an RMS of 7cm (Schwarz et al, 1993). Digital images are taken by three CCD cameras (this number can be expanded to more) in the grey level range of 0 - 255.

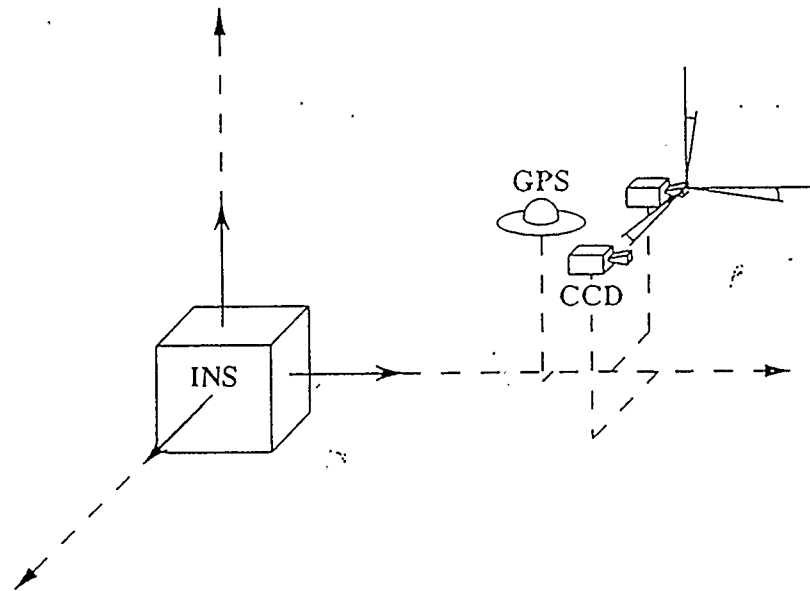


Figure 3.1 Spatial relationships in the system

The information extraction is implemented on the workstation - Geostation. Here, both basic photogrammetric theory and advanced image analysis strategies are applied. When an operator digitizes any object appearing in different images (at least two) on the Multiview screen, 3D spatial coordinates of this object can be calculated by using the collinearity condition which is the most popular mathematical model used in photogrammetry. This model describes the spatial relationships among image coordinates, perspective centers and object-space coordinates. Theoretically, the three points should lie on the same projective line. Normally, this relationship can be expressed in the following form:

$$x_i - x_p = -f \frac{m_{11}(X_i - X_0) + m_{12}(Y_i - Y_0) + m_{13}(Z_i - Z_0)}{m_{31}(X_i - X_0) + m_{32}(Y_i - Y_0) + m_{33}(Z_i - Z_0)} \quad (3.1a)$$

$$y_i - y_p = -f \frac{m_{21}(X_i - X_0) + m_{22}(Y_i - Y_0) + m_{23}(Z_i - Z_0)}{m_{31}(X_i - X_0) + m_{32}(Y_i - Y_0) + m_{33}(Z_i - Z_0)} \quad (3.1b)$$

where:

- x_i, y_i are image coordinates of point i ;
- x_p, y_p are the coordinates of the principal point expressed in the image coordinate system;
- X_i, Y_i, Z_i are the object-space coordinates of the point;
- f is the focal length of the camera;
- X_0, Y_0, Z_0 are the object-space coordinates of the perspective center;
- m_{ij} are the i (row) and j (column) coefficients of the rotation matrix.

For each image, we can get the above two equations, so that the object-space of a point coordinates X_i, Y_i, Z_i can be obtained by using two geo-referenced images.

3.2 Epipolar geometry

In a softcopy photogrammetric system, efficiency and reliability are the most important concerns. For the purpose of improving the above factors in the measurement procedure, the application of epipolar geometry is advantageous. Epipolar geometry is a classic concept in photogrammetry. An epipolar plane contains the base line, which links the two perspective centres and includes the corresponding points on the two images, and the object point. The epipolar plane intersects the two image planes along epipolar lines. Therefore, two conjugate image points should lie on the corresponding epipolar lines (Lu, 1988). In this way, with known points in the left image, search for the conjugate points in the right image can be performed along corresponding epipolar lines on the right image. Therefore, two dimensional search is simplified into one dimensional search. (Mazry, 1974)

There are many methods to determine the epipolar line. The following model was implemented in Geostation:

- 1.) Transformation from digital image coordinates to conventional photo coordinates.

(Figure 3.2a shows transformation from digital image coord to photo coord system)

$$x = x_{dig} - x_p \quad (3.2)$$

$$y = (y_{dig} - y_p) / k_y \quad (3.3)$$

where

- x, y are the coordinates of the point in a conventional image system;
- x_{dig}, y_{dig} are the coordinates in a digital image system;
- x_p, y_p are the coordinates of principal point in the digital image system;
- k_y is the scale factor in y direction.

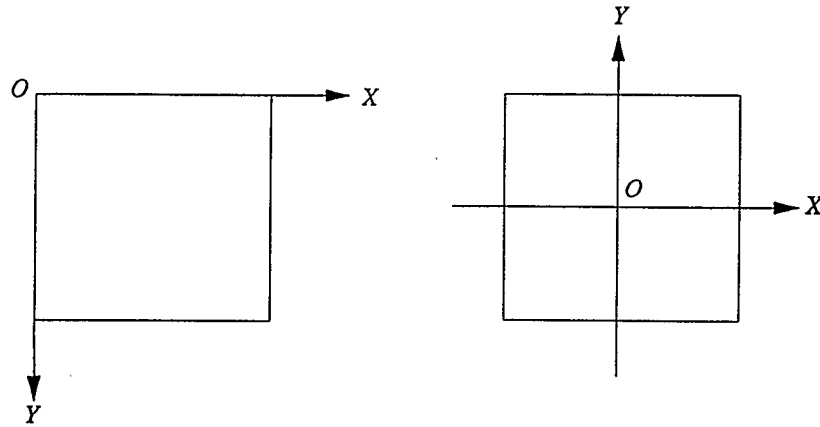


Figure 3.2a Transformation from digital image coord to photo coord

2.) Transformation from photo coordinate to object-space coordinate system (Figure 3.2b):

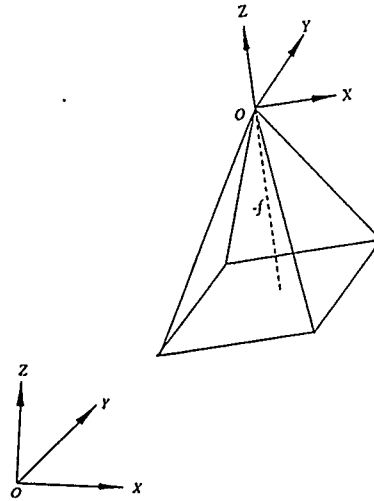


Figure 3.2b Transformation from photo coord to object coord

$$R_{lii} = R_{id} * R_{pi} * R_{oi} \quad (3.4)$$

where

I is the identification number of the corresponding image;

$R_{\omega i}$, $R_{\varphi i}$, $R_{\kappa i}$ are the rotation matrices about x, y, z axes, respectively.

R_{II} is the rotation matrix from image space coord. to spacial coords.

3.) Transformation from object-space coordinate to base line coordinate system:

$$B_X = X_{0i} - X_{0j} \quad (3.5a)$$

$$B_Y = Y_{0i} - Y_{0j} \quad (3.5b)$$

$$B_Z = Z_{0i} - Z_{0j} \quad (3.5c)$$

$$\alpha = \tan^{-1}(B_Y / B_X) \quad (3.6)$$

$$\beta = \tan^{-1}[B_Z / (B_X^2 + B_Y^2)^{1/2}] \quad (3.7)$$

$$R_{II} = R_{\beta} * R_{\alpha} \quad (3.8)$$

where

B_X , B_Y , B_Z are the projections of base length B on X, Y, Z axes, respectively

α , β are rotation angles about Z, Y' axes (Figure 3.2c), respectively

R_{α} , R_{β} are rotation matrices about Z and Y axes, respectively

R_{II} is the rotation matrix from spacial coords. to base line coords.

Combining the above three steps, the transformation from image coordinate to base line coordinate system is:

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = R * \begin{bmatrix} x \\ y \\ -f \end{bmatrix} \quad (3.9)$$

where

$$R = R_{II} * R_I = \begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix} \quad (3.10)$$

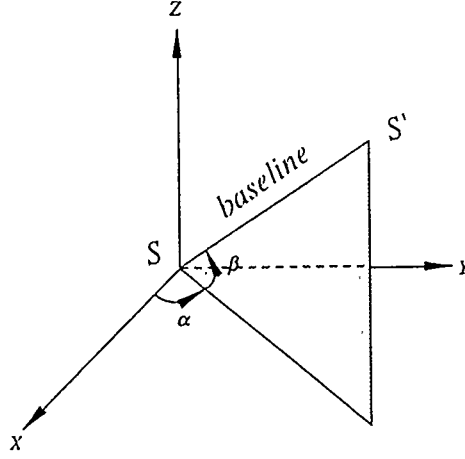


Figure 3.2c Transformation from space coord to baseline coord

In the base line coordinate system, we can derive the coplanarity condition equation:

$SS' \bullet (S_a \times S_b) = 0$, that is

$$\begin{vmatrix} B & 0 & 0 \\ U_a & V_a & W_a \\ U_b' & V_b' & W_b' \end{vmatrix} = B * \begin{vmatrix} V_a & W_a \\ V_b' & W_b' \end{vmatrix} = 0 \quad (3.11)$$

where

$$V_a = B_1 * x_a + B_2 * y_a - B_3 * f \quad (3.12a)$$

$$W_a = C_1 * x_a + C_2 * y_a - C_3 * f \quad (3.12b)$$

$$V_b' = B_1' * x_b + B_2' * y_b - B_3' * f' \quad (3.12c)$$

$$W_b' = C_1' * x_b + C_2' * y_b - C_3' * f' \quad (3.12d)$$

Here, the rotation matrix can be calculated from the geo-reference information corresponding to the images. In the equation (3.11), the number of unknowns is four. They are x_a , y_a , x_b and y_b . When clicking the point in the left (first) image, x_a, y_a are defined, and given one of x_b and y_b , the other one can be solved. In fact, point b can be fixed on one of the borders in the right (second) image. In this case, either x_b or y_b equals to zero, therefore, the epipolar line is determined.

3.3 The space intersection and bundle adjustment

The collinearity condition is the basic mathematical model used in calculating 3D coordinates. By using two stereo images, we can determine the 3D coordinates of an object through simple space intersection. In this system, the distance between an object and the camera is generally limited to 50 meters. The main reason for this constraint comes from the geometry of the intersection and the image resolution. A camera base line length of about 2.0 m is currently being used for the prototype system. From the view point of photogrammetry, the largest error appears in the along track. The ideal ratio of base line length to object distance is larger than 1/15, the acceptable ratio is larger than 1/25. On the other hand, when we perform measurement on the discrete digital image, the measuring accuracy we can reach is 1/2 pixel. This means that the digitizing error is within 1/2 pixel. The CCD camera we used has the viewing field of 512 in the horizontal by 480 in the vertical direction. So an error of 1/2 pixel causes an along-track error of 0.7 m in object distance over 50 m. Hence, if the distance to the object can be kept within 25 m, the corresponding along-track error can be reduced to 0.3 m. A significant improvement can be implemented by applying the following updating:

- Perform sub-pixel digitizing measurement. e.g. if 1/3 pixel accuracy is achieved, the corresponding maximum along-track error is reduced to 10 cm;
- Increase the base line length. e.g. The optimal combination of images can be selected from different exposure stations;
- Use multi images to calculate the object coordinates. The influence from the geometry can be reduced.

Aside from the above factors, the lens distortion of CCD camera also contributes to the final object error. Although it is considered in the calibration, it is sometimes hard to obtain perfect solutions because of the high correlation with other parameters (Shortis, 1988). Through the above discussion, we arrive at the conclusion that an adequate algorithm is needed to perform the photogrammetric triangulation. The flexibility of the bundle adjustment algorithm allows it to meet these requirements, thus, the bundle adjustment is chosen as the mathematical model of triangulation (Chapman, 1992).

The bundle adjustment is still based on the collinearity equations. Considering lens distortion and the transformation from digital image coordinate to photo coordinate system, Equation (3.1) can be written into the following format:

$$\bar{x} + \bar{x}(k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2\bar{x}^2) + 2p_2\bar{x}\bar{y} + f(m/q) = 0 \quad (3.13a)$$

$$\bar{y} + \bar{y}(k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1\bar{x}\bar{y} + p_2(r^2 + 2\bar{y}^2) + f(n/q)k_y = 0 \quad (3.13b)$$

where

$$\bar{x} = x - x_p$$

$$\bar{y} = y - y_p$$

$$m = m_{11}(X - X_0) + m_{12}(Y - Y_0) + m_{13}(Z - Z_0)$$

$$n = m_{21}(X - X_0) + m_{22}(Y - Y_0) + m_{23}(Z - Z_0)$$

$$q = m_{31}(X - X_0) + m_{32}(Y - Y_0) + m_{33}(Z - Z_0)$$

$x, y, x_p, y_p, X, Y, Z, X_0, Y_0, Z_0, k_y$ and m_{ij} as defined before, k_i, p_j are lens distortion coefficients from a camera calibration.

All the exterior orientation parameters and some of the interior orientation parameters are updated in the procedure. This flexible analytical tool offers the compensation of system errors by its self-calibration process. Additionally, either conventional stereopairs or multistation images can be incorporated in the unique adjustment procedure. Finally, highest accuracy can be achieved economically (Granshaw, 1980). In the prototype system, up to six points from six images can be used in the triangulation process to solve for the optimized results. Tests have shown that both the accuracy and reliability are improved by using multistation images in bundle adjustment.

3.4 Database generation and GIS link

3.4.1 Database generation

In addition to a simple description of geometry, such as type and coordinates, a measured object is associated with some non-spatial attributes as well, to become a GIS database element (Novak, 1992). These attributes include major type, attribute, time, coordinates, accuracy, layer name and other information. In the measurement procedure in Geostation, the operator can set the properties of data record associated with the

measurement result(s) according to different types. When the measurement is completed, it will be stored in the memory in the specified format that is acceptable to the GIS databases. In this procedure, every 3D coordinate triplet represents a vertex, with location updated on both Multiview and Overview upon creation. The data format is as follows:

Object ID

Time of acquisition

Time of measurement

Type

Attribute

Layer name

Number of vertex

Vertex index

x mx

y my

z mz

Vertex index

x mx

y my

z mz

..

Where x , y , z are the 3D coordinates of the vertex, m_x , m_y , m_z represent the corresponding standard deviations.

Blocks are usually used as the basic unit to which can be attached attributes and other information. Currently a block uses a format compatible with AUTOCAD DXF. One block may consist of several geometrical elements plus some non-spatial elements. For example, a road can be recorded as a block which consists of two arcs (polylines) with road name, etc. But a fire hydrant may only include 3D coordinates plus its type etc.

On the other hand, AUTOCAD is used as an editing tool for these objects in the Overview environment. During measurement interaction, objects can be deleted or modified on-line. The database is updated whenever the content of the data structure of any object has been changed. All of the data in Multiview, Overview and the database are kept consistent with each other at all times.

Additionally, we can make full use of AUTOCAD editing and database functionalities to process measurement results. In this case, both on-line processing or postprocessing may be performed. AUTOCAD is not just a drafting program but also a database manager. Although Autolisp or ADS allows users to manage and manipulate graphic and nongraphic data in ways that are impossible with AUTOCAD alone, the AUTOCAD database supplies higher efficiency in terms of managing huge data set and linking with other GIS package(s). We can add attribute information to a block to create an alphanumeric database within the drawing database. This database within a database can act as a link between the graphic information in the drawing and a more sophisticated and extensive database outside the drawing.

In the AUTOCAD internal database, a template file, which contains the desired structure for the attribute extract file should be created. It tells the ATTEXT (ATtribute EXtract) command which attribute tags to look for, and in what order to place the information contained in those attributes. Following is a simple template file that tells ATTEXT to extract any values associated with attribute tags ROADNO, AQTIME, MTIME and LYNAM from any blocks named ROADTAG:

```
BL:NAME  c00800  (Block name, 8 chars)

ROADNO   c00500  (Road number, 5 chars)

AQTIME   c00600  (Acquisition time, 6 chars)

MTIME    c00600  (Measurement time, 6 chars)

LYNAME   c01500  (Layer name, 15 chars)
```

The database file produced by using this template file should look like this:

```
ROADTAG 407 941120 950610 Laval Civil

ROADTAG 510 940818 950612 Montreal Civil
```

In fact, there is a practical limit to the number of attributes that can be linked to an individual block directly and the larger number of attributes may result in the consequence of information inflation, which causes a huge and slow drawing. Therefore, it would be a better strategy to track only identification and quantity data in the drawing, and then link that information to more detailed data in external database managers, such as ORACLE, Sybase etc. The outside database lets users store more detailed information about items referred to in the drawing. Another advantage of using (an) external database(s) to store

all kinds of attributes is the possibility for linking the Geostation system with other GIS packages or transferring the results to other CAD systems.

3.4.2 GIS link

At the current stage, DXF (Drawing eXchange Format) files are commonly used for transferring drawing database information among CAD or other vector based GIS systems. Geostation also uses this format to make the link with Arc/Info, which is widely used in many areas relating to geographical information. In the Overview environment, we can easily convert the drawing format file into DXF file by using the AUTOCAD DXFOUT command (Frederic, 1991). On the side of Arc/Info, the DXF file can be converted into an Arc/Info format spatial data set, which includes all the points, arcs and polygons, by using the DXFARC function.

The conversion or link between different systems is sometimes a difficult job. The main problems we met are: Firstly, this drawing format conversion is not 100 percent complete. For example, Arc/Info does not always have the same entity types as AUTOCAD. It does not accept some AUTOCAD format data structures or elements directly, such as some block entities and some layer information. Secondly, this transformation can only cover the basic information, although there is a great number of spatial and non-spatial data that need to be transferred. In this case, external database can play an important role. At the same time, the external database can be set up according to the designed data format. These data can be extracted either from the AUTOCAD internal database or directly from the data structure of measurement results.

The information system in Arc/Info consists of a spatial data set (the drawing information) and an attribute data set (the description of objects), which can provide information for:

- determining and maintaining the location of the measured objects in relation to features represented on the map,
- displaying a base map graphically and showing the attribute data, and
- merging both spatial and non-spatial information to some existing data set.

Based on the Arc/Info modules, the system can perform data query, display and analysis functions for GIS application. It can also supply a Graphic User Interface (GUI) by generating a User-friendly Menu System. Figure 3.3 shows the main structure of the system.

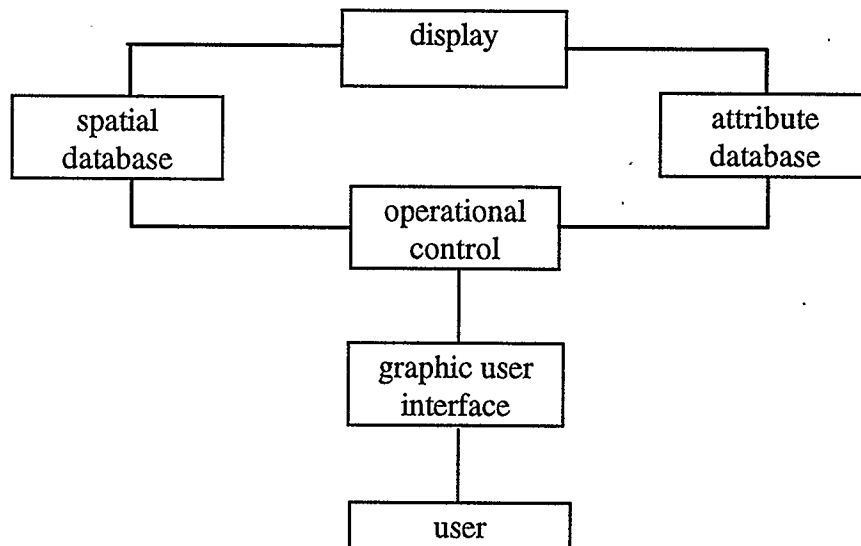


Figure 3.3 Main structure of the GIS system

Arc/Info system consists of two modules, ARC and INFO. ARC handles the spatial data set, and INFO handles the attribute data set. To generate all the attribute tables in the relational database and to connect the attribute tables to the points, arcs, and polygons, we need to use CLEAN and BUILD functions to create basic AAT (Arc Attribute Table) and PAT (Polygon Attribute Table) files so that we can add more description information for each element.

In AAT and PAT tables, the basic items are given in Table 1.

Table 1: Basic AAT and PAT

BASIC ITEMS OF AAT	BASIC ITEMS OF PAT
Entity Code #	Record #
From Node #	Area
To Node #	Perimeter
Left Polygon #	Polygon #
Right Polygon #	Polygon ID
Arc Length	
Arc #	
Arc ID	

Among these items, entity code # and record # are important because they are the inner connection between attribute database and spatial database. In the GIS application, all the items can be used as retrieval key to search for any spatial or attribute information. In practice, we can directly use the attribute data sets obtained in Geostation as the INFO

file and connect them with the spatial data. The most difficult task in the GIS link or transformation is to connect the attribute with spatial data. There are two ways that can be used to meet this requirement:

- 1.) In the editing process, entity code # and record # can be obtained manually through ARCPLOT IDENTIFY function. These ID can be added to the corresponding AAT or PAT file to form a new field (line). Each field includes one ID and several attribute items.

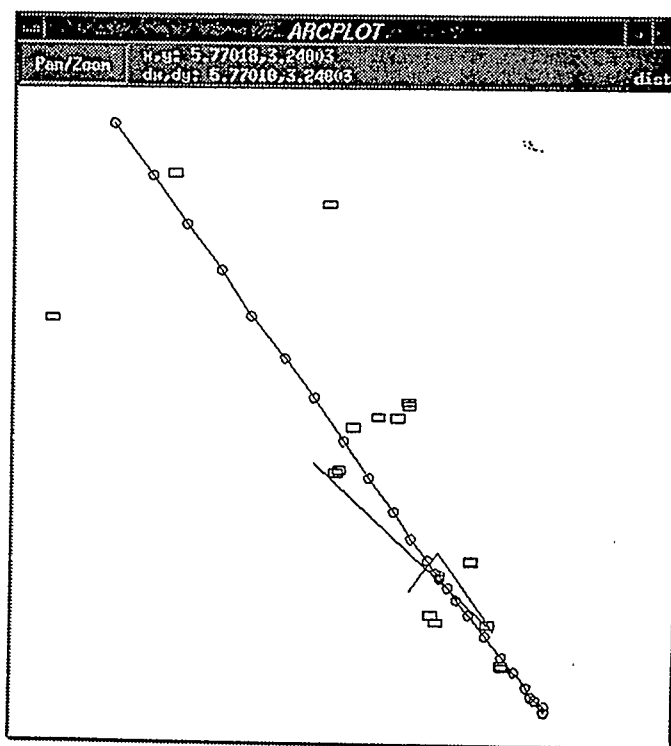


Figure 3.4 Display of measurement results in Arc/Info.

- 2.) In an automatic way, the Geostation measurement result file is used as INFO file and transferred to the INFO directory (Figure 3.4). Because AUTOCAD receives measurement results from GeoStation in blocks, with each block associated with

identification code which is the same as the corresponding item in INFO record whenever a block on the AUTOCAD is selected, its identification code can be obtained immediately. Therefore, the non-spatial attributes about that object can be found in INFO tables directly. Additionally, the operator may select and measure objects according to a specific theme, such as manholes or road center lines etc. in order that different kinds of results may be saved to a separate layer in AUTOCAD and DXF file through AutoLisp function. It will make the connection much easier.

Chapter 4

Image Processing

A very important component of digital image handling is image display which forms the basis of photogrammetric measurement procedures. Most digital photogrammetric systems display stereo images using a certain type of viewing devices connected to computers; thus permitting the viewing of the three dimensional model of the object can be conducted (Mori et al., 1992). But in the Geostation system, the display and measurement procedures are directly performed on the Multiview screen. In the VISAT implementation, huge numbers of digital images were taken, among which a considerable quantity of raw images acquired has poor contrast and low intensity because of poor exposure conditions and low camera quality. These images result in poor recognition of object features and a decrease in measurement accuracy. To enhance the images, different strategies have been implemented and tested in the system.

4.1 Grey level enhancement

Because a large number of images needs to be processed, efficient and practical global strategies are required. In the grey level enhancement, we use the point operations to process images, which modify the grey scale pixel by pixel and each processed pixel corresponds exactly to the pixel in the input image without any change of the coordinates.

4.1.1 Histogram equalization

Histogram equalization is one of the most popular forms of enhancement. It is based on the view point that the ideal distribution of grey level values is a uniform

distribution for most classes of images. Hence, a uniform distribution of grey level values equalizes each of the quantization levels to enhance the low contrast. The basic procedure of this algorithm can be listed as follows:

- 1.) Choose the value of the upper limit of output grey level, L , which is 255 in our case.
- 2.) Calculate the histogram of the original image. Each value contains n_i pixels to a total of N pixels in the entire image. The grey level frequency is:

$$P(r_i) = \frac{n_i}{N} \quad (4.1)$$

- 3.) Form the cumulative distribution curve function:

$$s_j = T(r_j) = \sum_{i=0}^j P(r_i) = \sum_{i=0}^j \frac{n_i}{N} \quad (4.2)$$

where

$0 \leq j \leq L$ and T is the transformation function of the histogram equalization.

- 4.) Invert the transformation function and calculate the updated grey level values:

$$r_j = T^{-1}(s_j) \quad 0 \leq s_j \leq 1, \quad (4.3)$$

Here s_j represents transformed grey level value. Since only a few equally spaced levels are allowed in a normal case, each of the transformed values must be assigned to its closest valid level. For example, s_j is rounded to the closest integer value l_j , that means:

$$s_j = \frac{l_j}{L} \quad (4.4)$$

The grey level value l_j will be assigned to the pixels with the original grey level value r_j . Figure 4.1 shows the original and processed image. From some results, we can

see that although the equalization method enhances most image details, the result images look unnatural and are not smooth.

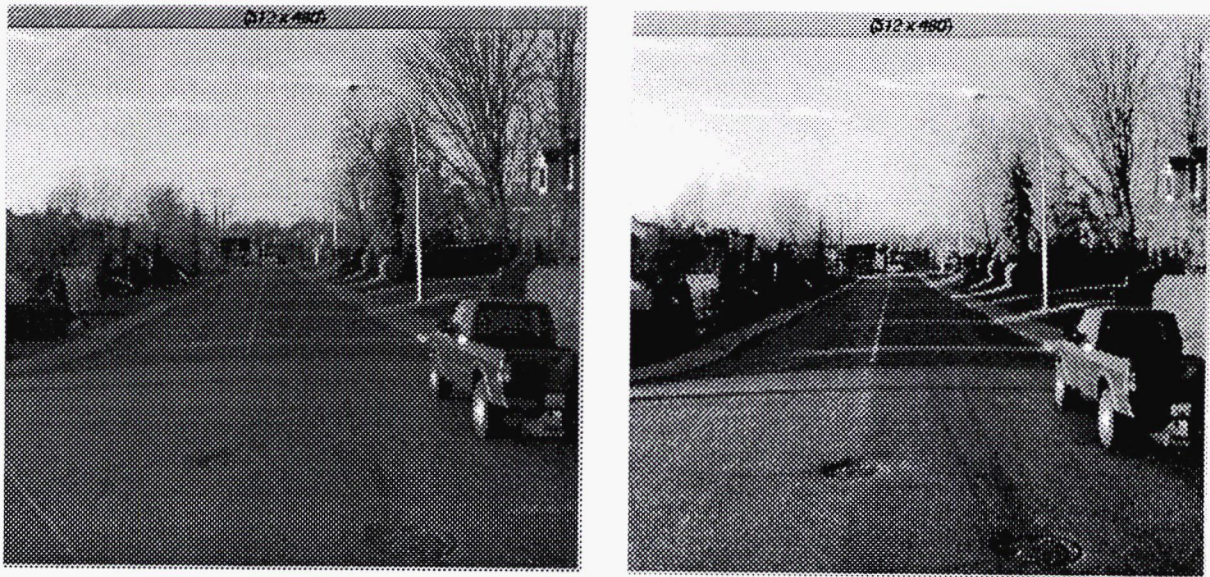


Figure 4.1 Original and equalized images

4.1.2 Linear stretch

When the exposure is poor, the grey level values in an image are rarely distributed in an appropriate pattern either in a narrow range or in several narrow ranges (Schowengerdt, 1989). In our case, it results in a decrease of both contrast and brightness. Therefore, linear contrast stretch can be used to expand grey level over the full range and to improve the brightness.

A simple linear stretch enhances the grey level of an image with a narrow range in a direct mode. Assuming that the image grey level covers a range (min, max), but the preferred range is (a, b). The transformation model is:

$$G(x, y) = \frac{g(x, y) - \min}{\max - \min} * (b - a) + a \quad (4.5)$$

where

$g(x, y)$ is the original grey level value at position (x, y) .

$G(x, y)$ is the corresponding resulting grey level value.

Normally, the grey level range of the resulting image is expanded to (0,255).

Through our experiment, it is evident that the grey level values of most images are distributed in 2-3 narrow ranges. As an example of multimodes, the bright sky is represented in one range, and most features belong to another narrow range. To stretch these images more appropriately, multi-segments linear stretch methods are applied. Assuming that *mean* is the mean of the grey level values of an entire image, and *pmean* is the preferred mean. The two segments linear stretch model is as:

$$G(x, y) = \frac{pmean}{mean - \min} * (g(x, y) - \min), \quad g(x, y) < mean, \quad (4.6a)$$

otherwise

$$G(x, y) = \frac{255 - pmean}{\max - mean} * (g(x, y) - mean) + pmean \quad (4.6b)$$

Figure 4.2 shows the transformation and the processed result. The test of the processing of a great number of images indicates that the two segments linear stretch can both enhance the grey level contrast and improve the brightness. In our system, we have a way to detect peaks of the histogram. Based upon the result of analysis, we conclude that a three segments linear stretch is the best strategy.

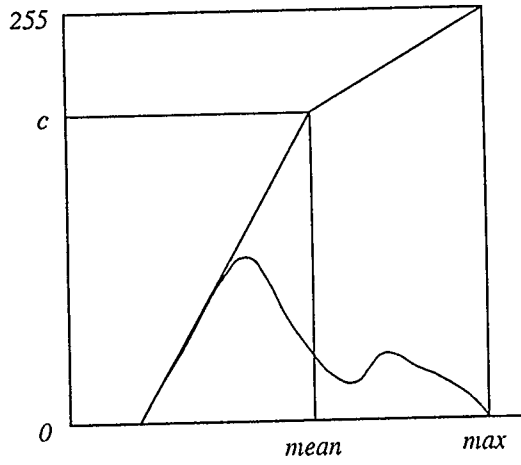


Figure 4.2 Two-segment linear stretch and processed result

The model is as follows:

$$G(x, y) = \frac{a'}{a - \min} * [g(x, y) - \min], \quad g(x, y) \leq a; \quad (4.7a)$$

$$G(x, y) = g(x, y) - a + a', \quad a < g(x, y) \leq b; \quad (4.7b)$$

$$G(x, y) = \frac{255 - [a' + (b - a)]}{\max - b} * [g(x, y) - b] + [a' + (b - a)], \quad g(x, y) > b \quad (4.7c)$$

where

$$a = \text{mean}$$

$$b = \text{mean} + 30$$

$$a' = 150$$

$$\text{mean} = \sum_i gL_i / N$$

After many tests, we found that most images are of a two-peak distribution type. From the grey level mean to mean + 30, we do not change the contrast because this segment is very clear for human viewing. For the first segment, however, we enhanced both the grey level and the brightness. This segment contains most details and features. Figure 4.3 shows the transformation and the processed result. We can also see the change of the histograms in Figure 4.4. In this way, we can greatly improve the image quality. This adaptive multi-segment linear stretch gives most reasonable results. It works for under-exposed images, over-exposed images, and normal images. Using this algorithm, we can preprocess a great number of raw images and also process individual images on-line.

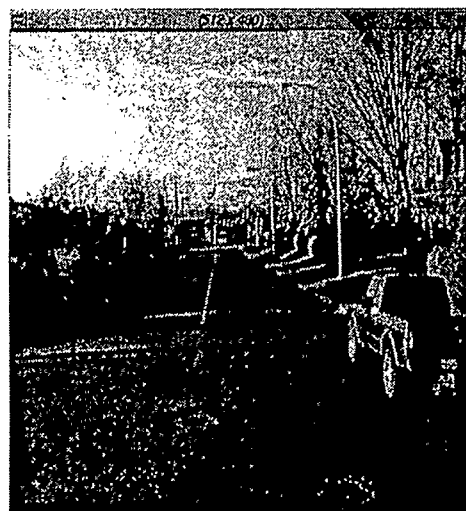
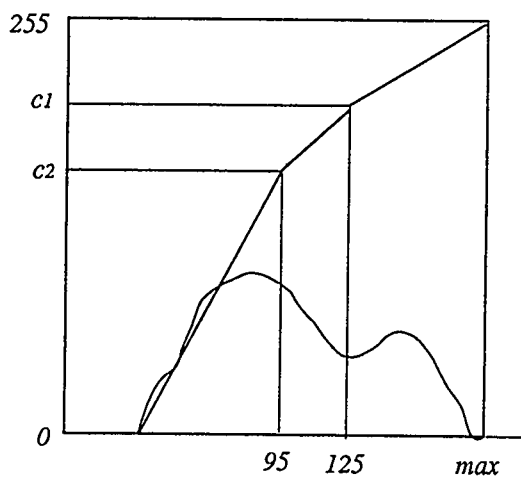
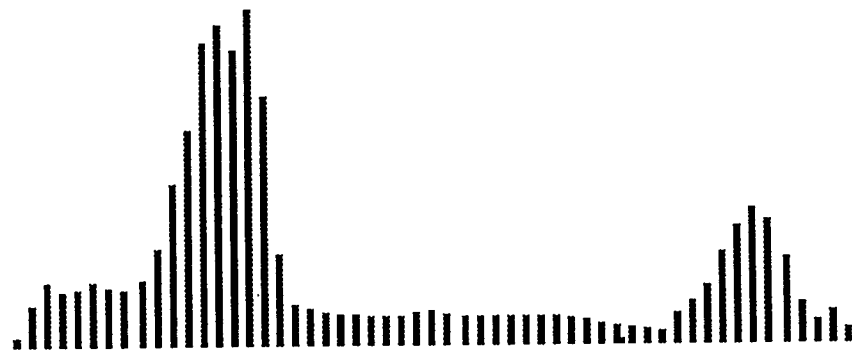
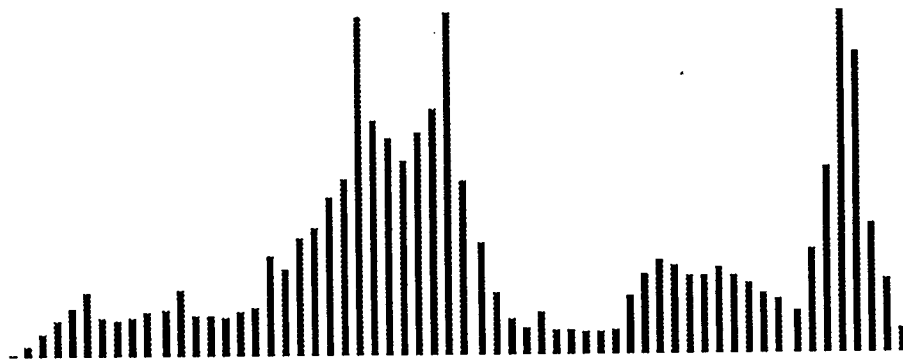


Figure 4.3 Three-segment stretch and processed result



Histogram Before Enhancement



Histogram After Enhancement

Figure 4.4 Change of histogram before and after processing

4.2 Image sharpening

Although the grey level enhancement improves both the brightness and contrast, some images are blurred in the edges of features. Therefore, image sharpening techniques are implemented and improved in order to provide a better recognition of features.

1.) Solbel filter method

Generally speaking, differentiation can have a sharpening effect on a blurred digital image. The most commonly used method of differentiation in image processing applications is the

gradient (Gonzalez 1987). Given a grey scales function $g(x,y)$, the gradient of g at coordinates (x, y) is defined as the vector:

$$G[g(x, y)] = \begin{bmatrix} \partial g / \partial x \\ \partial g / \partial y \end{bmatrix} \quad (4.8)$$

and the magnitude of $G[g(x,y)]$ can be given and simply written as:

$$G[g(x, y)] = [(\partial g / \partial x)^2 + (\partial g / \partial y)^2]^{1/2} \quad (4.9)$$

which represents the maximum rate of increase of $g(x, y)$ per unit distance in the direction of G . The most important steps are the detection and the enhancement of the edges on an image. For the discrete gradient digital images, some gradient filters can be used to find the edge pixels. Among these filters, a Sobel filter is commonly used. It is a direction dependent gradient filter. For a given pixel in position e , the 3×3 neighborhood a, b, \dots, i (see Figure 4.5a) is chosen and the grey level values of this neighborhood contribute the gradient components G_x and G_y of the position e :

$$G_x = (g+2h+i) - (a+2b+c) \quad (4.10a)$$

$$G_y = (c+2f+i) - (a+2d+g) \quad (4.10b)$$

a	b	c
d	e	f
g	h	i

Figure 4.5a 3x3 neighborhood for Sobel filters

Using Eq (4.10) the gradient amplitude G for position e can be calculated. Then the edge enhancement can be carried out. A global threshold is determined. If $|G| > \text{threshold}$ (or $|G_i| > \text{threshold}$, $i=x$ or y),

$$e = e + \text{scale} * G_i \quad (4.11)$$

where

$$G_i = \max\{|G_x|, |G_y|\} \text{ and } 0 < \text{scale} < 1.$$

This method can be concluded as an “addback” method (Pan, 1990). After the processing, the edges look sharper than before.

The disadvantage of a Sobel filter is that only the derivatives in x and y directions are considered and the x - y direction is ignored. This will result in losing some edge pixels. As an improvement, we consider an expanded neighborhood of e (see Figure 4.5b) and calculate

$$G_{xy} = 4/6[(j+2d+2h+u) - (p+2b+2f+o)] \quad (4.12a)$$

$$G_{yx} = 4/6[(m+2f+2h+s) - (r+2b+2d+l)] \quad (4.12b)$$

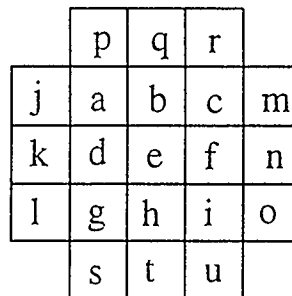


Figure 4.5b Expanded neighborhood

If $|G_i| > \text{threshold}$, $i=x$ or y , then

$$e = e + \text{scale} * G_i \quad (4.13)$$

where

$$G_i = \max\{|G_x|, |G_y|, |G_{xy}|, |G_{yx}|\} \quad (4.14)$$

In this way, we can enhance edges in every direction. Here the global threshold should be chosen properly. The scale factor is normally chosen in the range of 0.3 -- 0.7. If the scale factor is too small, the sharpened effect is not obvious, and when it is too large, the processed image will look unnatural. Figure 4.6 shows a result of processing.



Figure 4.6 Result of Sobel "addback"

2.) Grey level difference method

The improved Sobel filter method gives us good results, but sometimes it is not very sensitive. Here we developed another edge sharpening algorithm for the neighborhood of the position a is shown in Figure 4.7. The grey level derivatives are calculated as follows:

$$\begin{aligned}
 G_1 &= e - a & G_2 &= e - b \\
 G_3 &= e - c & G_4 &= e - d \\
 G_5 &= e - f & G_6 &= e - g \\
 G_7 &= e - h & G_8 &= e - i
 \end{aligned} \tag{4.15}$$

then

$$c = c + \text{scale} * G, \quad \text{if } |G| > \text{threshold.} \tag{4.16}$$

where

$$G = \max\{G_i\} \quad i = 1, 2, \dots, 8. \quad 0 < \text{scale} < 1$$

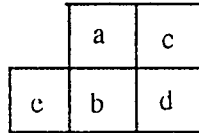


Figure 4.7 Grey level differential mask

In this approach, we still use the “addback” process. Through our implementation, we obtained very good visual results. The processed images are sharpened and more detailed feature information can be easily recognized. Therefore, more accurate measurements can be achieved. Figure 4.8 shows a result of this processing.

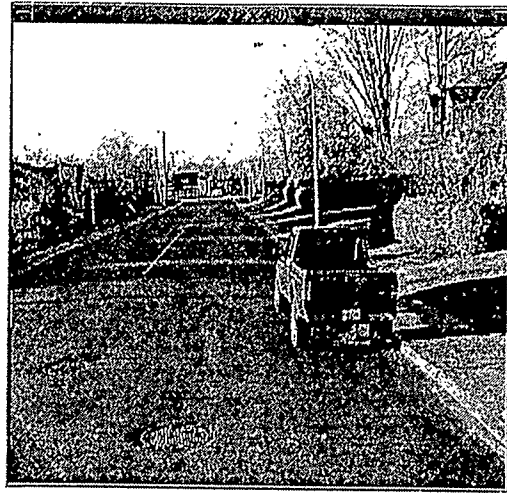


Figure 4.8 Processed image using grey level “addback”

The above “addback” image sharpening methods can be used in global image processing, or as an on-line image sharpening tool.

Chapter 5

Single Point Image Matching

Because of the rapid development of computer science, more and more research topics in digital photogrammetry have become practical concerns. One of the major difficulties in digital photogrammetry is that of the automatic determination of image coordinates of corresponding points for the computation of object space coordinates. This topic can be defined as image matching, and image matching is the location of corresponding phenomena in different images. The term of stereo matching refers to the positioning of an object in space and the reconstruction of surface from stereo pairs. Matching models are usually divided into two kinds: area based and feature based (User, 1988, Argialas, 1990). In our system, both methods have been improved and implemented.

5.1 Area-based single point matching

Area-based matching is performed point by point, and small rectangular pixel arrays with the interest point at the center are used as the basic unit. They are given in the form of a matrix array $g(x,y)$ of observed grey level values. The computational task is then to obtain an optimal matching by transforming one array of grey level values into the other (Ackermann, 1984). In this chapter, we mainly discuss single point matching, which is a semi-automatic matching method because the first chosen point needs to be selected interactively.

5.1.1 Correlation matching

One of the most commonly used area-based single point matching is the cross correlation technique. Although it is an old method, it works well most of the time, especially in the aerial photogrammetric images. It encounters difficulties, however, when two images are taken from quite different viewpoints of a scene that does not contain adequate visual texture, such as close-range photogrammetric images. In this approach, the image point to be matched is the center pixel of a small array (or small window) in the master (left) image. The grey level value for each pixel in this array is statistically compared with values in a same sized window as it moves pixel by pixel in a specific, larger search area in the slave (right) image (see Figure 5.1).

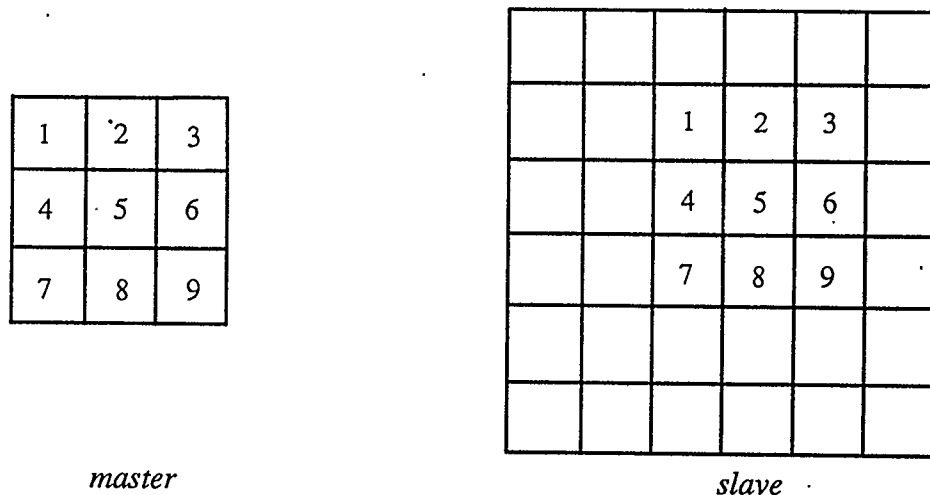


Figure 5.1 Target and searching masks

In semi-automatic mode, when the object is selected manually in the left image, a correlation coefficient is calculated for each possible position assumed by the target area in the right image. The maximum value of the coefficient factors is determined for

comparing all the corresponding coefficient factors within the search area. Then the matched position in the right image is obtained. Hence, the 3D space coordinates of this target can be calculated by using the space intersection method.

There are four main factors responsible for the grey values in an image: illumination, reflectivity, geometry, and viewpoint (Marr, 1980). To achieve unambiguous matches, all the four factors should be considered, but in fact, only the positions and the orientation of the exposure are known. Therefore, different mathematical models and constraints have been derived to reduce these influences.

There are several math models for the calculation of correlation coefficients for single point matching. In our system, the following ways are implemented and tested:

1.) cross-correlation coefficient

$$R_{ts} = \sum_i g_i^t * g_i^s / \sqrt{\sum_i (g_i^t)^2 * \sum_i (g_i^s)^2} \quad (5.1)$$

2.) normalized cross-correlation coefficient:

$$R_{ts} = \sum_i (g_i^t - \bar{g}^t) * (g_i^s - \bar{g}^s) / \sqrt{\sum_i (g_i^t - \bar{g}^t)^2 * \sum_i (g_i^s - \bar{g}^s)^2} \quad (5.2)$$

3.) weighted correlation coefficient:

$$R_{ts} = \sum_i w_i * g_i^t * g_i^s / \sqrt{\sum_i w_i * (g_i^t)^2 * \sum_i w_i * (g_i^s)^2} \quad (5.3)$$

4.) weighted cross-correlation coefficient:

$$R_{ts} = \sum_i w_i * (g_i^t - \bar{g}^t) * (g_i^s - \bar{g}^s) / \sqrt{\sum_i w_i * (g_i^t - \bar{g}^t)^2 * \sum_i w_i * (g_i^s - \bar{g}^s)^2} \quad (5.4)$$

where t means target, s means search and w is the weight coefficient.

When applying correlation matching by using VISAT images, stereo pairs of images are used. In this situation, the exposure of the two images have similar illumination and reflectivity conditions. Therefore, the Model 1.) works well. For most single targets and obvious objects, matching results give us exact conjugate points. However, when the matching is performed in the images taken from different exposure stations and different times, this model does not work as well, because the images have different grey level means. When normalizing using both sample means, it leads to use of the cross-correlation model. 2.) The cross-correlation model can reduce the influence of different exposure conditions and then results in improved matching reliability.

On the other hand, one of the main sources resulting in mismatching is the rich featured areas, such as tree leaves, grass and rough lands. If a search area partly includes this kind of area, a mismatched position may give a higher correlation coefficient factor than that of the correct one. Especially, in the case of cross-correlation, there may be a higher possibility for mismatching because of the normalization. Based on this view point and on the results of many tests, it can be concluded that the normalized cross-correlation model is more preferred to be used in different exposure cases, but the cross correlation works better in images of similar exposure conditions.

In close-range photogrammetry, due to the consideration of a wide coverage of objects and accuracy in 3D determination, the cameras have a considerably large difference in orientation. Consequently, the great difference of geometry in stereo images is produced from relatively short distances between cameras and objects. This degrades our matching results. Generally speaking, the central part of an object seen in an image

has less geometric differences with respect to that of its conjugate image, so that a weight function w_i can be applied to reduce the influence of the geometric differences between the target and the search area. Here, w_i has greater values on the central parts than on the surrounding parts.

In the implementation of Geostation, the weighted correlation and the weighted cross-correlation always give us good matching results. For a standard matching array, a square array, the assigned weight values decrease from the center to the bound. Figure 5.2 shows the weight distribution in a 5x5 array.

2	2	2	2	2
2	3	3	3	2
2	3	5	3	2
2	3	3	3	2
2	2	2	2	2

Figure 5.2 Weight distribution

The maximum value of a correlation coefficient should exceed a global threshold and, then, the corresponding matched pixel position is determined. For the purpose of improving the pixel matching accuracy, the distinct coefficient values can be considered as discrete sample values from a continuous surface. The neighborhood of R_{\max} are used to

approximate the correlation function by a continuous function (Lemmens 1988). Webber (1973) used a second order polynomial to model this function:

$$f(r) = a_0 + a_1x + a_2x^2 \quad (5.5a)$$

$$f(r) = b_0 + b_1y + b_2y^2 \quad (5.5b)$$

By assigning the partial derivatives to zero, the subpixel position of R_{\max} can be obtained as:

$$x = -a_1 / (2a_2) \quad (5.6a)$$

$$y = -b_1 / (2b_2) \quad (5.6b)$$

5.1.2 Least squares matching (LSM)

The previous section shows that the correlation matching method can give us good results for parallax measurement. In particular, it is obvious that a better matching result can be obtained if the target and search windows contain enough texture and minimal geometric and radiometric distortion. On the other hand, the geometric and radiometric distortion exist in both aerial and close-range images. For the purpose of compensating for these distortions and improving matching accuracy, the least squares matching (LSM) method was recommended by Gruen (1985) and Rosenholm (1987). LSM is the approach in which the optimum match is defined by a transformation of one array into another which minimizes the remaining grey value differences. The forward straight least squares solution minimizes the residual differences between the grey values of the pairs of corresponding points, thus representing the demanded optimum transformation which gives the optimal matching. In this method, the geometric differences can be modeled by

an affine transformation, while the radiometric differences are modeled by an additive and a multiplicative parameter. The main purpose for the single point LSM in our system is to refine the approximate matching position(s) and to get subpixel accuracy. This basic algorithm can be formulated in two dimensions with the following generalized equations:

$$g_t(x_t, y_t) + n(x_t, y_t) = g_s(x_s^0 + dx, y_s^0 + dy) \quad (5.7)$$

$$\Delta g + n = \frac{dg_s}{dx} \partial x + \frac{dg_s}{dy} \partial y \quad (5.8)$$

where

$g_t(x, y), g_s(x, y)$ grey level function of the target and the search windows

$n(x_t, y_t)$ error vector

$\frac{dg_s}{dx}, \frac{dg_s}{dy}$ gradients in x-direction and y-direction, which can be declared

as g_x, g_y .

In the implementation, both the radiometric quality and the geometrical differences are considered. Equation (5.7) can be written as:

$$g_t(x, y) = h_0 + h_1 * g_s[(a_0 + a_1x + a_2y), (b_0 + b_1x + b_2y)] \quad (5.9)$$

The observation equation can be obtained by linearizing the above equation:

$$\Delta g(x_i, y_j) + v(x_i, y_j) = g_x(x_i, y_j)da_0 + x_i g_x(x_i, y_j)da_1 + y_j g_x(x_i, y_j)da_2 + g_y(x_i, y_j)db_0 + x_i g_y(x_i, y_j)db_1 + y_j g_y(x_i, y_j)db_2 + dh_0 + g_1(x_i, y_j)dh_1 \quad (5.10)$$

where

$a_0, a_1, a_2, b_0, b_1, b_2$ unknown parameters for affine transformation

h_0, h_1 unknown coefficients for radiometric correction

$$\Delta g(x, y) = g_t(x, y) - g_s(x, y) \quad (5.11)$$

$$g_x(x, y) = [g_s(x+1, y) - g_s(x-1, y)] / 2 \quad (5.12)$$

$$g_y(x, y) = [g_s(x, y+1) - g_s(x, y-1)] / 2 \quad (5.13)$$

Through the least squares processing, the weighted mean of the right hand array (with the squares of the gradients as weights) is chosen to represent the right image area:

$$x_s = \sum_{i,j} x_{i,j} [g_x(x_i, y_j)]^2 / \sum_{i,j} [g_x(x_i, y_j)]^2 \quad (5.14a)$$

$$y_s = \sum_{i,j} y_{i,j} [g_y(x_i, y_j)]^2 / \sum_{i,j} [g_y(x_i, y_j)]^2 \quad (5.14b)$$

This procedure is iterative until the magnitude of the computed translation is smaller than a threshold. After each iteration the search window is resampled from the original image. The bilinear interpolation function is applied to calculate the new grey level values for each grid in the search window. The mathematical formula is:

$$g_{x,y} = g_{i,j}(x_{i+1} - x)(y_{j+1} - y) + g_{i+1,j}(x - x_i)(y_{j+1} - y) + g_{i,j+1}(x_{i+1} - x)(y - y_j) + g_{i+1,j+1}(x - x_i)(y - y_j) \quad (5.15)$$

where, i, j are the number of rows and columns on the image. $g_{i,j}$ is the corresponding grey level value, and $g_{x,y}$ is the interpolated grey level value at the position (x, y)

In each iteration, the design matrix and the misclosure vector are recalculated. The aerial images normally have low radiometric quality but small geometrical differences. Inversely, the close-range images have larger geometrical differences but high radiometric quality, especially in the same stereo image pair. The initial approximation is very important because both the matching reliability and convergence of least squares algorithm depend on it. If the approximated position is close to the ideal position, the iterative procedure will gradually converge and the resulting coordinates are improved. Here, we use the results of correlation matching as the initial approximation. It is sufficient that the accuracy of these approximations is within two pixels. Hence, the convergence of computation is quicker and easier, and final resulting coordinates are refined.

By using the LSM, we can make full use of grey level values in the image windows to perform pixel matching. This method improves both the matching reliability and the accuracy of parallax measurement. It also compensates for the radiometric and geometric distortions.

5.2 The reduction of searching range

Through the discussions in the above section, we can find that single point area-based image matching is an effective and precise way to perform semi-automatic point measurement, but the matching procedure needs complicated and time consuming computation because the optimal position search is carried out pixel by pixel. Normally the digital image size is around several hundred by several hundred to several thousand by several thousand. In VISAT project, for example, the image size is 512 by 480 or 640 by

480. Therefore, it is necessary to use some special methods to reduce search range and to improve the computation efficiency.

Since all the images are geo-referenced, the epipolar geometry can easily be used. The important concept of epipolar geometry to be realized in image matching is that conjugate imagery is always found along or near conjugate epipolar lines, regardless of photographic orientation. So that in the correlation matching, points situated on epipolar lines are void of gross shifts crossing the epipolar line direction. What remains along any epipolar line is the shift parallel to the epipolar line direction. Based on this characteristic, the conjugate of a point in one image can be found along the conjugate epipolar line using the automatic searching procedure. In this case, a two-dimensional search of a conjugate image point is reduced to a one-dimensional search. It improves both efficiency and reliability in matching calculation. In our real implementation, the geo-referenced information includes some errors, which result in the existence of errors in determining epipolar lines. Hence, the conjugate point search along the epipolar line is carried out on three lines on the corresponding image, i.e., the epipolar line plus its two adjacent parallel lines.

Since the correlation matching is based on two image windows and the search window moves one pixel along the epipolar line each time in the matching process, the correlation coefficient values gradually change. It may be assumed that the function of these values is smoothly changed. Through many tests, this characteristic has also been proven. According to this assumption, the matching can follow a special searching rule: The computation of the correlation coefficient is not performed at each pixel on the

epipolar line, but concentrated in some sections on the epipolar line, where larger correlation coefficient values appear. In this procedure, the rule is: If the current correlation coefficient is less than a threshold, the next pixel for the correlation computation jumps to the next point several pixels away, other than the next adjacent pixel on the epipolar line or corresponding lines. For example, in Geostation system, the experimentally determined rules are:

If the coefficient is less than 0.2, then jump +4 pixels;

If the coefficient is less than 0.4 but greater than 0.2, then jump +3 pixels;

If the coefficient is less than 0.6 but greater than 0.4, then jump +2 pixels;

Others, go to the next pixel.

In this way, the computation time is significantly reduced. On the other hand, some mathematical formulas can be simplified, so that the computation is changed to a simpler form. For example, Eq (4.1) can be changed to the following form (Lemmens, 1988):

$$R_s = \frac{\sum_i g_i^t g_i^s - (\sum_i g_i^t * \sum_i g_i^s) / n}{[(\sum_i (g_i^t)^2 - (\sum_i g_i^t)^2 / n) * (\sum_i (g_i^s)^2 - (\sum_i g_i^s)^2 / n)]^{1/2}} \quad (5.16)$$

Alternately, less costly similarity measures can be used, like the sum of the absolute values of differences (Heleva, 1972). Besides these, the matching window size is another critical factor affecting the computation time. A larger window size may result in a dramatic increase of computation time. In Geostation system, the best range of measuring object is about 2-50 m away from the cameras. When the point is selected in the left image and the

epipolar line is generated in the right image (or other corresponding images), the segment corresponding to the above effective range can be easily determined through the photogrammetric theory, so that the matching can be limited within this segment instead of on the entire epipolar line. The computation time will be reduced.

5.3 The optimal matching window size and shape

In the area-based single point image matching technique, including correlation and least squares matching, the size of the target and the search matching windows must be pre-defined. Generally, this kind of window refers to a rectangular array of grey values extracted from a corresponding image, which is symmetrical about the pixel position representing either the target point or the candidate matching point, and both windows have the same size for convenient comparison. Our system provides the user with the function for selecting matching window sizes. Through empirical investigations, we found that the window size and shape play an important role for a successful matching with regard to both reliability and accuracy. If the window is too small, the detailed information for the similarity measurement is not sufficient. However, when the window is too large, the geometric deformation and some extra features further away from the center pixel may seriously affect the matching process and degrade the matching quality. Additionally, large window size always results in huge increases of the computation time. On the other hand, the selection of different window size and shape also depends on the characteristics of the images. For example, the matching quality is more sensitive to the close-range images than to the aerial images because the former contains a larger geometric deformation and more fluctuations. Inversely, if the two images have an equal

or a very similar geometry, the error of the measured parallax will decrease with increasing window size (Rosenholm, 1987).

Since the matching quality is object and background dependable and the process is flexible, it is very difficult to determine the optimal window size. The preprocessing method was suggested by some researchers such as Dowman (1977) and Rosenholm (1987). One method is to pre-detect the fluctuation in an initially selected window on the target (normally left) image by calculating the deviation or the average absolute difference of the pixel grey level values. The more the grey level values fluctuate, the smaller must be the matching window size selected. Another method is to re-calculate the correlation coefficient between the target and the search windows with a default matching window size. If the maximum coefficient is larger than a threshold, the matching is completed. If less, the window size is reselected and the procedure is repeated. The other method is applied for LSM. In this method, a variance component model with three different components was used. The sum of the three variance components computed from empirical data by the least square method was taken as the standard to evaluate the matching quality; more detail can be seen in Rosenholm (1987).

Comparing the above methods, the second one is considered to be more reasonable and easier to implement in a practical system. It was investigated to use different sizes of windows to perform both correlation and least squares matching.

Reliability is normally connected to the number of gross errors in the matching results. We found that optimal reliability is reached at the window size of 15x15. For the correlation matching, the matching accuracy is always around one pixel if no gross error

(i.e., mismatching) occurs. In this case, the matching reliability is more important to the user. Furthermore, we found that matching results sometimes are more sensitive to the window shape. This arises from the large geometric difference between two images in the close-range application. This difference can be divided into two parts: the geometric difference of object projection on different images and the relative shift of object background on different images. Spatial objects can be divided as:

- symmetric objects, such as an electrical pole, a fire hydrant top and a manhole;
- nonsymmetric objects, such as traffic signs; and
- linear features, such as a road center line, a road edge and some building edges which extend along the depth direction.

For the first kind of symmetric object, the projections on the two images are very similar. The influence of geometric difference comes mainly from the shift of the backgrounds. By using weighted correlation matching with smaller window (e.g. 11×11), the geometric influence can be partly compensated, and better matching results can be obtained. As to the nonsymmetric object, there are some geometric limitations. If the orientation difference is too large, for example, the two focal length directions are close to crossing each other, the object projections on the two images are very different. In this situation, the matching will not be successful. If the orientation difference is small, the weighted correlation algorithm works well. In our implementation of highway information extraction, some linear objects are the main spatial features. However, this kind of objects usually yield larger geometric deformation on the images. It is almost impossible to perform correlation matching with a square window. We have, however, obtained good

results by using a rectangular window. For matching a single point on road center line or on a road edge, the optimal window size is about 3 rows by 15 columns. Figure 5.3 shows the matching results by using rectangular window.

The LSM is applied to refine the correlation matching results, so that the window size for LSM is not the critical factor for successful matching. Through the experiment, we found that the optimal window size is 9x9 for the general objects and 3x11 for the vertical linear objects.



Figure 5.3 Matching result by using 3x15 window

5.4 The application of triangulation from multi images

As we discussed in Section 3.3, the geometry of an intersection will influence the accuracy of object coordinates. The triangulation processing by using more corresponding points (more than two) can compensate for weak geometry and, consequently, improve

the accuracy. This procedure can also be implemented in the semi-automatic point measurement procedure. When an object is selected and clicked in one image, the epipolar lines corresponding to this point can be determined on the other related images. Theoretically, the correlation and then the least squares matching can be performed in these images. However, this application is constrained by large geometric differences in the images from different pairs. Since there may be more multiple correlation peaks occurring along epipolar lines in a close-range image, the system is likely to select the point with a false peak as the matched point. Besides, this procedure is very time consuming because the correlation computation needs to be carried out within a larger range on each selected images.

In fact, we can make full use of geo-referenced information in the multi-point matching procedure for the triangulation. In Geostation system, the stereo pair of images at the same exposure station is normally considered as the master image pair. In other words, we always choose two images with less geometric difference as the master image pair to start the multi-point matching procedure. For example, when a side viewing image is selected, another adjacent side viewing image may be chosen, and these two images become the master image pair. Firstly, the conventional correlation matching procedure is performed within the master image pair. When the 3D coordinates of the object are calculated through the normal space intersection processing, the approximation of image coordinates for this object can be obtained through a back projection procedure, which is based on the photogrammetric principle. From Eq (3.13), and ignoring the lens distortion, we can obtain:

$$x = -f \frac{m}{q} \quad (5.17a)$$

$$y = -f \frac{n}{q} \quad (5.17b)$$

Because the 3D coordinates of this object and all the orientation parameters are known, the m , n , q can be easily calculated. Then, the approximation of the image coordinates x , y in each selected image can be obtained. Because the geo-referenced parameters are accurate in our system, the approximation is very close to the real value (position) in the implementation, with the accuracy averaging around 3-4 pixels. Therefore, the matching procedure can be continued. The search area in each image is centered at its corresponding approximated position. Considering the error from the intersection, we can calculate the image coordinates of the perpendicular foot on the epipolar line, and take this position as the new approximate position. Then, the correlation matching works in a small neighbor range along the epipolar line. We usually search for the local maximum of correlation coefficients in 3×11 area, i.e., only 33 search windows in each image are used in the computation. In this way, the reliability of matching results from the extra images is greatly improved. When all the matching points are obtained, the triangulation is performed by applying the same technique as that mentioned in section 3.3, so that the final result can be achieved with higher accuracy.

In real cases, there are some obstructions or mismatching results which will degrade the quality of the results. For example, it is not guaranteed that the object appears in each of the selected images. Hence, a threshold is needed to check the matching quality. If a correlation coefficient is less than the threshold, the corresponding

matching result is rejected. The least squares matching can be performed only when the correlation matching is successful.

5.5 The feature-supported area based single point matching

For the application of area-based correlation matching, the reliability and the quality depend mainly on the information contained in the matching windows. More information, especially feature information included in the windows, will result in a higher reliability of matching (Schenk, 1990). From this view point, we have improved the area-based single point matching by including feature information in the windows to the matching procedure.

Reflect intensity and its changes (i.e., edges) are dependable factors for algorithms in semi-automatic object positioning. In the current approach, area based and feature-based matching are combined, and it also takes full advantage of epipolar geometry and multiple images. To obtain the edges, the improved Sobel filter is used within an area. Based on the result of edge detection, filtered line segments are extracted and represented in vector form, as shown in Figure 5.4. The vectors are merged, trimmed and reorganized to reduce trivial details.

At the present stage, when the operator selects a point in one image window, the program is required to find the rest of the conjugate points in other image windows so that the 3D coordinates of that point are obtained automatically. Considering one image window, using area based or feature based matching, there might be more than one candidate for the conjugate point. The following rules are used to eliminate the false candidates. The matched point should:

- lie on the corresponding epipolar line.
- be within a region so that triangulation gives a 2D position in front of the cameras, or even within a given distance.
- be a local maximum in terms of the correlation value.
- be on an edge if the selected point is also on an edge.
- have similar angle(s) formed by intersecting edges if the selected points are vertices.

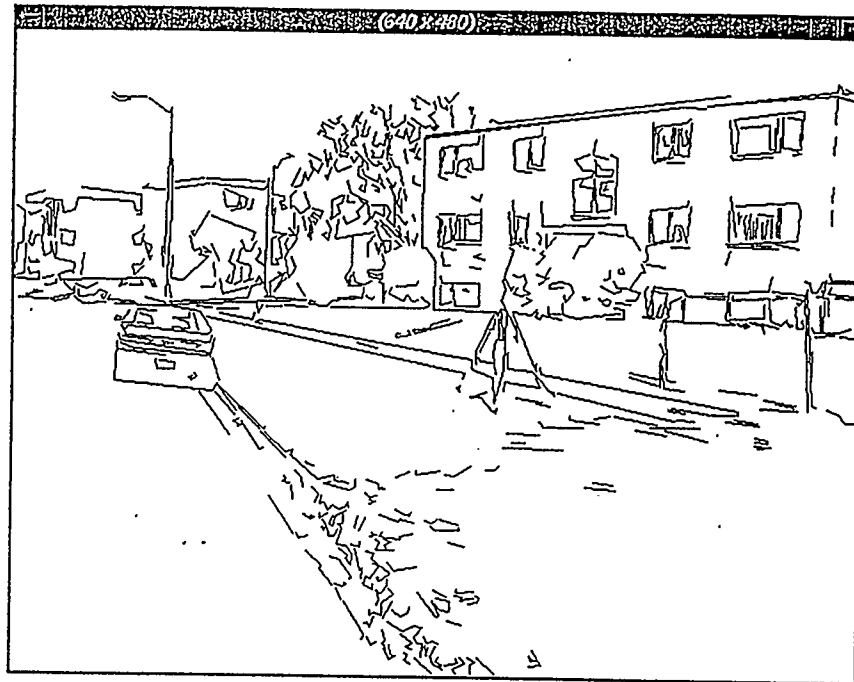


Figure 5.4 Edges in vector form

Matching can be verified by using more than two images. In the current system, the images are acquired in high volume and the redundant images are used for match verification, optimization and error estimation.

The following example shows the advantage of combining different matching methods. Figure 5.5 shows a case of mismatching caused by large geometrical differences of the same object captured in two images. When A is the selected point in the left image, the area-based matching program searches along the epipolar line in the right image and finds point B which yields the maximum correlation coefficient. However, B is not the corresponding point because it does not conform with the above matching rules. Figure 5.6 shows the correct matching result by combining different matching methods. C is the conjugate point selected in left image. The match point D should not only lie on the corresponding epipolar line, but also on the corresponding edge.

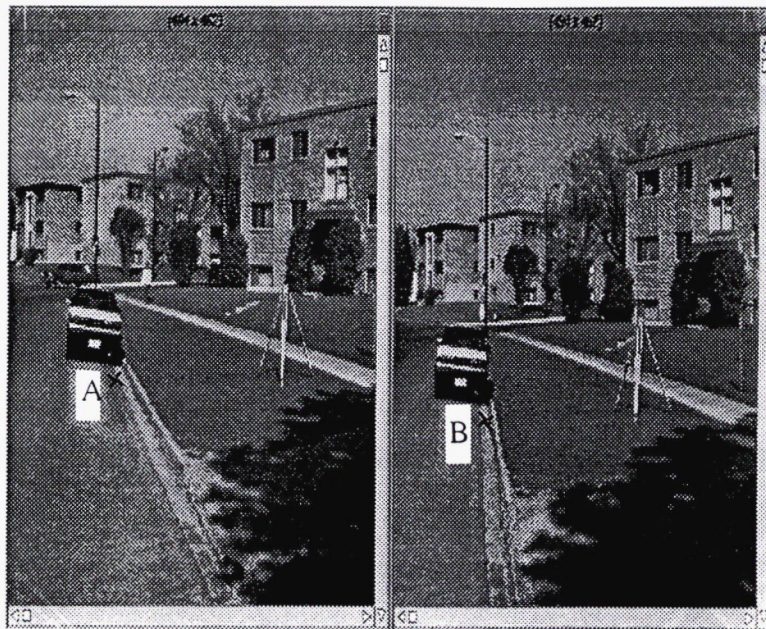


Figure 5.5 Mismatch by large geometric deformation

This method is the combination of the area-based matching and the feature-based matching. It takes feature information as the constraint, but is unlike the feature-based matching because the matching is still area-based. It not only has the advantages of traditional area-based matching, but also has the following:

- more reliable and accurate than in area-based matching
- mismatching is greatly reduced

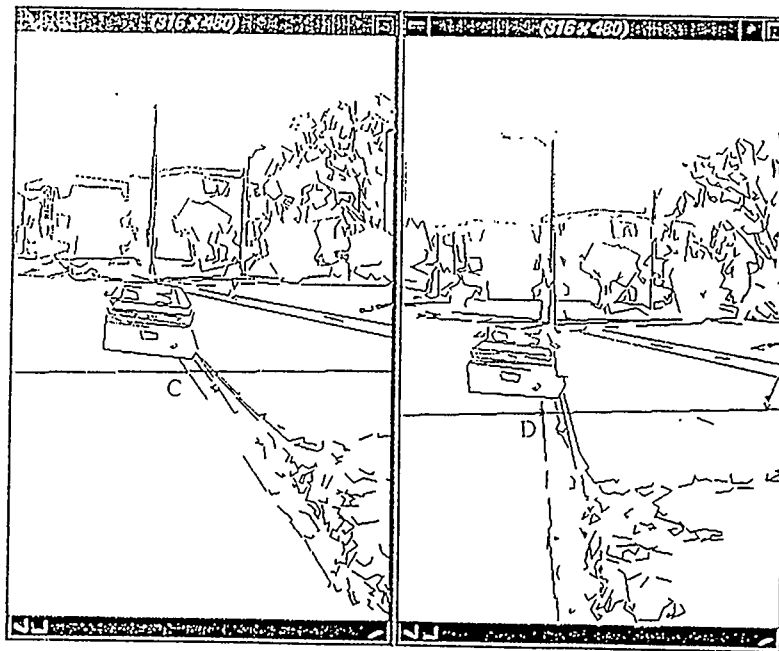


Figure 5.6 Correct match by combined conditions

Chapter 6

Road Feature Edge Extraction in 3D

In Geographic Information Systems, the collection of geo-data, including spatial and attribute, is a very important aspect. It is also one of the main purposes for developing the VISAT system. Because the data resource is very huge and complex, the automatic GIS data extraction from stereo images has become a critical topic in our system. In a highway information system, most of the objects we may survey are the transportation features. These features can be roughly grouped into two types: point and arc. For the point type objects, such as stop signs, we can apply the single point matching method discussed in Chapter 5 to perform the 3D determination. For the arc type objects, such as a road center line, feature-based image matching techniques need to be implemented to extract the 3D information of the whole object (Zlotnick, 1993). In order to realize a completely automatic system for image interpretation of topographic objects depicted in the images, their relations, the context in which they are situated, and also their functionality must be adequately modeled. The consequence, however, is that at least for practical applications, completely automatic systems are out of reach in the foreseeable future (Heipke, 1995). On the other hand, the main arc type features in our system are road center lines and road edges. The semi-automatic arc feature extraction method can be used for measuring this kind of special objects. Also, we can get the best benefit from the geo-reference sequence image data to greatly improve the efficiency of the matching procedure.

Although there are a number of feature-based image matching methods under research (Greenfeld, 1989), only a few of them can be really applied in our system. In this chapter, several feature-based semi automatic arc type objects extraction methods are discussed, implemented and compared in detail. The approach we are using is based on edge extraction, edge following and edge matching. Furthermore, the object space information is also integrated in the matching procedure to make the result more reliable.

6.1 Edge detection and noise removal

Roads are linear objects, their boundaries are normally represented by edges. An edge in the sense of image processing is a discontinuity in the two dimensional grey scale function. In applying the feature matching, the first step is the determination of the corresponding edge(s). Considering that the objects projections are almost vertical in an image, at least a starting point for the road has to be provided by the operator. Then, a relatively small window around the starting point is gradient filtered to detect edges. Through the 2D edge processing, sequential edge pixels can be determined. In Geostation, the starting point is selected by the operator with the visualization of the image (usually the left image) on the screen. This point needs to be close to the chosen linear edge. Secondly, a small rectangular area of interest is processed. The boundaries of the area are parallel to the image boundaries and the size of it is decided by the user. Normally, the area is much smaller than the whole image, so that the processing time for each area is relatively short. In practical cases, the initial search direction of the edge is defined in the vertical direction (assuming that the starting point is close to the bottom of

the image). As the edge pixels are extracted, the area and the search direction are continually updated (see Figure 6.1).



Figure 6.1 Edge detection in local areas

In the area of interest, the edge detection procedure is performed first. A gradient filter is used to pass this area to detect obvious grey value discontinuities. For the implementation, we still use the improved Sobel filter and the differential filter (see Chapter 4) to detect edges because they are very straightforward and efficient. These two filtering methods are both threshold-based, but in a real case, the result is often degraded by the random noises in the images. To suppress noise, the popular way is to use a low-pass filter, such as a medium filter or a mean filter, but it may change the geometry of objects in the image and result in positional shift of the edges. Therefore, the low-pass

filter was not used in our system. In fact, a line filter is added in the edge detection procedure. Line filter considers the grey values within a short horizontal line centered on each pixel. It can remove the bit-error noise consisting of discrete isolated pixel variations. As a simple example, when we detect the right edge of a road center line, a 1x7 line filter can be used. Only when the grey value of the center pixel is much lower than that of the first and second pixels, and also much higher than that of the sixth and seventh pixels, the center pixel is considered as a candidate of the edge pixel. That is:

if $g[3] - g[i] > \text{threshold}$, $i = 0, 1$. and

if $g[3] - g[j] < \text{threshold}$, $j = 5, 6$. then

the pixel corresponding to $g[3]$ is a candidate.

In this way, most of the bit errors (pixel values with no relation to the image scene) and unnecessary details are excluded from the edge pixels without destroying thin edge and blurring other edges. Actually, a simple template is used in the procedure. Different templates may be chosen for detecting different objects. From the results, we can see that the differential filter detects more details of the edge but the improved Sobel filter can give us a smooth edge line.

Noise is inherent in electrical circuits and is introduced from external sources. Noisy data are a combination of signal and noise, as distinguished from random bit errors that are entirely noise. The above filter method effectively works in the implementation of feature-based matching. For very noisy images, however, a box filter is applied. This method is based on an adaptive algorithm (Eliason and McEwen, 1990), which considers the values within a rectangular box centered on each pixel. It avoids the problem of

destroying thin lines and sharp corners when the threshold is properly determined. This algorithm can be summarized as follows:

$G(i, j)$ represents an integer sample of an image array G . A $(2k+1)$ by $(2m+1)$ rectangular box centered at position (i, j) is taken into processing. Then a delta function D is defined as:

$$D(i, j) = 0 \quad \text{if } G(i, j) < 0 \text{ or } G(i, j) > 255 \quad (6.1a)$$

$$D(i, j) = 1 \quad \text{if } 0 \leq G(i, j) \leq 255 \quad (6.1b)$$

The delta function is used to discriminate between valid and invalid pixel values. For each box, the following sums should be calculated. They are: the sum of valid points, $S(i, j)$; the number of valid points, $N(i, j)$; and the sum of the squares of the valid points, $SS(i, j)$.

$$S(i, j) = \sum_{m=i-K}^{i+K} \sum_{n=j-L}^{j+L} D(m, n) G(m, n) \quad (6.2)$$

$$N(i, j) = \sum_{m=i-K}^{i+K} \sum_{n=j-L}^{j+L} D(m, n) \quad (6.3)$$

$$SS(i, j) = \sum_{m=i-K}^{i+K} \sum_{n=j-L}^{j+L} [D(m, n) G(m, n)]^2 \quad (6.4)$$

Invalid pixels can result from missing image data due to data dropouts or points that have been identified as bit errors and set to zero on output by the bit-error removal algorithm (Eliason et al, 1990). The sample mean value of the box excluding invalid values $M(i, j)$ and the updated mean value excluding both invalid values and the center pixel $R(i, j)$ can be obtained as:

$$M(i, j) = S(i, j) / N(i, j) \quad (6.5)$$

$$R(i, j) = [S(i, j) - D(i, j)G(i, j)] / [N(i, j) - D(i, j)] \quad (6.6)$$

If $G(i, j)$ is identified as a bit error, it will be replaced by $R(i, j)$ because it does not contain the error value $G(i, j)$ itself. The root square of the variance for this box is:

$$\sigma(i, j) = \sqrt{[SS(i, j) / N(i, j)] - [S(i, j)^2 / N(i, j)^2]} \quad (6.7)$$

To determine whether $G(i, j)$ is a bit error, a neighborhood dependent threshold is computed from the statistics of the box. The condition for an bit error pixel is:

$$|G(i, j) - M(i, j)| > C\sigma(i, j) \text{ and } |G(i, j) - M(i, j)| > T \quad (6.8)$$

Where C is a consistent between 1 and 2, T is a pre-defined minimum value. The filter is adaptive to the local scene contrast because of the use of $\sigma(i, j)$. Low-contrast areas of a scene will have a smaller standard deviation than high-contrast areas. Therefore, a pixel value may be classified as a bit error with a relatively small value of the absolute value of the difference in Equation (6.8). In areas of high contrast, such as along boundaries, the value of $\sigma(i, j)$ increases correspondingly and so does the threshold value. It requires a larger value of the above difference for a pixel to be classified as a bit error. Therefore, the high contrast boundary pixels will neither be classified as invalid nor as bit error pixels.

In our implementation, the bit error is the main noise source. It degrades both the visual quality and the matching results. Especially when the grey level of images is enhanced, the bit error is also enhanced. The adaptive filtering method can effectively remove noise in the form of either global pre-processing or real time local processing.

6.2 Edge binarization and skeletonization

After the edge detection procedure, the detected edge pixels are recorded. In this case, all the pixels in the area of interest are divided into two classes: edge pixels and background pixels. This processing is called binarization. In order to distinguish between edge pixels and non edge pixels, a pre-defined global threshold is used. The result of the gradient filtering and the binarization are edges wider than one pixel. The required final position of the edge lies approximately in the middle of this wider edge (Heipke, 1995). To extract the center position of the edge, some algorithms should be applied. This is known as skeleton processing.

Because every pixel in the linear edge may only have two edge pixels in its closest neighborhood except at the end of the edge or an intersection position. For the purpose of skeletonizing the edge pixels, one way is to check the small area containing the candidate edge pixel to decide whether the represented pixel belongs to the skeleton. In this situation, the template method can be applied to solve the problem. The simple and popular template is based on a square array of image. Normally the size may be 3x3, 5x5, 7x7, etc. As the template size increases, the number of different combinations dramatically increases and the computation time will be unacceptable. Here, an algorithm proposed by Kreifelts (1977) is implemented. This approach works in a 3x3 array of grey level values. The center pixel is the represented pixel and it is also a detected edge pixel. Comparing different 3x3 templates (arrays), the examined array will be classified either as a skeleton pixel or a non skeleton pixel. The non skeleton pixel will be removed. Because all of the combinations of the eight neighboring pixels is 256, there may be 256 different

arrays appearing. Every 3x3 array needs to be checked to see whether it is identical to one of the template matrices which make up a subset of all the combinations. The skeletonization procedure must avoid the splitting of thick edges into parallel thin edges or the creation of holes inside thick edges. Based on these conditions, a set of template matrices have been selected in Kreifelts algorithm. When any array is matched to one of them, the corresponding center pixel is changed to a background pixel. As the procedure is performed, the thickness of the edges is gradually reduced until the width equals one pixel. Figure 6.2 shows the set of template matrices. Among them, the first one is used to suppress isolated bit error pixel.

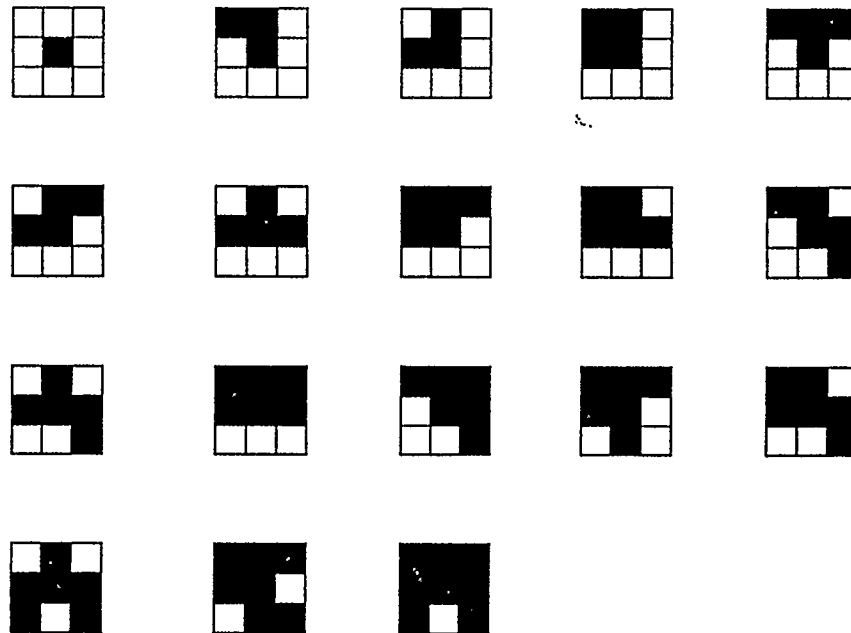


Figure 6.2 Template matrices

The other matrices filter out non skeleton pixels. Considering the symmetry, there are, in total, $1 + 4 \times 17 = 69$ template matrices in the subset. An example of this process is shown in Figure 6.3.



Figure 6.3 An example of edge filtering

Figure 6.4 shows part of the binary image after edge detection and binarization. The middle one (b) contains number 3 matrix. Therefore, the center pixel should be changed to a background pixel. The processing result is shown in the right one (c).

When the symmetric matrices are used at the same time, reduction from only one side can be avoided. This is an iteration procedure. When there is no more arrays to be matched to the template matrices, the processing is completed. The number of iterations depends on the thickness of the edges (Heipke, 1994). In the implementation of Geostation, three or four iterations are required for the processing of a whole image with

the size of 512x480. Actually, more than 98% of all the non skeleton pixels are processed in the first iteration. That means this algorithm is efficient.

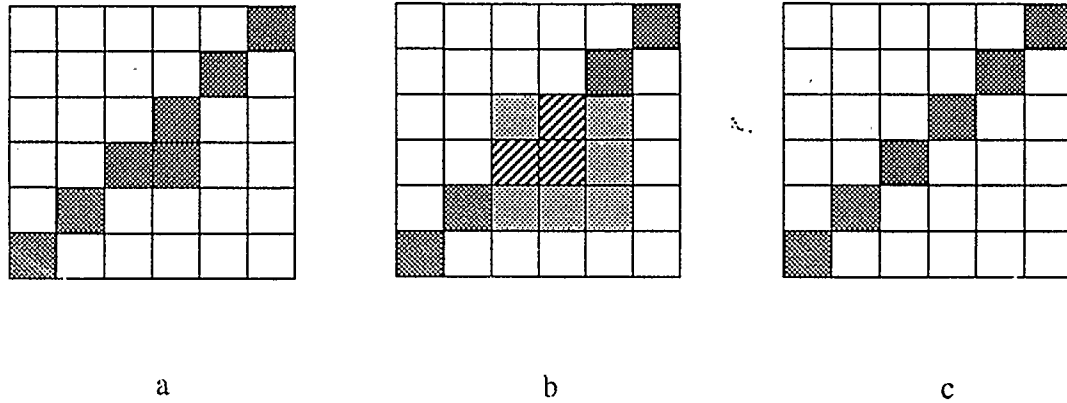


Figure 6.4 Description of the detailed skeletonization procedure

On the other hand, the computation efficiency is also a critical factor in the whole procedure. When comparing the 3x3 pixel array with the template matrices, it is necessary to pass over all the templates one by one. These templates can be considered as different conditions. Therefore, it would require a large computation time. In our system, a binary coding method of comparison is applied. Each position of the eight neighboring pixels in a 3x3 matrix has an index number from 0 to 7 (see Figure 6.5). If the pixel is an edge pixel, it is assigned as 1, otherwise assigned as 0. Therefore, we can form an 8 bit binary code from the neighboring pixels. On the other side, a one-dimensional array with 256 elements is allocated and also assigned values. The assigned values are either 1 or 0. The value 1 means that the center pixel is a skeleton pixel and 0 means not. Let's take the

number 3 matrix as an example. Its binary code can be defined as $[10000010]_2$, which corresponds to $[82]_8$ in octal and $[66]$ in digital. Therefore, we can assign:

$$\text{array}[66] = 1$$

When an array is checked, its binary code is first formed as I. If the binary number I is converted into digital format and it equals to the digital number 66, the array element value is 1. That means that the corresponding center pixel is not a skeleton pixel and should be changed to a background pixel. Because the array values have been pre-assigned in the beginning of the program, the processing is very fast compared to the checking through of all the 69 conditions (template matrices) one by one. In this way, the computation time is greatly reduced. When it works in the small area of interest, the efficiency is very high.

0	1	2
7		3
6	5	4

Figure 6.5 Ordering of neighborhood pixels

6.3 Edge following

6.3.1 Directional matrix line following

After edge skeletonization, we can obtain thin edge lines with one pixel width in the area of interest. For extracting the whole edge, we need to trace pixels and obtain

their positions. This process is known as edge following. As we mentioned in section 6.1, when starting the measuring of an edge in the semi-automatic mode, the initial point is provided by the operator. The initial direction can either be given by the operator or be determined through the automatic search procedure. In the latter case, the initial direction is actually approximated as 0 degrees (i.e., pointing upward in the image). Since the edge objects in which we are interested are only the road center lines, lane separating lines and road curbs, the directions of these objects are close to vertical. So the initial following direction as 0 degrees is reasonable. In the edge following algorithm, 11 directional matrices are pre-defined. They cover the range from -75° to $+75^\circ$ with one for every 15° . Figure 6.6 displays the directional matrices from 0° to $+75^\circ$ (By symmetry, the remainder of the matrices can easily be determined).

In each matrix, p is the current position of the edge in processing, the searching action is performed in the order of the numbered pixels until the edge pixel is found in the matrix or changes to an adjacent directional matrix and continues. Because a smaller number is normally located further away from the current position, it is selected with priority, so that the processing time in the current matrix can be reduced, and additionally small gaps in the edge may be bridged. When an edge pixel is found, it is used as the new starting point by the algorithm and the procedure is repeated. If no edge pixel is found through all the directional matrices, a major gap might exist. Then, a larger directional matrix is applied to bridge it (Heipke, 1994). In the current system, three larger matrices for each 45° are available. Figure 6.7 shows the matrices for 0° and 45° directions. In a larger searching matrix, there are more searching positions. A position which is closer to

the starting point and closer to the searching direction is assigned a smaller number. That means it will be searched earlier than positions further away.

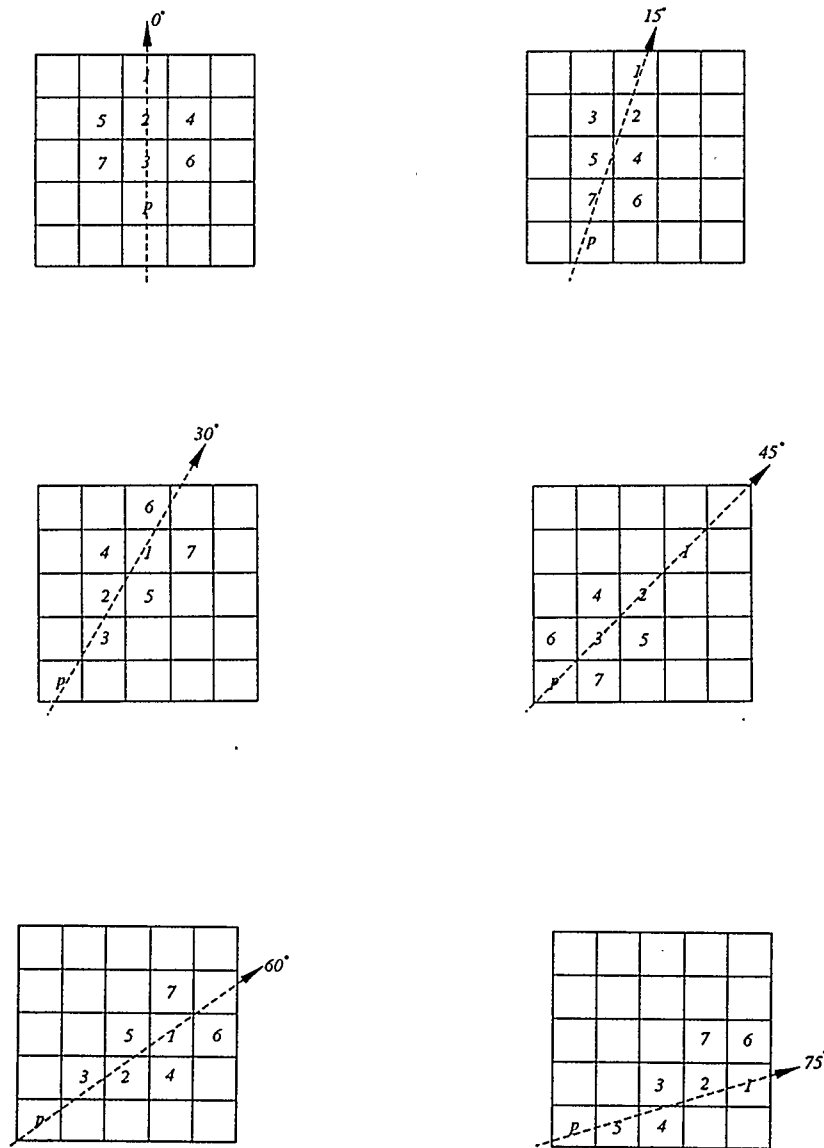


Figure 6.6 Directional matrices

The searching direction also needs to be updated continually in the process. The frequency of updating the direction depends on the characteristics of the edges. With VISAT front viewing images, the projections of road center lines and road edges are close to vertical straight lines or possibly they have a slight curve. Hence, the frequency of the edge direction updating is quite low. The initial starting point is considered to be the initial direction counting point. When the current starting point is over 20 pixels away from the initial point in vertical direction, the searching direction will be re-calculated from the two points mentioned above, and the current starting point will become the initial direction counting point. In the complete edge matching procedure, this direction is also used to determine the new area of interest.

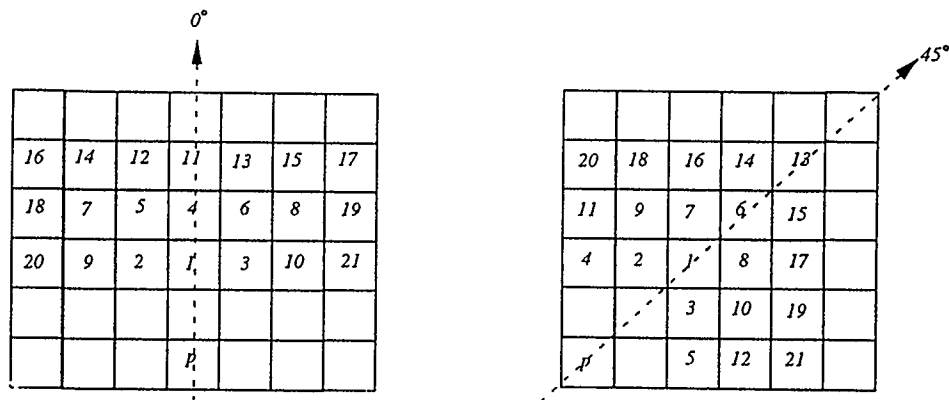


Figure 6.7 Large directional matrices

This edge following procedure will stop in the following cases: when reaching the end of the edge; meeting a large gap, which results from occlusions, shadows, or low texture and it cannot be bridged; or requiring too large a difference for the direction updating. For the first case, a dynamic matching procedure for sequential image pairs will be invoked. For the other two cases, an operator interaction is required either to provide the new initial condition to continue the procedure or to terminate it.

Although the results of edge following can provide us with all the detected edge points, only the initial direction counting points are required and recorded. By using them, edge data in a vector format will be obtained and edge matching in stereo pair can be implemented.

6.3.2 Least squares line following

The skeletonized edge following procedure is an effective and accurate method to position the desired edges in images, but the quality of the result depends on the result of edge detection. To extract an edge of a road center line, we can always get good results because the edge is very sharp and clear. Additionally, the constraints for the edge point in horizontal direction can easily be defined. However, when extracting road curbs, the skeletonized edge lines are sometimes of poor quality, since the edges in images are not clear and sharp. If this happens, we can not use the edge following algorithm discussed above to trace the edge(s). The least squares edge following algorithm works well in this case.

When the starting point for the current image is chosen, the least squares following procedure can be used. Considering that the processed edges are close to vertical lines,

when the edge is followed pixel by pixel from bottom to top, only the edge position off-set in horizontal direction for each horizontal line is the unpredictable component. In this typical case, a special least squares edge following procedure is applied. Here, only the horizontal shift parameters need to be solved.

Being different from the standard least squares single point matching procedure that we discussed in Chapter 5, the least squares edge following is performed within one image. Normally a rectangular window size with a larger extension in horizontal direction is selected for both target and searching windows. The center point of each matching window is still considered to be the representative point. The mathematical model is as follows:

$$g_t(x, y_i) = g_s(x^0, y_k) + g_x(x^0, y_k)dx \quad (6.9)$$

where

$g_t(x, y_i)$ is the grey level value in target window at (x, y_i) .

$g_s(x^0, y_k)$ is the grey level value in searching window at (x^0, y_k) .

$g_x(x^0, y_k) = [g_s(x^0 + 1, y_k) - g_s(x^0 - 1, y_k)] / 2$ is the x gradient value

dx is the off set value in horizontal direction for the matching window

In our implementation, a window size of 2x11 is applied. The offset in the vertical direction is kept as +1 pixel. For each pixel of the matching window, we can get one equation. Therefore, the above equation for the entire matching window is:

$$Adx = W \quad (6.10)$$

$$\text{where } W = g_t(x, y_i) - g_s(x^0, y_k) \quad (6.11)$$

Solved by least squares algorithm, the x offset value for the searching window is:

$$dx = -(A^T A)^{-1} A^T W \quad (6.12)$$

In the least squares processing, iterations are necessary. After each iteration, the updated x coordinate is computed as:

$$x^0 = x^0 + dx$$

$$\bar{x}^0 = (\text{integer})x^0$$

So that the resampled grey level value in searching window is:

if $dx > 0$

$$g(x^0, y_k) = g(\bar{x}^0, y_k) + \text{scale} * [g(\bar{x}^0 + 1, y_k) - g(\bar{x}^0, y_k)] \quad (6.14a)$$

or if $dx < 0$

$$g(x^0, y_k) = g(\bar{x}^0, y_k) + (1 - \text{scale}) * [g(\bar{x}^0 + 1, y_k) - g(\bar{x}^0, y_k)] \quad (6.14b)$$

where

$$\text{scale} = x^0 - \bar{x}^0 \quad (6.14c)$$

When $|dx|$ is less than a threshold value, the iteration process stops and the offset for the next edge pixel position in horizontal direction can be calculated as:

$$x_offset = \sum dx \quad (6.15)$$

In this way, the position of the center pixel for the searching window is updated and it becomes the position of the center pixel for the target window. Then, the procedure is continued. The result of the least squares edge following is a serial image coordinates

of the edge pixels from bottom to top and they will also be used in edge matching as well as the 3D edge extraction.

6.3.3 Polynomial edge fitting

In the field of feature based image matching, the edge following and the edge determination play an important role in the whole procedure. A good method of edge determination can reduce human interference in the automatic feature extraction processing and can also improve the quality of edge matching.

In our implementation, it is shown that the directional matrices edge following and the least squares edge following algorithms can usually work effectively on a road center line and a road edge when the quality of images is good, especially in the center lines or other solid lines because the edge is very clear and the image pattern is almost unique. On the other hand, road features, such as road curbs, sometimes are blurred and inconsistent. The normal edge following algorithms that we discussed above can not handle these cases. For example, some segments on the road edges are damaged, or there are too many shadows and other detailed features, so that we cannot obtain good results of edge detection and skeletonization. Furthermore, there are occlusions and intervals on the edges, or there are even dashed separating lines which need to be processed and extracted.

In order to achieve accurate and robust semi-automatic edge positioning, some special characteristics need to be considered and new approach needs to be developed. From the result of the edge detection, we found that most edge pixels can be detected by using the method we proposed. These kinds of points are normally discrete, and they are distributed along a narrow strip including the selected edge. To position the edge in the

image, a polynomial fitting algorithm is applied. On the left image, a starting point and an initial searching direction are first provided by the operator. Based on this initial condition, a strip region including the starting point and extending along the initial direction is determined. It has a width of several pixels in horizontal direction. Then, the edge detection and binarization discussed above are carried out within the area in order to arrive at the binarized edge detection result. Considering that the edges of the linear road features are normally very smooth within a limited distance from the cameras, a polynomial function of second order is assumed as the approximate format to model the chosen edge:

$$y = k_2x^2 + k_1x + k_0 \quad (6.16)$$

Where the k_0, k_1, k_2 are the unknown coefficients corresponding to the constant, the first order and the second order items respectively. Theoretically, all the detected candidate pixels of the edge in the strip should be satisfied for equation (6.16). Due to the existence of errors, a data processing procedure is required to determine the function. We used the linear least squares algorithm to solve the coefficients. For each pair of x, y coordinates of the candidate pixel, a linear equation with the unknown parameters k_0, k_1, k_2 can be formed. To the i th detected pixel (x_i, y_i) , we have:

$$(1, x_i, x_i^2)(k_0, k_1, k_2)^T - y_i = 0 \quad (6.17)$$

which can be written into the i th observation equation format:

$$v_i = a_i X + l_i \quad (6.18)$$

where

v_i	is the residual variable
a_i	is the coefficient vector of the observation equation
X	is the unknown parameters vector
l_i	is the consistent variable

Therefore, we can set up the least squares model for all the candidate edge pixels:

$$V = AX + L \quad (6.19)$$

The solution is given by:

$$X = -(A^T A)^{-1} (A^T L) \quad (6.20)$$

In this way, the parameters of this edge can be obtained. It does not need the skeletonization and vectorization processing, so the computing time is greatly reduced. Further, it can handle both continuous edges such as road center lines or interval edges such as lane separations with dashed lines. Additionally, this method can work on the edge with occlusions and/or shadows. However, there are still two questions that need to be answered. Firstly, it needs to determine whether the processing edge is either a straight edge or a curved edge. Secondly, it is necessary to remove the erroneous edge pixels in the strip area to reduce the influence of these errors.

In real implementation, a two-time processing procedure is applied. This procedure decreases the possibilities for the edge positioning error and improves the quality or accuracy of edge determination. When the starting point and the initial direction are given, the strip area is approximated as a straight strip. The center line, lane separating lines and road curbs are almost parallel, they should intersect at a vanishing

point. This vanishing point is always around the center position of the image. Furthermore, we detect road linear features from the stereo image sequence, and all detected linear features will be connected together from different image pairs. Since intervals between different exposure stations are only around 3-8 meters, we can only select a shorter range (e.g., 50 pixels) from the starting point for the strip area. Hence, a working area in the object space is always very close to the corresponding camera(s), and the quality of edge detection is, therefore, improved. We then put all of the detected edge candidate pixels into the linear least squares algorithm in order to find the coefficients k_0, k_1, k_2 in the Equation (6.16). Then we must to check the value of k_2 to see whether it is close to 0. If it is close to 0, that means that the edge is a straight line, so the second order item in the Equation (6.16) can be ignored. Otherwise the edge will be considered as a curved edge. At this time, because the coefficients of the function have been obtained, an updated strip area can be determined to repeat the procedure. By using Equation (6.19), all of the residuals can be easily calculated. If an individual residual is larger than a pre-defined threshold, the corresponding candidate edge pixel is removed. Finally, the above least squares fitting procedure is repeated. This excludes the error pixels. In this sense, the edge determination results are refined in the second round of edge determinations.

Through the above discussion, we can see that the quality of the polynomial fitting edge following method depends on the quality of the edge detection algorithm. Because we have obtained good results in edge detection in VISAT images, this edge following algorithm can successfully be used in our system.

6.4 Edge vectorization

For the directional matrices and the least squares edge following algorithms, the results consist of a sequence of edge pixels. For the purpose of edge matching and 3D edge extraction, the edge pixel chain needs to be converted into a vector format.

In our procedures, each time the algorithm stops after finding a certain number of edge pixels, each terminal point can be recorded as a vertex. These vertices will be directly used in the matching procedure. For data compression, the Douglas-Peucker-Algorithm (Douglas, et al 1973) is still popularly used in this field. It approximates the edge pixel by a polygon. The procedure starts at two end points of the whole edge in the current image. These end points are first linked by a straight line, and then the distances between each edge pixel and the straight line are calculated. If the maximum distance is larger than a pre-defined threshold, the corresponding edge pixel is considered to be a new vertex and also a polygon point. If two stereo images are used, this point might be considered as the pixel by which to calculate 3D coordinates. When two straight lines between the new vertex and the two end points are connected, the above procedure is repeated. This process will continue until all the maximum distances to corresponding straight lines lie below the threshold so that no more polygon points can be created (Heipke, 1994). Figure 6.8 shows the basic principle of this conversion.

Generally, the 2D projection of road feature edges in the VISAT images are straight lines, but sometimes curved edge lines might appear. In a vector format, the amount of data is much less than that of the edge pixel chain. Only the pairs of

coordinates for all the vertices need to be stored. In case of curved edges, it is necessary for some interpolated points. In this case, for every given length of segment, such as a length of 20 pixels, a pair of coordinates are recorded.

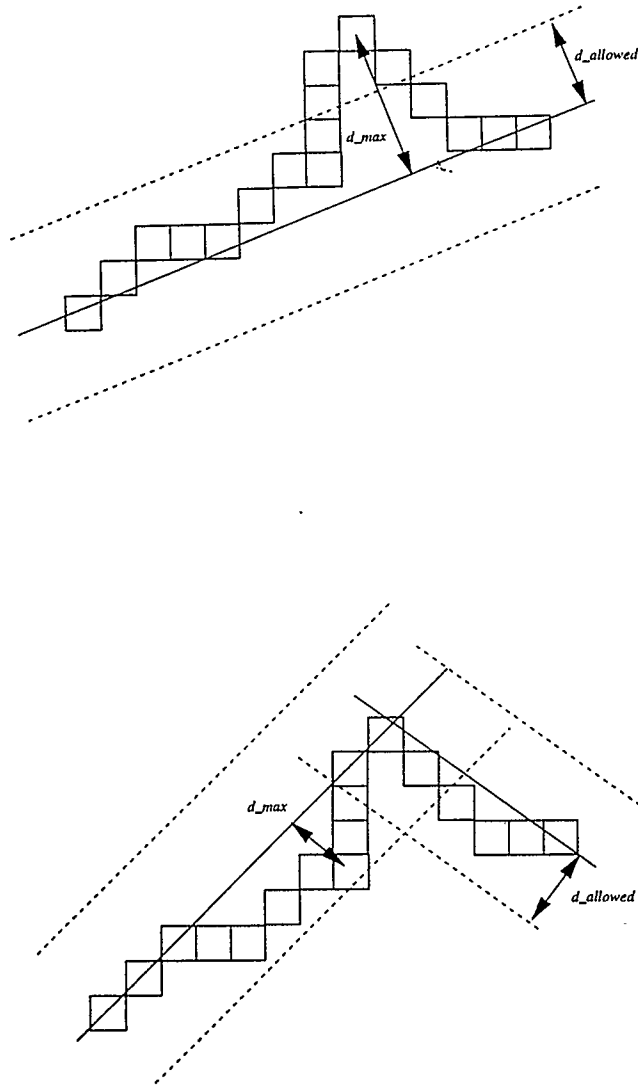


Figure 6.8 Conversion from raster to vector

6.5 Automatic extraction of 3D road feature edges

Automatic extraction of road features from images has been popular in the automation of photogrammetry for several years (Shapiro, 1987). In a highway information acquisition system, such as the VISAT system, the major objects to be dealt with are the linear road features. Therefore, it is the most important step for automating highway spatial data acquisition. Although automatic feature extraction is still a complicated task in general cases, it has been developed and implemented in Geostation by taking full advantage of geo-referencing, spatial orientation and the installation of the cameras.

In order to convert the road feature edges from an image into a three dimensional spatial domain, it is necessary to match those feature edges from the stereo image pair. That means when the feature object is selected in one image, the conjugate feature should be found in the conjugate image. Then the 3D information of this feature can be obtained through the triangulation. Based on the algorithms of the pre-processing for the edge matching discussed in the previous sections, we can easily detect, process and follow the edges in the left image. Subsequently, we will concentrate on how to locate, process and match it in the right image.

6.5.1 Matching the starting point

When starting the procedure, the starting point P of the edge is first identified in the left image. At the same time when pre-processing this edge in the left image, the conjugate starting point needs to be determined in the right image. Because of the large geometric distortion for the linear features, neither the cross correlation nor the least

squares matching algorithm works well in these cases. As an improvement, we consider the characteristics of the feature from the object space. We assume that the road surface is flat within a small area and that the whole road surface is smooth. Since the heights of the cameras above the ground corresponding to any image are known, the road surface directly below the van and nearby can be defined as a leveling flat plane with an estimated elevation. This plane is extended to include the selected starting point P which can be assumed to be close to the cameras. Then the following method is used to approximate the image coordinates of P in the right image. The elevation of the flat plane can be computed as:

$$Z_p = [(Z_1 - h_1) + (Z_2 - h_2)] / 2$$

where

Z_1, Z_2 are the elevations of left and right cameras.

h_1, h_2 are the heights of left and right cameras from the ground surface.

To obtain the approximation of the image coordinates for point P in right image, the collinearity equations:

$$x_p - x_0 = -f \frac{m_{11}(X_p - X_0) + m_{12}(Y_p - Y_0) + m_{13}(Z_p - Z_0)}{m_{31}(X_p - X_0) + m_{32}(Y_p - Y_0) + m_{33}(Z_p - Z_0)} \quad (6.22a)$$

$$y_p - y_0 = -f \frac{m_{21}(X_p - X_0) + m_{22}(Y_p - Y_0) + m_{23}(Z_p - Z_0)}{m_{31}(X_p - X_0) + m_{32}(Y_p - Y_0) + m_{33}(Z_p - Z_0)} \quad (6.22b)$$

are used. We can deduce these to:

$$A_1 x_p + B_1 y_p + C_1 = 0 \quad (6.23a)$$

$$A_2 x_p + B_2 y_p + C_2 = 0 \quad (6.23b)$$

where

$$A_1 = (x_p - x_0)m_{31} + fm_{11} \quad (6.24a)$$

$$B_1 = (x_p - x_0)m_{32} + fm_{12} \quad (6.24b)$$

$$C_1 = [(x_p - x_0)m_{33} + fm_{13}](Z_p - Z_0) - [(x_p - x_0)m_{31} + fm_{11}]X_0 - [(x_p - x_0)m_{32} + fm_{12}]Y_0 \quad (6.24c)$$

$$A_2 = (y_p - y_0)m_{31} + fm_{21} \quad (6.24d)$$

$$B_2 = (y_p - y_0)m_{32} + fm_{22} \quad (6.24e)$$

$$C_2 = [(y_p - y_0)m_{33} + fm_{23}](Z_p - Z_0) - [(y_p - y_0)m_{31} + fm_{21}]X_0 - [(y_p - y_0)m_{32} + fm_{22}]Y_0 \quad (6.24f)$$

The solution can be obtained as:

$$X_p = -(B_1 Y_p + C_1) / A_1 \quad (6.25a)$$

$$Y_p = -(A_2 C_1 - A_1 C_2) / (A_2 B_1 - A_1 B_2) \quad (6.25b)$$

The above spatial coordinates of point P are calculated from the external orientation parameters of the left image. Then, we perform the back projection onto the right image from the spatial coordinates of point P and get the image coordinates for the projection point, which are based on the external orientation parameters of the right image. The epipolar line corresponding to point P in the right image can still be generated. Therefore, we also calculate the position of perpendicular intersection point on the epipolar line for the projected image point. Since all the road features are on the surface of the road and the results of the back projection are accurate, the correlation method can still be applied along the epipolar line to refine the approximation of the back

projection point. Here, we choose a short range on the epipolar line as a searching area of central pixels for correlation matching. Normally the dimension of this area is about 15 pixels (along the epipolar line direction). The correlation procedure with a rectangular matching window size extended in horizontal direction is performed in this small searching area. Thus the conjugate point can be quickly matched. In this way, we can find the starting point for the feature matching in the right image with high reliability.

6.5.2 Matching the vertices

In the directional matrices or the least squares methods, the results of the edge following and the edge determination are sequential line segments represented by the vertices. They are stored in vector format. The performance of the edge matching can be focused on matching those vertices to obtain their 3D coordinates.

When the starting point in the right image is matched, the edge detection, the skeletonization following, and the vectorization algorithms will be performed, independently and in the same way as in the left image. The conjugate points for the selected road edge line on the right image are calculated based on the epipolar geometry. Limiting the searching range is still necessary to improve the matching efficiency. We can calculate each approximated position corresponding to each of the vertices using Equations (6.21) to (6.25). The search range can be defined with a fixed width on the epipolar line and centered at the position with the same x coordinate of the approximated position. Then, the exact position can be found by searching in the small one dimensional range. The 3D coordinates of each vertex is calculated using the pair of conjugate image points from left and right images by the photogrammetric triangulation. The 3D feature

edge extraction is then completed and the format of a linear object is sequential triples of coordinates. In the VISAT system, all the geo-reference information can be transformed into several kinds of popularly used coordinate systems, such as the standard 3TM coordinate system and the geographic coordinate system. Therefore, the results from triangulation can be directly applied in the fields of mapping, civil engineering, and GIS applications, etc.

In the implementation of the least squares edge fitting algorithm, the result of the edge determination in the left image is an analytical function for the selected edge with the starting point. To match edges more precisely, we consider interpolating some interval points on the target edge in the left image in case of curved edges. For a certain number of pixel increases in the vertical direction on the target edge, an interval point can be calculated from the analytical function, then the approximated positions of the conjugate points in the right image for the interval points are calculated using Equations (6.21) to (6.25). Based on these approximated positions, a searching area in the right image can be defined in order to be applied for the least squares edge fitting procedure which was discussed in the previous section, so that the analytical function for this edge in the right image can be determined. Then, each intersection point of the edge and each epipolar line corresponding to the interval points in the left image can be calculated in the 2D image coordinate system. Through the same triangulation procedure, the 3D edge can be extracted in the format of sequential triplets of interval points.

In the above method, the choice of the pixel number in the vertical direction between two adjacent interval points depends on the curvature of the edge. If the edge

fitting function in left image is linear, we choose a larger number, such as 30 - 40. If it is a second order function, the number 20 may be applied. Additionally, if it is necessary to approximate more accurate positions of the conjugate interval points, the elevations of these points should be updated by calculating the change of elevations along the road direction.

6.5.3 Dynamic link of matching edge between adjacent stereo image pairs

In the procedure of feature edge extraction, geo-referenced stereo image pairs are required. Normally, we work on one pair of stereo images and extract sections of the road edge or other linear features. The lengths of the extracted features are limited by:

- errors propagated from the geo-reference data errors, such as GPS positioning error, calibration error, or INS navigation error. This kind of errors are proportional to the distances between the objects and the cameras.
- error resulted from the geometry. The geometry of space intersection will get weaker with an increase of the object distance.
- pixel resolution of the digital images. The error will get larger with an increase of the object distance.

It is noted that the object distance is a critical factor and it influences the final accuracy of 3D object positioning and it should be limited within a range. Therefore, we always measure a small section of a linear feature on one pair of images. By using the sequence of image pairs, the short sections can be linked together to form the complete feature edge(s). Applied in the edge matching procedure, this link is performed

automatically and dynamically. In real implementation, we limit the image coordinate of y being larger than a fixed number when approaching this limitation, the edge matching will be switched to the next pair of images.

On the first pair of images, when an edge section has been matched and extracted, the end point of the section and its direction are known. Based on the large overlapped area between the adjacent images, we can back project the end point to the next image in the image sequence by using the collinearity equations. Because of errors, it is necessary to refine the position of the projected point on the second left image through correlation based point matching. This point is considered as the new starting point for the edge matching. As for the initial searching direction for the directional matrices edge following and the least squares fitting approaches, the initial searching direction can be calculated from the extracted edge section in the object space. In a simple way, another point on this section will be back projected to the second image, then, this point and the projected end point can determine the initial searching direction. In this procedure, the matching will be continually carried out until it reaches the end of the edge or stops where interaction is required.

We use the same method to obtain the starting point in the second right image as in the previous one. In this case, the elevation of the starting point is exactly known, so that the approximation of its position is more accurate. Figure 6.9 shows an example of edge matching, and Figure 6.10 is the results of the edge extraction projected to the Overview displaying environment.

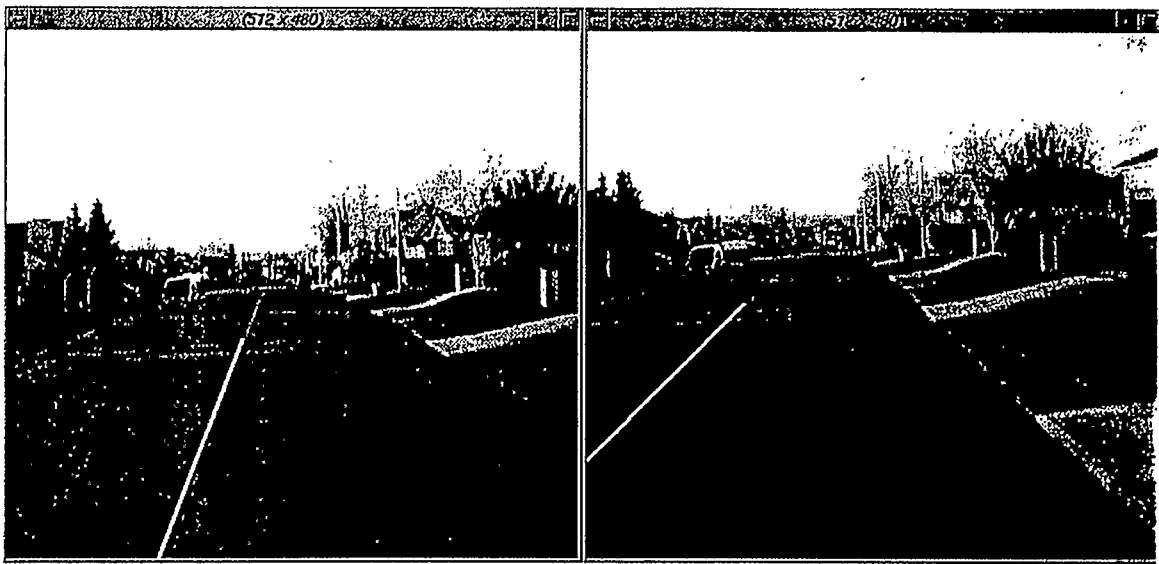


Figure 6.9 An example of edge matching

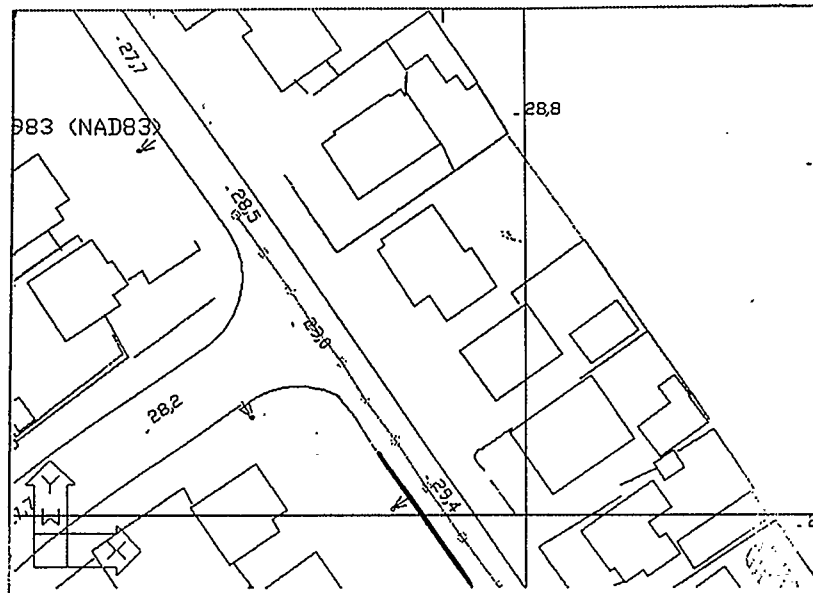


Figure 6.10 Projection on Overview

Chapter 7

Conclusions and Recommendations

7.1 Conclusions

The primary objective of this research was to investigate techniques used in improving , developing and implementing some advanced algorithms for the automation of spatial data acquisition from geo-referenced digital images. Besides this, some image processing methods, GIS data generation and linking strategies were also discussed and tested. In addition, a Sun-workstation based softcopy photogrammetric system has been developed. It can perform multi-functionalities of highway spatial data acquisition, processing and transformation with high reliability and high efficiency in a user-friendly environment.

Through the real development and implementation and from the testing results of a large number of images, the following conclusions are made:

1. The presented system is successfully implemented for fast, accurate, efficient and reliable spatial object measurements and 3D positioning in a comfortable office environment. It is intended for commercial applications. By using this system, it is expected that the surveying and data managing cost will be greatly reduced.
2. The system has shown that it has the potential to be useful for GIS data acquisition, management, and transformation from geo-referenced CCD digital images to popularly used GIS software packages, such as ArcInfo.

3. Some semi-automatic measuring methods, image matching procedures, were developed and implemented. By using these methods, the quality, the accuracy and the efficiency of object measurement have been improved.
4. When performing measurement in this system, a spatial positioning accuracy of 0.3 m can be achieved within a distance of 50 m in most cases. The base length and the image pixel resolution are the most significant factors influencing the positioning accuracy. For the currently used 2 m camera base length and 512x480 pixel image resolution, an optimal object distance range between 4 m to 30 m is recommended.
5. The combination of GPS, INS positioning, orientation techniques and CCD image acquisition strategy has been proven to be an ideal way to implement close-range photogrammetric applications. It is advantageous over the traditional method, in terms of simplification, convenience, efficiency and flexibility.
6. Image processing methods were successfully used in processing VISAT images. As a global processing method, the multi-section linear stretching grey level enhancement gave us the best results. It improved the image quality which was of poor exposure. During edge enhancement, the add-back procedure greatly enhanced the details on images, and we were provided with a clearer viewing on the displayed images. Both the grey level enhancement and edge enhancement algorithms can be applied either in global image processing or in local image processing.

7. The area-based single point matching algorithms were implemented and tested in our system. The weighted correlation matching gave us the best result. The cross correlation matching sometimes failed with rich-featured matching windows. Least squares single point matching can refine the accuracy of correlation matching results, but mostly it is not significant in comparison with the influence of other error resources. The matching window size is also a factor that influences the matching reliability. A size of 15x15 pixels is optimal for correlation matching, whereas the rectangular window size with a larger extension in horizontal direction is very appropriate for the vertical direction edge point matching.
8. Segment-based edge matching algorithms were also developed and implemented. The least squares edge following method was affected by some feature changes, such as shadow, shading, or noises. The directional matrices approach gave us better results. By using the improved Sobel filtering edge detection, template skeletonization and directional matrices edge following, we can get very accurate edge positions. However, this method cannot bridge over large gaps and occlusions, and is more sensitive to the image quality. The least squares edge fitting algorithm proved to be a good method in edge matching procedure. The advantage is that it can bridge over larger gaps or occlusions, so that it can work on both solid edge lines and dash lines. By using the two time processing procedure, it is less influenced by poor image

quality. The dynamic edge matching and linking method is a very efficient way for automatic edge extraction.

9. The Multiview subsystem and the Overview subsystem cooperate well in current status. The combination of these two systems provides us with an efficient way to access all the functions, to analyze the displayed images and to make an on-line display of the results, as well as to navigate the image system.
10. Some issues still remain to be investigated or to be solved, such as the reliability of image matching, both area-based or feature-based. For practical applications, completely automatic systems are out of reach in the foreseeable future and we cannot get rid of the interactive operation by human beings now.

7.2 Recommendations

For the purpose of improving the reliability of automatic feature extraction and the efficiency of softcopy photogrammetric system, several recommendations could be made:

1. For the semi-automatic single point object extraction, some object space information should be taken into consideration. Image pattern recognition techniques and knowledge-based object matching strategies might be used to identify the measured objects so as to improve the reliability of the matching procedures. One of the practical ways is to set up more constraint conditions to reduce matching candidates.
2. Color images might be introduced to our system. The main advantages are a better viewing result to users and improved quality of image matching.

Because the color resource comes from the three original colors, the point matching can be performed in each of them. A combination of the matching results in the three color components would give a more reasonable choice for the matching. On the other hand, the edge detection results would be much better by using the color images.

3. To improve the final positioning accuracy of objects, we can either increase the base length or increase the image resolution. Multi image point intersection would also improve accuracies. Furthermore, sub-pixel level measuring should be achieved to reduce measurement errors.
4. Better communication methods should replace the current one. Now, a special file is used as the buffer for the communication between Multiview and Overview subsystems. This method can not handle the large amount data exchange. So some new techniques, such as the SOCKET technique, should become the main communication method.
5. In edge detection processing, more templates for the small cross profile should be investigated. With the assistance of additional templates, edge pixels will be detected with fewer noises.
6. Some image clustering and classification techniques can be developed and implemented to extract well-defined objects and to manage the relative information.

7. A database should be set up to store and manage a huge number of surveying results and also to link with GIS packages.
8. The stability of the whole system should be considered as a critical point and be improved continually to meet the requirement of commercial applications.

REFERENCES

- Ackermann, F., 1984. Digital Image Correlation: Performance and Potential Application in Photogrammetry. *Photogrammetry Record*, 11(64): 429-439.
- Argialas, D.P. and C.A. Harlow, 1990. Computational Image Interpretation Models: An Overview and a Perspective. *Photogrammetric Engineering and Remote Sensing*, 56(6): 871-886.
- Beyer, H.A., 1992. Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System. The Dissertation of Ph.D. Institut fur Geodasie und Photogrammetrie, Eidgenossischen Technischen Hochschule-Honggerberg, Zurich.
- Case, J., 1982. The Digital Stereo Comparator/Compiler (DSCC), *International Archives of Photogrammetry and Remote Sensing*, 24(2): 23-29
- Chapman, M.A., 1992. Stereo Digital Image Processing. Lecture Notes, Department of Geomatics Engineering, The University of Calgary.
- Curry, S., S. Baumrind and J.M. Anderson, 1987. A Technique for Anatomic Feature Extraction and Tracking on sequential Digital X-Ray Images. *Photogrammetria*, 42: 126-136.
- Douglas, D.H. and T.K. Peucker, 1973. Algorithms for the Reductin of the Number of Points Required to Represent a Digitized Line of Its Caricature. *The Canadian Cartographer*, 10(2): 112-122.

- Dowman, I.J. and A. Haggag, 1977. Digital image Correlation Along Epipolar line, Proceeding of the International Symposium on Image Processing, Interactions with Photogrammetry and Remote Sensing, pp. 47-49.
- Dowman, I.J., 1992. Overview of European Developments in Digital Photogrammetric Workstations. Photogrammetric Engineering and Remote Sensing, 58(1): 51-56.
- El-Shiemy, N. and K.P. Schwarz, 1993. Kinematic Positioning in Three Dimension Using CCD Technology, Proceedings of VNIS'93, pp. 102-105.
- Eliason, E.M. and A.S. McEwen, 1990. Adaptive Box Filters for Removal of Random Noise from Digital Images. Photogrammetric Engineering and Remote Sensing, 56(4): 453-458.
- Frederic H.J. and L. Martin, 1991. The AutoCAD Database Book. Ventana Press.
- Gagnon, P.A., J.P. Agnard, C. Nolette, and M. Boulianne, 1990. A Microcomputer-Based General Photogrammetric System. Photogrammetric Engineering & Remote Sensing, 56(5): 623-625.
- Gonzalez, R.C. and P. Wintz, 1987. Digital Image Processing, Addison-Wesley Publishing Company.
- Granshaw, S.I., 1980. Bundle Adjustment Methods in Engineering Photogrammetry. Photogrammetric Record, 10(56): 181-207.
- Greenfeld, J.S., 1989. Experiments with Edge-Based Stereo Matching. Photogrammetric Engineering and Remote Sensing, 55(12): 1771-1777.

- Gruen, A.W. 1985. Adaptive Least Squares Correlation: A Powerful Image Matching Technique, South Africa Journal of Photogrammetry, Remote Sensing and Cartography. 14(3): 175-187.
- Heipke, C., A. Englisch, T. Speer, S. Stier, and R. Kutka, 1994. Semi-automatic Extraction of Roads From Aerial Images. Proceedings of International Society of Photogrammetry and Remote Sensing, Commission III Symposium, Munich, pp. 353-360.
- Heipke, C., 1995. State-of-the-Art of Digital Photogrammetric Workstations for Topographic Applications. Photogrammetric Engineering and Remote Sensing, 61(1): 49-56.
- Helava, U.V. 1976. Digital Correlation in Photogrammetric Instruments, Bendix Research Laboratories, Southfield, Michigan 48076, U.S.A. :26-26.
- Kreifelts, T., 1977. Skelettierung und Linienverfolgung in raster-digitalisierten Linienstrukturen, in Nagel, H.-H. (Ed.), Digitale Bildverarbeitung, pp. 223-231.
- Lemmens, M, 1988. A Survey on Stereo Matching Techniques. International Archives of Photogrammetry and Remote Sensing. Commission V, 27(B8): 11.
- Li, R., K.P., Schwarz, M.A. Chapman, and M. Gravel, 1994. Integrate GPS, INS and CCD Cameras for Rapid GIS Data Acquisition. GIS World, 7(4): 40-43.
- Li, R., M.A. Chapman, L. Qian, and Y. Xin, 1994. Prototype Development of Geostation System and The Pilot Project: Project Report. Department of Geomatics Engineering, The University of Calgary.

- Lu, Y. and Z. Zhang, 1988. Fast Implementation for Generating Epipolar Line Images with One-dimensional Resampling. No. 145 Report, Institut fur Geodasie und Photogrammetrie, Eidgenossische Technische Hochschule Zurich.
- Marr, D. and E. Hildreth, 1980. Theory of Edge Detection, Proceedings of Remote Sensing Society of London, 207: 187-217.
- Masry, S.E., 1974. Digital Correlation Principles, Photogrammetric Engineering, 40(3): 303-308.
- Mori, N., S. Masuda, and S. Marai, 1992. Development of a PC-Based Digital Photogrammetric System. International Archives of Photogrammetry, 28(2): 319-322.
- Novak, K. and S.L. Sperry, 1992. Integrating Digital Photogrammetry and GIS. Proceedings of American Society of Photogrammetry and Remote Sensing, pp. 57-64.
- Pan, J. and J.O. Domingue, 1990. Differential Design and Performance for Edge Sharpening. Photogrammetric Engineering and Remote Sensing, 56(5): 573-578.
- Qian, L. Y. Xin, R. Li, and M.A. Chapman, 1994. Reconstruction of Geometric Elements in GIS Databases Using GPS/INS Controlled CCD Images. Proceedings of GIS'94, (2): 395-398.
- Qian, L., Y. Xin, R. Li, and M. A. Chapman, 1995. Graphic User Interface Design for Softcopy Photogrammetry, Proceedings of GIS'95, (1), 752-755.

- Rogers, J.A. and R.M. Bennett, 1988. An AutoCAD-Based Mapping System for Encoded Stereoplotters. *Photogrammetric Engineering and Remote Sensing*, 54(3): 353-355.
- Rosenholm, D., 1987. Empirical Investigation of Optimal Window Size Using the Least Squares Image Matching Method. *Photogrammetria*, 42: 113-125.
- Rosenholm, R., 1987. Digital Image Correlation with an Analytical Plotter, *Photogrammetria*, 40: 9-19.
- Sarjakoski, T., 1981. Concept of A Completely Digital Stereo Plotter, *The Photogrammetric Journal of Finland*, 8(2): 95-100.
- Schenk, T. and O.Zilberstein, 1990. Experiments with a Rule-Based System for Interpreting Linear Map Features. *Photogrammetric Engineering and Remote Sensing*, 56(6): 911-917.
- Schowengerdt, R.A. and L. Wang, 1989. A General Purpose Expert System for Image Processing. *Photogrammetric Engineering and Remote Sensing*, 55(9): 1277-1284.
- Schwarz, K.P., H. Martell, N. El-Shiemy, R. Li, M.A. Chapman, and D. Cosandier, 1993. VISAT -- A Mobile Highway Survey System of High Accuracy. *Proceedings of VNIS'93*, pp. 476-481.
- Shapiro, L.G. and R.M. Haralick, 1987. Relatinal Matching. *Applied Optics*, 26(10): 1845-1851.

- Shortis, M.R., 1988. Precision Evaluations of Digital Imagery for Close-Range Photogrammetric Applications. *Photogrammetric Engineering and Remote Sensing*, 54(10): 1395-1401.
- Sizlet, C.D., G.Y.G. Lee, and L.J. Ladner, 1992. Implementation of Softcopy Photogrammetric Workstations at the U.S. Geological Survey. *Photogrammetric Engineering and Remote Sensing*, 58(1): 57-63.
- Stevens, W.R., 1990. *Unix Network Programming*, Prentice-Hall Inc.: 87-170.
- Usery, E.L. and P. Altheide, 1988. Knowledge-Based GIS Techniques Applied to Geological Engineering. *Photogrammetric Engineering and Remote Sensing*, 54(11): 1623-1628.
- Webber, W.F., 1973. Techniques for Image Correlation, *Proceedings of IEEE Conference on Machine Processing of Remotely Sensed Data*, 16-18: 1B_1-1B_7.
- Zlotnick, A. and P.D. Carnine, 1993. Finding Roads in Aerial images, *Computer Vision, Graphics, Image Processing: Image Understanding*, 57(2): 243-260.